Testing for Leverage Effect in Financial Returns

Christophe Chorro, Dominique Guegan, Florian Ielpo, Hanjarivo Lalaharison

UNIVERSITÉ PARIS 1 – PANTHÉON - SORBONNE –

Documents de Travail du
Centre d'Economie de la Sorbonne

CES Working Papers

Testing for Leverage Effect in Financial Returns

Christophe CHORRO, Dominique GUEGAN,
Florian IELPO, Hanjarivo LALAHARISON

2014.22

Maison des Sciences Économiques, 106-112 boulevard de L'Hôpital, 75647 Paris Cedex 13
http://centredeconomiesorbonne.univ-paris1.fr/
ISSN : 1955-611X

Testing for Leverage Effect in Financial Returns *

Christophe Chorro† Dominique Guégan‡ Florian Ielpo§ Hanjarivo Lalaharison¶

February 18, 2014

Abstract

This article questions the empirical usefulness of leverage effects to describe the dynamics of equity returns. Using a recursive estimation scheme that accurately disentangles the asymmetry coming from the conditional distribution of returns and the asymmetry that is related to the past return to volatility component in GARCH models, we test for the statistical significance of the latter. Relying on both in and out of sample tests we consistently find a weak contribution of leverage effect over the past 25 years of S&P 500 returns, casting light on the importance of the conditional distribution in time series models.

Keywords: Maximum likelihood method, related-GARCH process, Recursive estimation method, Mixture of Gaussian distributions, Generalized hyperbolic distributions, S&P 500, Forecast, Leverage effect.

**Acknowledgements:* We are grateful to Giovanni Barone-Adesi, Juan Pablo Ortéga and Benoit Sévi for useful remarks. Any errors remain ours.

†Corresponding author. Centre d'Economie de la Sorbonne, University Paris 1 Panthéon-Sorbonne. Email: cchorro@univ-paris1.fr.

‡Paris School of Economics and Centre d'Economie de la Sorbonne, University Paris 1 Panthéon-Sorbonne, Maison des Sciences Economiques, 106 bd de l'Hôpital, 75013 Paris, France. Email: dguegan@univ-paris1.fr. Tel: +33 1 44 07 82 98. Fax: +33 1 44 07 83 01.

§Lombard Odier Darier Hentsch & Cie, 11 rue de la corraterie 1204 Genève, Suisse. Email: f.ielpo@lombardodier.com.

¶Centre d'Economie de la Sorbonne, University Paris 1 Panthéon-Sorbonne. Email: lalaharisonh@gmail.com.

1 Introduction

An accurate description of the dynamics of stock index returns requires a mix between a time varying volatility structure and an asymmetric and fat-tailed distribution. This is the conclusion of a long stream of research papers published over the past decades. The time varying volatility structure accounts for the potential changes in the level of risk in financial markets when the conditional distribution deals with the rare occurrence of extreme events, usually seen as jumps. Such an approach turned out to be both successful when applied in continuous (see e.g. Bates (2000)'s extension of Heston (1993)'s work) and in discrete time models (see e.g. Christoffersen et al. (2006, 2010), Badescu et al. (2008), and Chorro et al. (2010, 2012) for references in GARCH settings). The estimation of such models remains however a complex matter as components in the volatility structure and in the conditional distribution can impact the model implied density in a very similar way. This is the case of the asymmetric past return to volatility feedback effect in discrete time models usually refereed to as "leverage effect": its impact over the goodness of fit of a time series model can be compared to that of the asymmetry of the conditional distribution. By using a recursive estimation scheme that consistently disentangles both effects, this article questions the usefulness of leverage effect. Our estimates indicate that its distributional usefulness is largely below what has been considered previously, the conditional distribution playing somewhat a more crucial role for financial time series modeling.

Leverage effect consists of an asymmetric reaction of volatility to past returns, volatility rising more rapidly when returns are negative than positive (see Aydemir et al. (2006)). This stylized fact first highlighted in Black (1976) and Christie (1982) is usually explained in two ways: first, an increase in volatility should coincide with a higher expected return – as prices drop – when the price of risk is constant. Second, a drop in the price of a stock increases the underlying company's financial leverage and the volatility of its stock increases as a response to this increase in the firm's risk. Nevertheless, the lack of consensus around the rationale to leverage effect is important as illustrated by the various contributions of Schwert (1989), Campbell and Hentschel (1992), Duffee (1995), Bekaert and Wu (2000), Figlewski and Wang (2000), Wu (2001) and Aydemir et al. (2006). Regardless of the conclusions raised in each of these articles regarding the economic origin of the phenomenon, they unanimously acknowledged the empirical existence of leverage effect in individual stocks. Now, when aggregating these firm-specific phenomenon into equity indices, we are unsure that either the volatility feedback or financial arguments hold anymore as the idiosyncratic risk is diversified anyway. Still, a significant number of articles maintain an asymmetric component in their modeling of volatility usually in order to generate time varying negative skewness. Examples of these contributions are Poon and Granger (2003), Awartani and Corradi (2005), Corsi and Reno (2012) and Bandi and Reno (2012). In the continuous time literature, a similar case can be made out of various extensions of the Heston (1993)'s model where the leverage effect component is used to generate asymmetric implied volatility surfaces as extracted from option prices. All in all, leverage effect probably turned out to be a useful component improving continuous and discrete time financial models' ability to fit the observed distribution of returns (Awartani and Corradi (2005)) or to match the observed price of options on stock indices (Christoffersen and Jacobs (2004)) especially when model's conditional distribution is Gaussian and thus do not incorporate any asymmetry or fat tails. Alternatively, a properly selected conditional distribution can provide an interesting goodness-of-fit of the returns' distribution as well (see e.g. Curto et al. (2009)), somewhat in a very comparable way to that of the leverage effect.

Thus, the theoretical economic reliability of leverage effect in the dynamics of returns on indices and the possibility to use alternative flexible distributions cast doubts on the necessity of a leverage component in conditional volatility structures. This article aims to provide a reliable measure of it using 25 years of S&P 500 returns. Following Badescu et al. (2008, 2011) and Chorro et al. (2010, 2012) we combine two classical asymmetric GARCH specifications used in the financial literature, namely, the Exponential GARCH (EGARCH) introduced by Nelson (1991) and the Asymmetric Power ARCH model (APARCH) of Ding et al. (1993), with two families of conditional distributions that are able to generate various levels of skewness and kurtosis: the Generalized Hyperbolic distribution as introduced by Barndorff-Nielsen (1977) and the mixture of two Gaussian distributions (see e.g. Behboodian (1970)). By Combining these two components we want to be able to disentangle the part of the skewness effectively coming from the leverage effect and the part coming from the left tail of the conditional distribution.

Given the potential overlap of both components in terms of goodness of fit, the estimation strategy used will be key: we compare three of them through Monte Carlo experiments in order to gauge their respective ability to estimate consistently both components without mixing them. We compare (1) maximum likelihood estimates, (2) quasi-maximum likelihood estimates and (3) estimates obtained from a recursive estimation scheme introduced in this article. Our empirical experiments suggest that the recursive estimation method is the one providing the most accurate estimates. Based on this conclusion, we then rely on this methodology in order to test the statistical significance of leverage in the dynamics of the S&P 500's returns. Our conclusions unfold as follows: first, we find that leverage effects only account for 10 to 20 % of the returns' total skewness, the rest of it being driven by the asymmetry of the conditional distribution. Second, by performing in-sample Hansen (1992)'s test, we consistently accept the hypothesis that the parameter driving the leverage effect in the asymmetric GARCH models can be set to zero. Finally, we consistently find that leverage effects do not statistically improve the out-of-sample forecast accuracy of our time series models. All results are stable when splitting the sample to check for the robustness of our findings. Hence, our results indicate that modeling the S&P 500's returns through an asymmetric GARCH model is not statistically relevant as long as the conditional distribution is flexible enough to capture the behavior of returns' distribution's tails.

The article is organized as follows. Section 2 presents the GARCH-type models used in this paper and studies related estimation methodologies. In particular, Monte Carlo experiments are performed in order to compare their finite sample properties. Section 3 details the tests used to assess the significance of leverage effect along with their results. Section 4 concludes.

2 Models and estimation methodologies

In this section we briefly present the classical asymmetric and non Gaussian GARCH-type models that will be used in the empirical part of this article. We also tackle the question of estimation challenges: a new recursive estimation methodology for non linear time series is presented and its finite sample properties are compared via Monte Carlo experiments, to classical competing methods.

2.1 Models

We consider a pure stationary and non Gaussian GARCH(1,1)-type model: $\forall t \in \{1, ..., T\}$,

$$Y_t = \sqrt{h_t} z_t, \quad z_0 = x \in \mathbb{R}, \tag{1}$$

$$h_t = F_{\theta^V}(z_{t-1}, h_{t-1}) \tag{2}$$

where the $(z_t)_{t \in \{1,...,T\}}$ are independent and identically distributed real random variables defined on a complete probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and where $F_{\theta^V} : \mathbb{R}^2 \to \mathbb{R}_+$ is compatible with realistic GARCH(1,1)-type volatility models indexed by a set of parameters θ^V. We define the associated information filtration by $(\mathcal{F}_t)_{t \in \{0,...,T\}}$ where $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $(\mathcal{F}_t = \sigma(z_u; 1 \leq u \leq t))_{t \in \{1,...,T\}}$. We will suppose from now on that h_0 is a constant in such a way that the information filtration is also generated by the (Y_t). Concerning the innovations $(z_t)_{t \in \{1,...,T\}}$ we additionally assume that their distributions lie in a parametric class of Lebesgue densities with mean 0 and variance 1 (d_{θ^D}) driven by a set of parameters θ^D.

In the empirical part, the volatility structures we use are the Nelson (1991)'s EGARCH model

$$\log(h_t) = a_0 + a_1 |z_{t-1}| + \gamma z_{t-1} + b_1 \log(h_{t-1}), \tag{3}$$

and the Ding et al. (1993)'s APARCH model

$$h_t^\delta = a_0 + a_1 \left(|z_{t-1}| - \gamma z_{t-1}\right)^\delta + b_1 h_{t-1}^\delta. \tag{4}$$

These models are chosen for their ability to generate both time varying volatility and leverage effect. Moreover, it is now well documented that time-varying volatility alone has difficulty explaining the scale of the skewness and kurtosis usually found in assets' returns (see e.g. Bollerslev (1987), Nelson (1991) and Terasvirta and Zhao (2011)). In particular, many authors paid attention to the importance of conditional skewness for option pricing in discrete time setting (Christoffersen et al. (2006), Christoffersen et al. (2010), Chorro et al. (2012) and Guégan et al. (2013)). This implies that the standardized innovations are not normal and that a more general specification for the conditional distribution is needed in order to capture accurately the degree of tail fatness and asymmetry in returns. The conditional distributions that we retained are the Generalized Hyperbolic distribution (GH hereafter) and the Mixture of Gaussian distributions (MN hereafter). These two distributions are interesting candidates as they encompass both the fat tails and the asymmetry of the conditional distribution of financial returns. The GH distribution has been introduced by Barndorff-Nielsen (1977) and applied with success to finance in Eberlein and Prause (2002), Badescu et al. (2011) and Chorro et al. (2010, 2012): For $(\lambda, \alpha, \beta, \delta, \mu) \in \mathbb{R}^5$ with $\delta > 0$ and $\alpha >| \beta |> 0$, it is defined by the following density function

$$f(x) = \frac{(\sqrt{\alpha^2 - \beta^2}/\delta)^\lambda}{\sqrt{2\pi} K_\lambda(\delta\sqrt{\alpha^2 - \beta^2})} e^{\beta(x-\mu)} \frac{K_{\lambda-1/2}\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right)}{\left(\sqrt{\delta^2 + (x-\mu)^2}/\alpha\right)^{1/2-\lambda}} \tag{5}$$

where $K(.)$ is the modified Bessel function of the third kind. The MN distribution has been applied to financial returns modeling in Kon (1984), Akgiray and Booth (1987), Tucker and Pond (1998) and Alexander and Lazar (2006). We focus on the mixture of two Gaussian distributions having the density

$$f(x) = \phi f(x, \mu_1, \sigma_1) + (1 - \phi) f(x, \mu_2, \sigma_2), \tag{6}$$

where $(\phi, \mu_1, \mu_2, \sigma_1, \sigma_2) \in [0, 1] \times \mathbb{R}^2 \times (\mathbb{R}^*)^2$ and where $f(., \mu_i, \sigma_i)$ is the density of a Gaussian random variable with expectation μ_i and standard deviation σ_i. In such a case, the number of parameters driving the distribution is five, that is the number of parameters characterizing the GH one, making the two distributions highly comparable and flexible as well.

From the theoretical work of Bai et al. (2003) that provide the unconditional kurtosis of a GARCH process when innovations are conditionally not Gaussian, it is well known that volatility and distribution parameters can have a similar impact on the returns' third and fourth order moments. The fact that both the variance structure and the conditional distribution can generate asymmetry and fat tails make the parameter estimation difficult through maximum likelihood as the goodness of fit of the model can be equally improved by both sets of parameters. In the next section we present three competing estimation methods that can be used to estimate such non-Gaussian GARCH models, and evaluate their respective ability to disentangle the contribution of each of these models' components.

2.2 Estimation methodologies

Starting from an observed data set $(y_1, ..., y_T)$, we now present three different estimation strategies to estimate the vectorial parameter $\theta = (\theta^V, \theta^D)$. We distinguish a direct method (estimating all the parameters at the same time) based on the classical Maximum Likelihood (ML) estimator, a sequential method using two steps (where the volatility and the distribution parameters are estimated separately) based on the so-called Quasi Maximum Likelihood (QML) estimator and a new empirical recursive method (REC) that starts from the QML estimate and that iteratively maximizes the likelihood over either the volatility's or the distributions' parameters.

When $(y_1, ..., y_T)$ is a realization of model (1)-(2) the set of parameters $\theta = (\theta^V, \theta^D)$ can be estimate using a conditional version of the classical ML estimation. In this setting, the conditional log-likelihood based on the observations $(y_1, ..., y_T)$ is

$$L_T(y_1, ..., y_T \mid \theta) = \sum_{t=1}^{T} l_t(y_t \mid y_1, ..., y_{t-1}, \theta) \tag{7}$$

where

$$l_t(y_t \mid y_1, ..., y_{t-1}, \theta) = -\frac{log(h_t)}{2} + log\left[d_{\theta^D}\left(\frac{y_t}{\sqrt{h_t}}\right)\right]. \tag{8}$$

The ML estimator is defined by $\hat{\theta}_T = argmax_{\theta\in\Theta} L_T(y_1, ..., y_T \mid \theta)$. We refer the reader to Straumann (2005) Chap.6 for general conditions for the consistency and the asymptotic Normality of this estimator. However and despite its efficiency this method is very sensitive to the misspecifications of the innovations' density. Newey and Steigerwald (1997) were the first explaining why the density used for the conditional distribution of the time series model has to be rich enough to avoid inconsistency problems. From a numerical point of view this leads to optimization issues that grow with the number of parameters, such as local maxima.

Since it overcomes the misspecification dilemma, the QML estimator is probably the most famous estimation strategy for conditionally heteroscedastic time series. The likelihood (7) is written as if the distribution of the z_t were a Gaussian $\mathcal{N}(0,1)$:

$$L_T(y_1,...,y_T \mid \theta^V) = \sum_{t=1}^{T} -\frac{1}{2}\left(log(2\pi) + log(h_t) + \frac{(y_t)^2}{h_t}\right) \tag{9}$$

and the QML estimator $\widetilde{\theta_T^V}$ is the argmax of (9). A remarkable feature is that this Gaussian assumption is not necessary in general to ensure good asymptotic properties of $\tilde{\theta_T^V}$. In the case of pure GARCH processes very mild conditions are necessary, as illustrated in Francq and Zakoian (2010) Chap.7. However for more general GARCH specifications only recent and partial theoretical results are available (see Straumann (2005) Chap.5 for the pure AGARCH case, Meitz and Saikkonen (2011) for nonlinear AR.GARCH models, Wintenberger (2012) for the pure EGARCH(1,1) model and Francq et al. (2012) for the pure log-GARCH process). In a second step, the nuisance parameter θ^D is estimated from the previously obtained standardized residuals $\left(\frac{y_1}{\sqrt{h_1(\widetilde{\theta_T^V})}},...,\frac{y_T}{\sqrt{h_T(\widetilde{\theta_T^V})}}\right)$ by maximizing

$$\sum_{t=1}^{T} log\left[d_{\theta^D}\left(\frac{y_t}{\sqrt{h_t(\widetilde{\theta_T^V})}}\right)\right].$$

The main interest of this two-step approach is to consider separately the distribution and volatility parameters, reducing in particular the dimension of the optimization problems, even if the estimators may become inefficient (see Francq and Zakoian (2010) Chap.9). However, this estimation approach is likely to estimate leverage effects that are stronger than necessary, in order to fit the returns' skewness as much as possible during the first step, and therefore undermining the conditional distribution's role.

In our empirical experiments, we use another estimation approach that specifically aims at dealing with the potential inefficiency of the two previously mentioned methodologies. Such efficiency issues are all the more likely in our experiments as asymmetric GARCH models mixed with asymmetric conditional distributions use the volatility and distribution set of parameters in order to span a single dimension: the potential dynamic skewness of the underlying time series. We propose a new empirical recursive method (REC) that starts from the QML estimate and that iteratively maximizes the Likelihood over either the volatility or the distributions parameters:

1. We start from the QML estimate $\hat{\theta}_T^1 = (\widetilde{\theta_T^D}, \widetilde{\theta_T^V})$.
2. We re-estimate the volatility parameter θ^V by maximizing

$$\sum_{t=1}^{T} -\frac{log(h_t)}{2} + log\left[d_{\widetilde{\theta_T^D}}\left(\frac{y_t}{\sqrt{h_t}}\right)\right],$$

   obtaining $\hat{\theta}_T^{2,V}$.
3. We re-estimate the distribution parameter θ^D from the standardized residuals

$$\left(\frac{y_1}{\sqrt{h_1}(\hat{\theta}_T^{2,V})},...,\frac{y_T}{\sqrt{h_T}(\hat{\theta}_T^{2,V})}\right)$$

   by maximizing

$$\sum_{t=1}^{T} log\left[d_{\theta^D}\left(\frac{y_t}{\sqrt{h_t}(\hat{\theta}_T^{2,V})}\right)\right]$$

   and obtain $\hat{\theta}_T^{2,D}$ and $\hat{\theta}_T^2 = (\hat{\theta}_T^{2,D}, \hat{\theta}_T^{2,V})$.
4. We iterate this procedure until a good trade off between precision and computational cost is reached.

The precision and computational costs of each of these methods will be assessed numerically through Monte Carlo experiments in the next subsection.

2.3 Finite sample properties of the three estimation methodologies

In this subsection, we review and detail the results of the Monte Carlo experiments developed to test the finite sample properties of the three estimation strategies presented earlier. The comparisons are based on simulated processes mixing a special time-varying volatility structure with a non-Gaussian distribution as detailed in section 2.1. In such a case, a single step maximum likelihood estimation is due to be perturbed by the similar role played by parameters governing the volatility structure and the conditional distribution. In order to compare the estimation approaches, we propose to rely on the following criteria:

– The first criterion that we rely on is the Root Mean Square Error (RMSE hereafter) of the estimated parameters. It is due to help us deciding upon which methodology yields estimated parameters that are the closest to the true values used to sample the process. Let θ_i^0 be the true value for the parameter i and $\hat{\theta}_i^{j,n}$ be the estimated parameter using the methodology j, obtained from the n^{th} simulated sample. This parameter can belong either to the volatility structure or to the conditional distribution one. With a total number of N simulations and I parameters, the total RMSE is

$$RMSE_j = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\sum_{i=1}^{I}\left(\frac{\theta_i^0 - \hat{\theta}_i^{j,n}}{\theta_i^0}\right)^2}. \tag{10}$$

$RMSE_j$ should be seen as a score obtained by the estimation approach j. Larger parameters are usually affected by larger estimation errors: in the criterion that we propose, the errors are weighted by the true value of the parameter, making the aggregation of the estimation errors easier.

In the end, we compute three different scores. The first one is the previous score computed over the total number of parameters that we refer to as the "Total RMSE". The second and third scores focus on specific parameters: we are interested in the differences obtained with each estimation strategy when it either comes to the volatility parameters or to the distribution ones. We thus compute the previous RMSE criterion for these two different subsets of parameters. We refer to them as "Volatility RMSE" and "Distribution RMSE". These criteria are computed as follows:

$$\text{Volatility } RMSE_j = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\sum_{i=I_1}^{I_1'}\left(\frac{\theta_i^0 - \hat{\theta}_i^{j,n}}{\theta_i^0}\right)^2} \tag{11}$$

$$\text{Distribution } RMSE_j = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\sum_{i=I_2}^{I_2'}\left(\frac{\theta_i^0 - \hat{\theta}_i^{j,n}}{\theta_i^0}\right)^2}. \tag{12}$$

The parameters $i \in [I_1; I_1']$ are the volatility parameters and the parameters $i \in [I_2 : I_2']$ are the distribution ones.

– Maximizing a likelihood function over parameters usually requires to use a numerical optimizer. Most of these optimizers are gradient-based method, such as the Newton-Raphson algorithm. In our experiments we rely on the BFGS optimization algorithm. This method is a quasi-Newton method also known as "variable metric algorithm". Interested readers in such aspects can refer to Nocedal and Wright (1999). Such algorithm requires a certain number of iterations before it delivers the output of the likelihood maximization and take time. The second numerical aspect we are interested in is thus the time used by each method: computational burden is an aspect of numerical optimization that should not be left apart, as the feasibility of the approaches investigated here should be an important trigger for readers interested in our results. Hence, the second criterion used in this article is the average time required to perform the estimation of each simulated model using the different estimation strategies presented before.

The numerical features of the simulation strategy implemented here unfold as follows: the total number of experiments with each strategy for each model is 2,000 and the sample size used each time is 1,500 observations. The simulations are based on parameters selected so that to mimic the salient features of financial markets' returns. The parameters used are the following:

– For the APARCH volatility structure (4) we take $\delta = 1.2$, $a_0 = 0.04$, $a_1 = 0.3$, $b_1 = 0.6$ and $\gamma = 0.75$,

– For the EGARCH (3) volatility structure we take $a_0 = -0.4$, $a_1 = 0.1$, $b_1 = 0.96$ and $\gamma = -0.1$,

– For the MN density (6), the selected parameters are $\phi = 0.23$, $\mu_1 = -0.4$, $\sigma_1 = 1.3$, $\mu_2 = 0.12$ and $\sigma_2 = 0.86$,

– For the GH distribution (5) the selected parameters are $\alpha = 1.9$, $\beta = -0.55$, $\delta = 3.6$, $\mu = 0.55$, $\lambda = -5.5$.

For each experiment, the starting values of the parameters are obtained by perturbing the true values stated above in the following way:

$$\theta_i^{j,n} = \theta_i^0 \left(\frac{1}{2} + u \right), \tag{13}$$

for a given estimation approach j and the n^{th} replication of the Monte Carlo experience. u is a random variable which follows a uniform distribution over $[0 : 1]$. There are numbers of constraints required with these volatility structures and distributions. Once the parameters are perturbed, we check for their consistency with these constraints and discard those that do not match these requirements. What is more, we impose these constraints numerically within the optimization process. However, this is of little impact on the results as the starting point is selected to be close to the true value of the parameters. This would have a sharper influence on the results in the case of a real data set, involving the difficult step of the initialization of the parameters without knowing them. We do not recall these constraints here as they are detailed in the relevant literature on these volatility processes and on these distributions.

The Monte Carlo exercise is performed as follows: for a given model, we generate a sample using the true value for the parameters. Then, we perturb these values to obtain a starting point to optimize the likelihood function using the previous sample. At the end of this process, we store the relevant information and start again these steps. In the case of volatility, we initialize volatility to its long term average as estimated from the sample using the method of moment estimator. The REC estimation approach is run 10 times: this number was selected on the ground of various information. First, the improvement over this tenth step was unclear during our initial trial-and-error. Second, the average time required to obtain these 10 optimization steps is around 6 to 9 times the time required for the QML estimator to be run. We wanted to make the approaches broadly comparable, even from a time consumption perspective.

Now, we turn our attention to the detailed analysis of the results, specifying the outcome of this simulation-based assessment of the various estimation approaches detailed earlier:

– Case 1: the GH-APARCH model.

  The results for this model are presented in Table 1.

  [Table 1 about here]

  For this case, the lowest Total RMSE is obtained with the final step of the recursive estimation approach, whose value is 2.66. It improves the QML starting point whose value is 3.65. The ML approach behaves poorly, yielding a Total RMSE equal to 4.69. The REC10 improvement stems mainly from the estimation of the volatility parameters, as the ML approach is the worst performing competitor, with a Volatility RMSE equal to 3.99, vs. 3.45 for the QML and 2.54 for the REC10. When it comes to the distribution's parameters estimation, the ML becomes the best competitor with a RMSE of 0.64, vs. 1.21 for the QML and 0.8 for REC10. In the end, the improvement over the volatility structure estimation is so important with the REC10 approach that it dominates the Total RMSE obtained. Finally, the average time required to perform the estimation is 6 seconds for QML, 12 seconds for ML and 55 seconds for REC10[1].

– Case 2: the MN-APARCH model.

  The results for this model are presented in Table 2.

[1]This exercise has been done using a Intel Xeon E5420 PC (2,5 GHz, 1333 MHz, 2X6 Mo)

[Table 2 about here]

This case is very similar to the latter as the REC10 approach delivers the best Total RMSE (2.16), improving not only the QML score (3.08), but also the ML case (4.69). Again, the main explanation stems from the volatility's parameters estimation as for the Volatility RMSE, the hierarchy between the approaches remains the same: ML obtains 4.67, QML 2.99 and REC10 2.09. The distribution's parameters are best estimated with the ML method (0.43), whereas the REC10 one comes second (0.55) while still improving the QML case (0.75). However, these figures remain remarkably close to each other. From a convergence time REC10 uses 61 seconds when the ML estimators are obtained after 14 seconds and the QML ones after 7 seconds.

– Case 3: the GH-EGARCH model.

  The case of GH-EGARCH is presented in Table 3.

[Table 3 about here]

The results obtained here are different from those obtained before: it appears that QML gets the lowest Total RMSE (3.12), while being very close to the ML estimates (3.14). The REC10 delivers a score of 3.64, which remains however close to the others. This is again the result of the estimation of the volatility parameters, as a similar ranking is obtained with the Volatility RMSE: QML yields 2.96, ML gets 3.04 and finally REC10 delivers a ratio of 3.58. Here again the scores remain close. Finally, the QML approach is the fastest strategy, with an average of 5.97 seconds, when ML yields estimates after 9.82 seconds and REC10 after 62.14 seconds.

– Case 4: the MN-EGARCH model.

  This final model's results are presented in Table 4.

[Table 4 about here]

For such a case, the REC10 dominates again the Total RMSE ranking (1.38), whereas QML is the worst performing competitor (3.16), right behind ML (1.77). In this case, the volatility results explain most of this performance, as the REC10 shrinks the QML's Volatility RMSE from 3.02 to 1.15. ML gets a score of 1.68. The story for the Distribution RMSE is different again, as ML appears to be the best competitor with a score of 0.55, when QML gets a score of 0.92 that is reduced by the REC10 to 0.76. The gap between ML and REC10 is thus small. This result is thus similar to the one obtained in the GH-EGARCH. The time consumption in this case for the REC10 approach is smallest of all our experiments, as it requires only 42 seconds. The same comment applies for the ML whose average time needed to estimate the parameters is 8 seconds.

We can thus summarize the salient features of this exercise as follows: first, the REC10 approach always improves the QML one, whatever the model and the subset of parameters investigated. Second, the REC10 dominates 3 out of the 4 Total RMSE investigated, mostly because it provides estimates of the volatility parameters with smaller estimation errors. The results obtained in the case of the distribution's parameters do not provide such sharp conclusions. Even though the ML estimates remain the most precise ones, the difference between the REC10 estimates and the ML ones remain small in every case. A notable exception is obtained in the GH-EGARCH case, with a lower Distribution RMSE in the REC10 case than in the ML one. This overall nice performance of the REC10 approach comes at the cost of a longer estimation time. Given the REC performances, we will rely on it throughout the rest of the paper.

3 Testing for leverage effect

This section is dedicated to testing for leverage effect in S&P 500's returns.

3.1 Description of the data set

We estimate the various models presented earlier using a data set of returns on the S&P 500. The data set starts on January, 2^{nd} 1987 and ends on July, 20^{st} 2011. It includes different market phases and four marked market crashes in 1987, 1998, 2001 and 2008. It is made of daily closing prices in US Dollar. The full sample contains 6190 observations. Figure 1 charts the evolution of the index over the sample. Descriptive statistics are provided by Table 5.

[Figure 1 about here]

[Table 5 about here]

The S&P 500 returns are characterized by fat tails, as the excess kurtosis is positive in the full sample (29.075). The sample's skewness is equal to -1.341 highlighting the asymmetry to the left of the S&P 500's returns. Both stylized facts are rather stable in the two sub-samples presented in Table 5. However, the skewness associated to the 1999-2011 sample is very close to zero. Investigating Figure 1, the balance between rising and falling phases explains part of this surprising figure. Now, as discussed earlier, a zero skewness can be the result of the combination of leverage effect with positive conditional skewness, leaving ample room for our measurement experiment.

3.2 Comparison of estimation methodologies for S&P 500's returns

This section proposes estimates that cast additional light on the ability of the recursive estimation strategy to disentangle leverage from unconditional skewness in financial time series. We estimate the parameters of the models combining the volatility structures and the distributions used in our Monte Carlo experience and the three estimation strategies: ML, QML and REC. The starting values to perform these estimates are the same for each method and are set to be equal to the values used in the Monte Carlo experience. All estimations converged with those starting points. However the time consumption – that is not provided here – of the recursive method remains the highest of the three methods under the scope of our investigations. The estimated parameters for the EGARCH-GH, the EGARCH-MN, the APARCH-GH and the APARCH-MN models are respectively provided in Tables 6, 7, 8 and 9.

[Table 6 about here]

[Table 7 about here]

[Table 8 about here]

[Table 9 about here]

The examination of the estimation tables gives the impression that the estimated parameters can be different, even as the sample remains the same. For example, focusing on the results of Table 7, the size of the leverage effect is clearly different across the estimation approaches. In the case of the EGARCH model, the returns-to-volatility spillover effect is captured in the conditional variance equation:

$$\log(h_t) = a_0 + a_1|z_{t-1}| + \gamma z_{t-1} + b_1 \log(h_{t-1}). \tag{14}$$

Whenever z_t is positive, the volatility is increased by a factor equal to $a_1 + \gamma$. When it is negative, this factor becomes $a_1 - \gamma$. The leverage effect is characterized by the difference between these two figures. In the EGARCH-MN case, QML estimation yields a difference that equals 0.33. In the case of the ML strategy this figure becomes 0.177. Finally, REC estimations point to a difference equal to 0.224. Such a disparity between the figures obtained from the three estimation methods put their reliability into questions. The question is now to detect which of these estimations provide the fairer results beyond the intuitions gathered in the previous Monte Carlo experiments. We propose to answer this question using two different arguments. First, we propose to test which of these estimation strategies yield the best fit to the joint distribution of the sample considered here, analyzing model by model results. Then, we propose to investigate the stability of the news impact curves obtained for each model and of the estimated conditional distribution, showing how different models applied to the same data set can produce various estimates of similar quantities. All in all, we find that

REC estimates are the most stable and accurate ones.

We now focus on the fact that one of these estimation approaches should be better than the remaining two others, making the underlying model as likely as possible given the sample at hand. To compare the estimation methods we propose to use the joint density of the sample as a score. The test we use can be seen as an in-sample version of the test proposed for density forecast in Amisano and Giacomini (2007) as presented in Vuong (1989). Say we deal with a time series model for the log-returns whose estimated conditional density at time t is $f(Y_t|\underline{Y_{t-1}}, \hat{\theta}_1)$, where $\underline{Y_{t-1}} = (Y_1, ..., Y_{t-1})$ and $\hat{\theta}_1$ is the vector of parameters describing the shape of this conditional distribution and the volatility structure estimated from methodology 1. We compare this estimation method to another one defined by the conditional density $f(Y_t|\underline{Y_{t-1}}, \hat{\theta}_2)$, with $\hat{\theta}_2$ being the estimated parameters obtained from this second method. The null hypothesis of the test is "methods 1 and 2 provide a similar fit of the log-return's conditional distribution using the same underlying model". The corresponding test statistic is then:

$$t_{1,2} = \frac{1}{n} \sum_{t=1}^{n} \left(\log f(Y_t|\underline{Y_{t-1}}, \hat{\theta}_1) - \log f(Y_t|\underline{Y_{t-1}}, \hat{\theta}_2) \right), \tag{15}$$

where n is the total number of observations available. Under the null hypothesis

$$\frac{t_{1,2}}{\hat{\sigma}_n} \sqrt{n} \underset{n \to +\infty}{\to} \mathcal{N}(0, 1), \tag{16}$$

where $\hat{\sigma}_n$ is a properly selected estimator for the statistic volatility. Here, as proposed in Amisano and Giacomini (2007), we use a Newey-West estimator, with a lag empirically retained to be large (around 25).

Table 10 provides the results of this test, displaying the test statistics as presented in equation (16).

[Table 10 about here]

We use three tests: the first one compares the ML vs. the QML approach; the second one compares the REC vs. the QML one and the third one compares the REC to the ML estimates. We perform such a comparison over the full sample of returns described earlier as well as on two sub-samples, one starting in 1987 and ending in 1998 and the second starting in 1999 and ending in 2011 so that to test the stability of our empirical findings. The core results can be summarized as follows: (1) ML and REC dominates the QML results, in a stronger manner than in our Monte Carlo simulations. (2) The REC estimates outperform the ML ones in most cases. When it is not the case, both sets of estimates are found to be equivalent. The global argument is thus that the REC approach advocated here is an improvement both over the ML and the QML methods: the estimates obtained stand a greater chance to be closer to the true set of parameters that makes the likelihood function maximum. Those results are however remarkably consistent with our Monte Carlo experiments.

[Figure 2 about here]

[Figure 3 about here]

[Figure 4 about here]

[Figure 5 about here]

[Figure 6 about here]

[Figure 7 about here]

Figures 2, 4, 6 and Figures 3, 5, 7 present our second argument supporting the relative domination of the recursive estimation methodology when it comes to models mixing different sources of skewness and kurtosis: these figures chart the estimated news impact curves for each model using each estimation strategy (Figures 2, 4, 6) and similar figures for the estimated conditional densities (Figures 3, 5, 7). We first turn specifically toward news impact curves. This concept introduced in the financial econometrics literature in Engle and Ng (1993) measures the contribution of past returns to volatility in GARCH-type models. In a general setting, Figures 2, 4, 6 chart the volatility on date t as a function of the return on date $t-1$. As this function clearly depends on the model used, we do not provide the analytical expression for each of them. Defining a plausible support for the returns' value and setting the past volatility to its long term average, we compute volatility's expected value for the next day. Investigating Figure 6, we clearly obtain a high variability between the different shapes obtained for the QML and to a lesser extent for the ML methods. Our main finding here is that the REC approach provides a shape that is stable across all models: the scale of values obtained for this news impact curves is the same across the four models. These results can be directly extended to the estimated conditional distribution presented in Figure 7: the REC method provides similar shapes for the conditional density across models unlike the ML and the QML methods. If this phenomenon is not exactly a surprise when it comes to the QML method, the case of ML is more striking. Both findings are stable across the two sub-samples presented in Figures 2, 4 and 3, 5. Examining Tables 6 to 9 yields a hint regarding a potential explanation to our findings: the problem that arises with the ML estimation strategy seems to be partly related to the relation between leverage effects and skewness. The mixing effects advocated earlier seems to have a disruptive effect, leading to the instability in the estimated news response function and conditional distribution. A similar comment can be made when it comes to the kurtosis implied by the conditional distribution. From Table 6 for example, the estimated value for α in the GH distribution is clearly different with the ML approach, as it reaches 0.793 whereas it is equal to 0.265 with REC and 0.245 with QML. REC's figures across GH-based models are fairly comparable, whereas ML estimates are very different. Again, such a feature holds in the results obtained from the sub-samples.

This instability could lead to a wrong model selection. Using the density test presented earlier, we propose to test which model should be favored for the data set at hand and for a given estimation strategy. For example, in the case of the REC estimation method, we test the different combinations of estimated models that mix the EGARCH/APARCH volatility structures to the MN/GH conditional distributions using the density test that we used previously to compare the estimation methods. We now apply this test methodology to test which model dominates the others for a given estimation methodology. The test results are presented in Table 11.

[Table 11 about here]

Given the transitivity of this test statistics, it is possible to rank the different models and thus decide on the model(s) that should be favored. In the QML estimates case, all models are found to be equivalent but the APARCH-MN model that is favored to the EGARCH-MN one. In the case of the ML estimates, the APARCH-GH model is dominated by all the other models that are found to be equivalent. The conclusion obtained with the REC estimates is different: each model provides an equivalent fit. Hence, depending on the estimation strategy, the ranking obtained is different and misleading conclusions can be reached in terms of model specification. Given the various test results discussed above and the stability of the results obtained with the REC estimation approach, the estimation results obtained with the latter method are apparently the most trustworthy. We will now rely on it to perform an assessment of the statistical significance of the leverage effect in S&P 500's returns.

3.3 Testing the empirical significance of leverage effects

This section details the results of our empirical experiments that aim to test for the empirical relevance of a return-to-volatility asymmetric feedback component in volatility models. We rely on three types of analysis: first, we simply report an estimate of the percentage of the sample's total skewness that can be accounted for by the leverage component of the GARCH-type models, once these models have been properly estimated using the recursive method detailed in the previous section. Second, we properly test for the in sample significance of the leverage component in GARCH-type models using Hansen (1992)'s test that is used to compare these nested models contrary to the Vuong (1989)'s methodology presented in section 3.2. Finally, we confirm the findings obtained using the first two steps by gauging the usefulness of the leverage effect when it comes to forecasting the future density of returns. To do so, we rely on Amisano and Giacomini (2007)'s forecast density test methodology.

3.3.1 The explanatory power of the leverage effect over return's skewness

All our experiments are based on the S&P 500 data set presented earlier. In order to measure the share of the sample's skewness that comes from the leverage effect and the share that comes from the conditional distribution, we rely on two sets of estimates. The first set comes from a recursive estimation of EGARCH and

APARCH models, mixed with a GH and a mixture of two Gaussian distributions, with no restrictions on the parameters. A second set of estimates is obtained by constraining the parameter in the asymmetric GARCH models that control the leverage effect γ (as in equations (3) and (4)) to be equal to zero. The metric that will help us to measure the importance of the leverage effect to explain the sample's skewness will be based on the skewness of the residuals obtains with each set of parameters for a given models. Let $s_{\gamma=0}$ be the skewness obtained in the residuals when γ is constrained to zero and $s_{\gamma\neq 0}$ be the same quantity when γ is unconstrained. The following ratio measures the decrease in the residuals' skewness that comes from the GARCH models' leverage component:

$$r = \frac{s_{\gamma=0} - s_{\gamma\neq 0}}{s_{\gamma=0}}. \tag{17}$$

When the leverage component fully explains the returns' skewness, $s_{\gamma\neq 0} = 0$ and $r = 100\%$. When this component does not explain any part of the sample's skewness, then $s_{\gamma\neq 0} = s_{\gamma=0}$ and $r = 0$. Tables 12, 13 and 14 provide the parameters' estimates along with r, $s_{\gamma=0}$ and $s_{\gamma\neq 0}$ for each models and sub-samples of our main data set.

[Table 12 about here]

[Table 13 about here]

[Table 14 about here]

Focusing on the full sample results presented in Table 14, we obtain consistent evidence that the leverage effect only explains a weak part of the returns' skewness. For the four models used in our experiments, the ratio described by equation (17) ranges from around 11% in the APARCH-MN case to around 20% in the APARCH-GH case. The EGARCH based models lead to a ratio approximately equal to 16%. These conclusion hold as well when the ratio is computed from subsets of our full sample. In all our experiments, the maximum r is obtained with the APARCH-GH model over a sample of returns covering the 1987-1998 period. In this case $r = 37.6\%$: our findings thus show that when properly estimated, the leverage effect only account for a third of returns' skewness. However, we consistently find that the γ parameter across models is estimated to be statistically different from zero: there is a leverage effect at work in the dynamics of the S&P 500 returns, but it does not play a role that is as significant as what the literature usually assumes. In order to better measure the extent in which leverage effect statistically matters, we now use in and out of sample likelihood ratio tests.

3.3.2 Testing for leverage effect in-sample

Beyond the previous evidence showing that leverage effect has a weaker than expected quantitative contribution to financial returns' skewness, we now propose to test for the significance of the leverage parameter in the various GARCH specifications considered previously. To test the in-sample performances of GARCH type models with or without leverage parameters we apply the methodology developed by Hansen (1992) that may be used to compare nested models contrary to the Vuong (1989)'s methodology presented in section 3.2.

Viewing the likelihood as a function of the unknown parameter (here the leverage parameter), it is possible to find, under mild conditions, a bound for the asymptotic distribution of a standardized likelihood ratio statistic using empirical process theory. Even if this flexible framework can take into account unidentified nuisance parameters the presentation we adopt here focus on the case (implied by the GARCH specification we empirically consider in the paper) where they are all identified. We consider a sample of size T $(y_1, ..., y_T)$ for the log-returns. In this part we denote by $\gamma \in \Gamma$ the leverage parameter of the volatility structure we consider and $\theta \in \Theta$ all the remaining parameters. We start from a log-likelihood of the form

$$L_T(\gamma, \theta) = \sum_{i=1}^{T} l_i(\gamma, \theta),$$

where $l_i(\gamma, \theta)$ is the conditional log-likelihood given $\mathcal{F}_{i-1}$. This form typically arises in the estimation of GARCH-type models. The null and alternative hypotheses are:

$$H_0 : \gamma = 0 \qquad vs \qquad H_1 : \gamma \neq 0.$$

We consider the sequence of maximum-likelihood estimate of θ for fixed values of γ based on the sample $(y_1, ..., y_T)$:

$$\hat{\theta}_T(\gamma) = argmax_{\theta \in \Theta} L_T(\gamma, \theta). \tag{18}$$

The concentrated and standardized likelihood ratio process becomes

$$\hat{LR}^*_T(\gamma) = \frac{\hat{LR}_T(\gamma)}{V_T(\gamma)^{1/2}}$$

where $\hat{LR}_T(\gamma) = L_T(\gamma, \hat{\theta}_T(\gamma)) - L_T(0, \hat{\theta}_T(0))$ is the concentrated likelihood ratio and

$$V_T(\gamma) = \sum_{i=1}^{T} \left(\underbrace{l_i(\gamma, \hat{\theta}_T(\gamma)) - l_i(0, \hat{\theta}_T(0)) - \frac{\hat{LR}_T(\gamma)}{T}}_{q_i(\gamma, \hat{\theta}_T(\gamma))} \right)^2$$

its sample variance. It can be proved from the empirical process theory that the standardized likelihood ratio statistic $\hat{LR}^*_T = sup_{\gamma \in \Gamma} \hat{LR}^*_T(\gamma)$ fulfills $\forall x \in \mathbb{R}$,

$$\mathbb{P}\left(\hat{LR}^*_T \geq x\right) \leq B_T(x) \underset{T \to \infty}{\longrightarrow} \mathbb{P}\left(sup_{\gamma \in \Gamma} Q^*(\gamma) \geq x\right) \tag{19}$$

where Q^* is a centered Gaussian process completely characterized by its covariance K^* that is not exactly known but may be approximated for large T by its sample analogue

$$\hat{K}^*_T(\gamma_1, \gamma_2) = \frac{\hat{K}_T(\gamma_1, \gamma_2)}{V_T(\gamma_1)^{1/2} V_T(\gamma_2)^{1/2}},$$

$$\hat{K}_T(\gamma_1, \gamma_2) = \sum_{i=1}^{T} \hat{q}_i(\gamma_1)\hat{q}_i(\gamma_2) + \sum_{k=1}^{M} w_{k,M} \left[\sum_{i=1}^{T-k} \hat{q}_i(\gamma_1)\hat{q}_{i+k}(\gamma_2) + \sum_{i=1+k}^{T} \hat{q}_i(\gamma_1)\hat{q}_{i-k}(\gamma_2) \right],$$

where $\hat{q}_i(\gamma) = q_i(\gamma, \hat{\theta}_T(\gamma))$, $w_{k,M} = 1 - |k|/(M+1)$ is the Bartlett kernel and M is a bandwidth number. Now, the approximated p values of the test can be computed by Monte Carlo methods: when $(u_i)_{1 \leq i \leq T+M}$ are i.i.d $\mathcal{N}(0,1)$,

$$\tilde{LR}^*(\gamma) = \frac{\sum_{k=0}^{M} \sum_{i=1}^{T} \hat{q}_i(\gamma) u_{i+k}}{\sqrt{1+M} V_T(\gamma)^{1/2}},$$

is a Gaussian process with covariance $\hat{K}^*_T$. Moreover, we can compute its supremum over γ forming a reasonable grid $[\gamma_0, ..., \gamma_N]$ centered on the value 0. Practically, the theory does not give any particular guidance for the choice of M, thus, all the numerical work is calculated for $M = 0, 1, \ldots, 5$. We also consider a grid with step size 0.05 containing 10 points centered on 0 and in order to achieve some efficiency in the computation of $\tilde{LR}^*$ we compute along the grid each $\hat{\theta}_T(\gamma)$ using as a starting point the preceding obtained value. Then, the distribution of $sup_{\gamma \in \Gamma} Q^*(\gamma)$ and the associated p values of the test are approximated using 1000 independent realizations of $sup_{\gamma \in \Gamma} \tilde{LR}^*(\gamma)$.

To conclude, let us remark that, under Hansen's test, we do not know the exact asymptotic distribution of the standardized likelihood ratio statistic but only an asymptotic bound for its tail function. This approach may appear as conservative but it is important to remark that the results are obtained under remarkably mild hypotheses. In particular, the classical maximum likelihood estimation of the parameters in (18) may be replaced by another estimation approach once it is consistent. In the following, this test is performed using the REC estimation methodology to provide an additional qualitative analysis to measure the impact of the leverage effect on the general shape of the unconditional density of the financial returns.

[Table 15 about here]

Table 15 presents the results obtained from Hansen (1992)'s test. As for our previous results, we obtain consistent conclusions across the various sets of models used in our experiments: at a 5% risk level, none of the γ estimates are found to be different from zero, whatever the value of M. For example, in the case of the EGARCG-MN model with $M = 5$, the test p-value is found to be equal to 13% that is the minimum value for each of the M used in our tests. The lowest p-value obtained is 6% in the case of the APARCH-GH model with $M = 2$. As for our previous results, breaking down our full sample into sub-periods does not change the conclusions raised from our test: the γ parameter is found to be statistically equal to zero. Still, the p-value obtained with the 1987-1998 sample are higher than the one obtained with the 1999-2011 one, probably in relation with the increase in the number of crisis periods in the second sample. We now turn out our attention to an out of sample analysis of leverage effect.

3.3.3 The role of leverage effect in forecasting the density of returns

If our previous tests consistently highlighted the weak role played by leverage effect to accurately describe the density of returns, it says little regarding its role in forecasting the future density of returns. To assess the contribution of leverage effect when it comes to returns forecasting we use the forecasting density test described in Amisano and Giacomini (2007). In our presentation we follow Maheu and McCurdy (2011), focusing on the ability of the approach to test multi-period forecasts.

For $M \in \{A, B\}$ we consider two competing GARCH-type models of the following form:

$$Y_t = log\left(\frac{S_t}{S_{t-1}}\right) = \underbrace{\sqrt{h_t} z_t}_{\varepsilon_t}, \quad S_0 = s \in \mathbb{R}_+, \tag{20}$$

where the $(z_t)_{t\in\{1,\dots,T\}}$ are independent and identically distributed random variables following an arbitrary distribution $D^M(0,1)$ with mean 0 and variance 1 and where

$$h_t = F^M(z_{t-1}, h_{t-1}) \tag{21}$$

where $F^M : \mathbb{R}^2 \to \mathbb{R}_+$ is compatible with realistic GARCH(1,1)-type volatility models. We denote by θ^M the parameters of the model M including the GARCH and the distribution parameters. These parameters are evaluated using an estimation method denoted by $Est(M)$.

Starting from a sample $(y_1, ...y_T)$ of size T we want to test forecast horizons $1 \leq k \leq k_{max}$ through rolling-window forecasting schemes of size τ. Thus, for $k \geq 1$, the average predictive likelihood is given by

$$D_{M,k} = \frac{1}{T - \tau - k_{\max} + 1} \sum_{t=\tau+k_{\max}-k}^{T-k} \log f_{M,k}(y_{t+k}, \widehat{\theta^M_{t+k}} | \mathcal{F}_t), \tag{22}$$

where $\widehat{\theta^M_{t+k}}$ is estimated via $Est(M)$ using the sample $(y_{t-\tau-k_{max}+1}, ..., y_t)$ and where

$$f_{M,k}(y_{t+k}, \widehat{\theta^M_{t+k}} | \mathcal{F}_t)$$

is the conditional density of Y_{t+k}, given $\mathcal{F}_t$ and $\widehat{\theta^M_{t+k}}$, evaluated at the realized log-return y_{t+k}. The particular form of (22) allow us to obtain a term structure of average predictive likelihoods, $(D_{M,1}, ..., D_{M,k_{max}})$ to compare the performance of alternative models, M, over an identical set of out-of-sample data points $(y_{\tau+k_{\max}}, ..., y_T)$.

In our GARCH(1,1) setting the conditional densities $f_{M,k}(x, \widehat{\theta^M_{t+k}} | \mathcal{F}_t)$ don't have a closed form except for $k = 1$ and have, in general, to be evaluated generating independent realizations of Y_{t+k} given $\mathcal{F}_t$ and using classical density kernel estimators. Nevertheless, remarking that

$$f_{M,k}(y_{t+k}, \widehat{\theta^M_{t+k}} | \mathcal{F}_t) = \int_{\mathbb{R}_+} f_{M,k}(y_{t+k}, \widehat{\theta^M_{t+k}} | h_{t+k}) p(h_{t+k} | \mathcal{F}_t) dh_{t+k} \tag{23}$$

we have

$$f_{M,k}(y_{t+k}, \widehat{\theta^M_{t+k}} | \mathcal{F}_t) \approx \frac{1}{N} \sum_{i=1}^{N} f_{M,k}(y_{t+k}, \widehat{\theta^M_{t+k}} | h^i_{t+k}) \tag{24}$$

where $f_{M,k}(y_{t+k}, \widehat{\theta^M_{t+k}}|h_{t+k})$ is the density of a

$$\sqrt{h_{t+k}} D^M(0,1)$$

evaluated at the realized log-return y_{t+k} and where h^i_{t+k} are independent realizations of h_{t+k} generated from the GARCH(1,1) structure (21) with parameters extracted from $\widehat{\theta^M_{t+k}}$ and starting values h_t and z_t.

According to Amisano Giacomini (2007), under the null hypothesis of equal performance, the statistic based on predictive likelihoods of horizon k for models A and B,

$$t^k_{A,B} = \frac{(D_{A,k} - D_{B,k})\sqrt{T - \tau - k_{max} + 1}}{\hat{\sigma}_{A,B,k}} \tag{25}$$

is asymptotically standard Normal, where $\hat{\sigma}_{A,B,k}$ is a properly selected estimator for the variance of

$$\log f_{A,k}(y_{t+k}, \widehat{\theta^A_{t+k}}|\mathcal{F}_t) - \log f_{B,k}(y_{t+k}, \widehat{\theta^B_{t+k}}|\mathcal{F}_t).$$

Here, as proposed Amisano Giacomini (2007), we use a Newey-West estimator that take into account heteroskedasticity and autocorrelation. One of the main interest of this approach comes from the fact that the two models can be nested or not and can be estimated using very different techniques from the moment that they are based on a finite estimation window. Moreover, if the forecaster wants to focus on the predictive performance in the right or the left tail of the distribution, the following weights can be used:

- Center of distribution: $\omega_1(Y) = \phi(Y)$, ϕ standard normal density function
- Tails of distribution: $\omega_2(Y) = 1 - \phi(Y)/\phi(0)$,
- Right tail: $\omega_3(Y) = \Phi(Y)$, Φ standard normal distribution function
- Left tail: $\omega_4(Y) = 1 - \Phi(Y)$.

In this case, the preceding method may be extended considering in (22)

$$\omega_i(y^{sd}_{t+k}) \log f_{M,k}(y_{t+k}, \widehat{\theta^M_{t+k}}|\mathcal{F}_t) \tag{26}$$

instead of

$$\log f_{M,k}(y_{t+k}, \widehat{\theta^M_{t+k}}|\mathcal{F}_t)$$

where $y^{sd}_{t+k} = \frac{y_{t+k} - \hat{\mu}_{k,t}}{\hat{\sigma}_{k,t}}$ is the standardized realized log-returns taking respectively for $\hat{\mu}_{k,t}$ and $\hat{\sigma}_{k,t}$ the mean and the standard deviation of $(y_{t-\tau-k_{max}+1}, ..., y_t)$.

[Table 16 about here]

Table 16 presents the tests results obtained using a rolling window of 2000 trading days. The first sample ends on May 8, 2003. We use a fixed window scheme rolling from one day to another and reestimating each time each model's parameters using the recursive estimation methodology. Here again, the consistency of our results is significant: for forecasting horizons ranging from 1 to 60 days the models with leverage effect are found to be equivalent to models without leverage effects. This result holds for the different weight functions. For example, in the case of the EGARCH-MN model, when using the w_0 weight function and for a forecast horizon of 30 days, the Amisano and Giacomini (2007)'s test statistics is equal to -1.356: a model with or without leverage effect provides statistically equivalent density forecasts at a 5% risk level. Now, our results contain three cases for which both models are not found to be statistically equivalent: in the EGARCH-GH with the w_0 and w_2 weights and in the EGARCH-MN with the w_4 weight the model without leverage effect is found to dominate the one with leverage effect. Hence across all our forecasting results, the model without leverage effect is found to be equivalent or better than the one with leverage effect.

Considering our three previous tests and measures, we obtain very consistent results showing that once properly estimated, the leverage effect does not seem to have a statistical importance when it comes to equity returns as proxied by the S&P 500's returns.

4 conclusion

This article questions the empirical usefulness of leverage effects to describe the dynamics of equity returns. Using a recursive estimation scheme that accurately disentangles the asymmetry coming from the conditional distribution of returns and the asymmetry that is related to the past return to volatility component in GARCH models, we test for the statistical significance of the latter. Relying on both in and out of sample tests we consistently find a weak contribution of leverage effect over the past 25 years of S&P 500 returns, casting light on the importance of the conditional distribution in time series models.

References

[1] Aydemir, A.C., Gallmeyer, M. and Hollifield, B., 2006. Financial Leverage Does Not Cause the Leverage Effect. Working Paper, Carnegie Mellon University.

[2] Akgiray, V., Booth, G.G., 1987. Compound Distribution Models of Stock Returns: An Empirical Comparison. The Journal of Financial Research 10, 269-280.

[3] Alexander, C., Lazar, E., 2006. Normal Mixture GARCH(1,1): Applications to Exchange Rate Modeling. Journal of Applied Econometrics 21, 307-336.

[4] Amisano, G., Giacomini, R., 2007. Comparing Density Forecasts via Weighted Likelihood Ratio Tests. Journal of Business & Economic Statistics, American Statistical Association 25, 177-190.

[5] Awartani, B.M.A., Corradi, V., 2005. Predicting the volatility of the S&P -500 stock index via GARCH models: the role of asymmetries. International Journal of Forecasting, 21, 167-183.

[6] Badescu, A., Kulperger, R., Lazar, E., 2008. Option Valuation with Normal Mixture GARCH Models. Studies in Nonlinear Dynamics and Econometrics 12 (2), 1580-1580.

[7] Badescu, A., Elliott, R.J., Kulperger, R., Miettinen, J. and Siu, T.K., (2011). A comparison of pricing kernels for GARCH option pricing with generalized hyperbolic distributions. International Journal of Theoretical and Applied Finance, 14(5), 669-708.

[8] Barndorff-Nielsen, O.E., 1977. Exponentially decreasing distributions for the logarithm of particle size. Proceedings of the Royal Society of London Series A, 353, 401-419.

[9] Bai, X., Russell, J. R. and Tiao, G. C., 2003. Kurtosis of garch and stochastic volatility models with non-normal innovations. Journal of Econometrics, 114, 349360.

[10] Bandi, F.M, Reno, R., 2012. Time varying leverage effects. Journal of Econometrics, 169(1), 94-113.

[11] Bates, D., 1996. Jumps and stochastic volatility: exchange rate processes implicit in deutsche mark options. Review of Financial studies, 9(1), 69-107.

[12] Behboodian, J., 1970. On the Modes of a Mixture of Two Normal Distributions. Technometrics, 12(1), 131-139.

[13] Bekaert, G., Wu, G., 2000. Asymmetric volatility and risk in equity markets. Review of Financial Studies, 13, 142.

[14] Black, F., 1976. Studies of stock prices volatility Changes. Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economics Statistics Section, 177-189.

[15] Bollerslev, T., 1987. A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return, Review of Economics and Statistics, 69, 542-547.

[16] Campbell, J.Y., Hentschel, L., 1992. No news is good news: An asymmetric model of changing volatility in stock returns. Journal of Financial Economics, 31, 281318.

[17] Chorro, C., Guégan, D., Ielpo, F., 2010. Martingalized Historical approach for Option Pricing. Finance Research Letters 7 (1), 24-28.

[18] Chorro, C., Guégan, D. and Ielpo, F., 2012. Option pricing for GARCH type models with generalized Hyperbolic Innovations. Quantitative Finance, 12(7), 1079-1094.

[19] Christie, A.A., 1982. The stochastic behavior of common stock variances. Journal of Financial Economics, 10, 407432.

[20] Christoffersen, P. and Jacobs, K., 2004. Which Garch model for Option valuation. Management science , 50(9), 1204-1221.

[21] Christoffersen, P., Heston, S., Jacobs, K., 2006. Option Valuation with Conditional Skewness. Journal of Econometrics 131, 253-284.

[22] Christoffersen, P., Elkamhi, R., Feunou, B., Jacobs, K., 2010. Option Valuation with Conditional Heteroskedasticity and Non-Normality. Review of Financial studies, 23(5), 2139-2183.

[23] Corsi, F., Reno, R., 2012. Discrete-Time Volatility Forecasting With Persistent Leverage Effect and the Link With Continuous-Time Volatility Modeling. Journal of Business & Economic Statistics, 30(3), 2012.

[24] Curto, J., Pinto, J., Tavares, G., 2009. Modeling stock markets volatility using GARCH models with normal Students t and stable Paretian distributions. Statistical Papers, 50 (2), 311-321.

[25] Ding Z., Granger, C.W.J., Engle, R.F., 1993. A Long Memory Property of Stock Market Returns and a New Model. Journal of Empirical Finance 1, 83-106.

[26] Duffee, G.R., 1995. Stock returns and volatility: A firm level analysis. Journal of Financial Economics, 37, 399420.

[27] Eberlein, E., Prause, K., 2002. The Generalized Hyperbolic Model: Financial Derivatives and Risk Measures, in: Geman, H., Madan, D., Pliska, S., Vorst, T. (Eds.), Mathematical Finance-Bachelier Congress 2000. Springer Verlag, pp. 245-267.

[28] Engle, R. F., Ng, S., 1993. Measuring and Testing the Impact of News on Volatility. Journal of Finance, American Finance Association 48 (5), 1749-1778.

[29] Figlewski, S., Wang, X., 2000. Is the leverage effect a leverage effect?. Working Paper, New York University.

[30] Francq, C. and Zakoian, J.M., 2010. GARCH Models: Structure, Statistical Inference and Financial Applications. John Wiley and Sons, Ltd.

[31] Francq, C., Wintenberger, O. and Zakoian, J.M., (2012). GARCH models without positivity constraints: Exponential or Log GARCH?. Working paper.

[32] Guégan, D., Ielpo, F. and Lalaharison, H., 2013. Option Pricing with Discrete Time Jump Processes. To appear in the Journal of Economic Dynamics and Control.

[33] Hansen, B.E., 1992. The Likelihood Ratio Test under Nonstandard Conditions: Testing the Markov Switching Model of GNP. Journal of Applied Econometrics, 7(1), 61-82.

[34] Heston, S.L., 1993. A closed-form solution for options with stochastic volatility, with applications to bond and currency options. The Review of Financial Studies, 6, 327-343.

[35] Kon, S. J., 1984. Models of Stock Returns: A Comparison. The Journal of Finance 39, 147-165.

[36] Maheu, J.M. and McCurdy, T.H., 2011. Do high-frequency measures of volatility improve forecasts of return distributions?. Journal of Econometrics, 160, 69-76.

[37] Meitz, M. and Saikkonen, P., 2012. Parameter estimation in nonlinear AR.GARCH models. Econometric Theory, 27(6), 1236-1278.

[38] Nelson, D.B., 1991. Conditional Heteroskedasticity in Asset Returns. Econometrica 59, 347-370.

[39] Newey, W. and Steigerwald, D., 1997. Asymptotic bias for quasi-maximum-likelihood estimators in conditional heteroscedasticity models. Econometrica, 65, 587599.

[40] Nocedal, J., Wright, S. J., 1999. Numerical Optimization. Springer, Berlin.

[41] Poon, S. H., Granger, C.W.J., 2003. Forecasting volatility in financial markets: A review. Journal of Economic Literature, 41, 478539.

[42] Schwert, G.W., 1989. Why does stock market volatility change over time?. Journal of Finance, 44, 1115?1153.

[43] Straumann, D., 2005. Estimation in Conditionally Heteroscedastic Time Series Models. Lecture Notes in Statistics 18. Springer, Berlin.

[44] Terasvirta, T. and Zhao, Z., 2011. Stylized facts of return series, robust estimates and three popular models of volatility. Applied Financial Economics, 21 (1-2), 67-94.

[45] Tucker, A. L., Pond, L., 1998. The Probability Distribution of Foreign Exchange Price Changes: Tests of Candidate Processes. Review of Economics and Statistics, 11, 638-647.

[46] Vuong, Q.H., 1989. Likelihood Ratio Tests for Model Selection and Non-nested Hypotheses. Econometrica, 57, 307-333.

[47] Wintenberger, O., 2012. Continuous Invertibility and Stable QML Estimation of EGARCH(1,1) Model. Working Paper.

[48] Wu, G., 2001. The determinants of asymmetric volatility. Review of Financial Studies, 14, 837 859.

Figures and tables

Method	Time	Distribution RMSE	Volatility RMSE	Total RMSE
ML	12.02	0.64	3.99	4.04
QML	6.16	1.21	3.45	3.65
REC2	21.81	2.08	11.47	11.66
REC3	27.1	2.23	10.62	10.86
REC4	31.9	2.15	9.51	9.75
REC5	36.28	1.98	8.26	8.49
REC6	40.18	1.66	6.9	7.09
REC7	43.63	1.53	5.89	6.08
REC8	47.08	1.19	4.29	4.45
REC9	50.77	1.15	3.91	4.08
REC10	54.49	0.8	2.54	2.66

Table 1: **Comparison of the estimation methodology across the two criteria in the case of the GH-APARCH model**

The model specifications are the following:

– For the APARCH model:

$$h_t^\delta = a_0 + a_1 \left(|z_{t-1}| - \gamma z_{t-1}\right)^\delta + b_1 h_{t-1}^\delta, \tag{27}$$

with $\delta = 1.2$, $a_0 = 0.04$, $a_1 = 0.3$, $b_1 = 0.6$ and $\gamma = 0.75$.

– The GH distribution is given by:

$$f(x) = \frac{(\sqrt{\alpha^2 - \beta^2}/\delta)^\lambda}{\sqrt{2\pi} K_\lambda(\delta\sqrt{\alpha^2 - \beta^2})} e^{\beta(x-\mu)} \frac{K_{\lambda-1/2}\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right)}{\left(\sqrt{\delta^2 + (x-\mu)^2}/\alpha\right)^{1/2-\lambda}} \tag{28}$$

where $\alpha = 1.9$, $\beta = -0.55$, $\delta = 3.6$, $\mu = 0.55$ and $\lambda = -5.5$.

The number of simulations is 2000 and the sample size is 1500. ML stands for maximum likelihood, QML for Quasi-Maximum Likelihood and RECx for the recursive algorithm at step x. The column ”Time” is specified in terms of seconds. Distribution RMSE presents the quantity computed from equation (12), ”Volatility RMSE” presents the quantity computed from equation (11) and ”Total RMSE” presents the quantity computed from equation (10).

Method	Time	Distribution RMSE	Volatility RMSE	Total RMSE
ML	13.99	0.43	4.67	4.69
QML	6.67	0.75	2.99	3.08
REC2	25.43	1.03	10.81	10.86
REC3	29.68	1.02	9.08	9.14
REC4	34.05	0.94	7.91	7.97
REC5	38.18	0.84	6.47	6.52
REC6	42.32	0.75	5.25	5.3
REC7	46.46	0.68	4.06	4.12
REC8	50.87	0.62	3.4	3.45
REC9	55.55	0.58	2.67	2.73
REC10	60.54	0.55	2.09	2.16

Table 2: **Comparison of the estimation methodology across the two criteria in the case of the MN-APARCH model**

The model specifications are the following:

– For the APARCH model:

$$h_t^\delta = a_0 + a_1 \left(|z_{t-1}| - \gamma z_{t-1}\right)^\delta + b_1 h_{t-1}^\delta, \tag{29}$$

with $\delta = 1.2$, $a_0 = 0.04$, $a_1 = 0.3$, $b_1 = 0.6$ and $\gamma = 0.75$.

– The MN density is given by:

$$f(x) = \phi f(x, \mu_1, \sigma_1) + (1 - \phi) f(x, \mu_2, \sigma_2), \tag{30}$$

where $f(., \mu_i, \sigma_i)$ is the density of a Gaussian random variable with expectation μ_i and standard deviation σ_i. The parameters selected are: $\phi = 0.23$, $\mu_1 = -0.4$, $\sigma_1 = 1.3$, $\mu_2 = 0.12$ and $\sigma_2 = 0.86$.

The number of simulations is 2000 and the sample size is 1500. ML stands for maximum likelihood, QML for Quasi-Maximum Likelihood and RECx for the recursive algorithm at step x. The column ”Time” is specified in terms of seconds. Distribution RMSE presents the quantity computed from equation (12), ”Volatility RMSE” presents the quantity computed from equation (11) and ”Total RMSE” presents the quantity computed from equation (10).

Method	Time	Distribution RMSE	Volatility RMSE	Total RMSE
ML	9.82	0.76	3.04	3.14
QML	5.97	0.99	2.96	3.12
REC2	14.34	1.15	5.12	5.25
REC3	26.32	1.01	8.26	8.32
REC4	33.56	0.91	6.13	6.2
REC5	47.04	0.85	7.93	7.98
REC6	49.47	0.77	9.43	9.47
REC7	52.05	0.73	7.35	7.38
REC8	54.95	0.7	7.37	7.41
REC9	58.09	0.68	4.64	4.69
REC10	62.14	0.66	3.58	3.64

Table 3: **Comparison of the estimation methodology across the two criteria in the case of the GH-EGARCH model**

The model specifications are the following:

– For the EGARCH model:

$$\log(h_t) = a_0 + a_1|z_{t-1}| + \gamma z_{t-1} + b_1 \log(h_{t-1}), \tag{31}$$

with $a_0 = -0.4$, $a_1 = 0.1$, $b_1 = 0.96$ and $\gamma = -0.1$.

– The GH distribution is given by:

$$f(x) = \frac{(\sqrt{\alpha^2 - \beta^2}/\delta)^{\lambda}}{\sqrt{2\pi}K_{\lambda}(\delta\sqrt{\alpha^2 - \beta^2})} e^{\beta(x-\mu)} \frac{K_{\lambda-1/2}\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right)}{\left(\sqrt{\delta^2 + (x-\mu)^2}/\alpha\right)^{1/2-\lambda}} \tag{32}$$

where $\alpha = 1.9$, $\beta = -0.55$, $\delta = 3.6$, $\mu = 0.55$ and $\lambda = -5.5$.

The number of simulations is 2000 and the sample size is 1500. ML stands for maximum likelihood, QML for Quasi-Maximum Likelihood and RECx for the recursive algorithm at step x. The column "Time" is specified in terms of seconds. Distribution RMSE presents the quantity computed from equation (12), "Volatility RMSE" presents the quantity computed from equation (11) and "Total RMSE" presents the quantity computed from equation (10).

Method	Time	Distribution RMSE	Volatility RMSE	Total RMSE
ML	7.91	0.55	1.68	1.77
QML	7.15	0.92	3.02	3.16
REC2	8.5	3.4	3.14	4.63
REC3	14.05	3.26	10.44	10.94
REC4	19.38	2.63	10.51	10.84
REC5	21.27	2.3	10.24	10.49
REC6	25.91	1.96	8.72	8.94
REC7	28.85	1.31	8.73	8.83
REC8	33.26	1.13	8.64	8.71
REC9	37.67	0.94	1.36	1.65
REC10	42.03	0.76	1.15	1.38

Table 4: **Comparison of the estimation methodology across the two criteria in the case of the MN-EGARCH model**

The model specifications are the following:

- For the EGARCH model:

$$\log(h_t) = a_0 + a_1|z_{t-1}| + \gamma z_{t-1} + b_1 \log(h_{t-1}), \tag{33}$$

  with $a_0 = -0.4$, $a_1 = 0.1$, $b_1 = 0.96$ and $\gamma = -0.1$.

- The MN density is given by:

$$f(x) = \phi f(x, \mu_1, \sigma_1) + (1 - \phi) f(x, \mu_2, \sigma_2), \tag{34}$$

  where $f(., \mu_i, \sigma_i)$ is the density of a Gaussian random variable with expectation μ_i and standard deviation σ_i. The parameters selected are: $\phi = 0.23$, $\mu_1 = -0.4$, $\sigma_1 = 1.3$, $\mu_2 = 0.12$ and $\sigma_2 = 0.86$.

The number of simulations is 2000 and the sample size is 1500. ML stands for maximum likelihood, QML for Quasi-Maximum Likelihood and RECx for the recursive algorithm at step x. The column "Time" is specified in terms of seconds. Distribution RMSE presents the quantity computed from equation (12), "Volatility RMSE" presents the quantity computed from equation (11) and "Total RMSE" presents the quantity computed from equation (10).

	Number of observations	Mean	Standard Deviation	Skewness	Kurtosis
S&P 500 from 02/01/87 to 20/07/11	6190	0.027	0.190	-1.341	29.075
S&P 500 from 02/01/87 to 31/12/98	3034	0.052	0.164	-3.936	84.788
S&P 500 from 02/01/99 to 20/07/11	3156	0.002	0.212	-0.115	7.412

Table 5: **Descriptive statistics for the S&P 500 log returns.**

This table presents the descriptive statistics for the three samples of the S&P 500 log returns considered in the paper. The average and standard error statistics are annualized

From January 2, 1987 to December 31, 1998									
	α	β	δ	μ	λ	a_0	γ	a_1	b_1
QML	0.152	-0.151	1.793	0.131	-2.809	-0.384	-0.099	0.164	0.971
ML	0.134	-0.134	1.754	0.124	-2.647	-0.354	-0.072	0.138	0.973
REC	0.164	-0.163	1.836	0.142	-2.885	-0.222	-0.058	0.118	0.985
From January 2, 1999 to July 20, 2011									
	α	β	δ	μ	λ	a_0	γ	a_1	b_1
QML	0.960	-0.959	4.130	0.857	-10.556	-0.226	-0.119	0.093	0.983
ML	2.330	-0.333	0.152	0.480	3.826	-0.212	-0.102	0.076	0.985
REC	1.026	-1.025	3.969	0.914	-9.824	-0.198	-0.128	0.090	0.986
From January 2, 1987 to July 20, 2011									
	α	β	δ	μ	λ	a_0	γ	a_1	b_1
QML	0.245	-0.244	2.182	0.224	-3.569	-0.501	-0.132	0.198	0.962
ML	0.793	-0.126	1.673	0.119	-2.369	-0.230	-0.107	0.133	0.984
REC	0.265	-0.264	2.246	0.243	-3.744	-0.215	-0.094	0.118	0.986

Table 6: **Estimated parameters for the EGARCH - GH model using the S&P 500 data set**

The returns computed are logarithmic returns. ML stands for maximum likelihood, QML stands for Quasi Maximum Likelihood and REC stands for Recursive Likelihood. The model specifications are the following:

– For the EGARCH model:

$$\log(h_t) = a_0 + a_1|z_{t-1}| + \gamma z_{t-1} + b_1 \log(h_{t-1}), \tag{35}$$

– The GH distribution is given by:

$$f(x) = \frac{(\sqrt{\alpha^2 - \beta^2}/\delta)^\lambda}{\sqrt{2\pi} K_\lambda(\delta\sqrt{\alpha^2 - \beta^2})} e^{\beta(x-\mu)} \frac{K_{\lambda-1/2}\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right)}{\left(\sqrt{\delta^2 + (x-\mu)^2}/\alpha\right)^{1/2-\lambda}} \tag{36}$$

From January 2, 1987 to December 31, 1998									
	Φ	μ_1	σ_1	μ_2	σ_2	a_0	γ	a_1	b_1
QML	0.055	-0.735	2.142	0.038	0.828	-0.384	-0.098	0.164	0.971
ML	0.055	-0.803	2.215	0.039	0.865	-0.276	-0.067	0.134	0.981
REC	0.047	-0.863	2.228	0.038	0.836	-0.275	-0.068	0.138	0.981
From January 2, 1999 to July 20, 2011									
	Φ	μ_1	σ_1	μ_2	σ_2	a_0	γ	a_1	b_1
QML	0.997	0.012	0.991	-2.786	0.074	-0.225	-0.119	0.093	0.983
ML	0.544	-0.249	2.120	0.394	1.552	-0.538	-0.073	0.072	0.958
REC	0.922	0.008	0.912	-0.998	1.474	-0.200	-0.130	0.088	0.985
From January 2, 1987 to July 20, 2011									
	Φ	μ_1	σ_1	μ_2	σ_2	a_0	γ	a_1	b_1
QML	0.082	-0.631	1.766	0.053	0.856	-0.501	-0.132	0.198	0.961
ML	0.031	-1.273	2.960	0.043	1.148	-0.232	-0.076	0.101	0.986
REC	0.046	-0.932	1.954	0.044	0.882	-0.232	-0.100	0.124	0.984

Table 7: **Estimated parameters for the EGARCH - MN model using the S&P 500 data set**

The returns computed are logarithmic returns. ML stands for maximum likelihood, QML stands for Quasi Maximum Likelihood and REC stands for Recursive Likelihood. The model specifications are the following:

– For the EGARCH model:

$$\log(h_t) = a_0 + a_1|z_{t-1}| + \gamma z_{t-1} + b_1 \log(h_{t-1}), \tag{37}$$

– The MN density is given by:

$$f(x) = \phi f(x, \mu_1, \sigma_1) + (1 - \phi) f(x, \mu_2, \sigma_2), \tag{38}$$

where $f(., \mu_i, \sigma_i)$ is the density of a Gaussian random variable with expectation μ_i and standard deviation σ_i.

From January 2, 1987 to December 31, 1998										
	α	β	$\delta_{distrib.}$	μ	λ	a_0	a_1	γ	b_1	δ_{vol}
QML	0.135	-0.134	1.821	0.123	-2.749	0.000	0.094	0.599	0.912	1.040
ML	0.137	-0.135	1.990	0.188	-2.873	0.000	0.083	0.500	0.905	1.102
REC	0.137	-0.136	1.824	0.125	-2.758	0.000	0.091	0.597	0.907	1.037
From January 2, 1999 to July 20, 2011										
	α	β	$\delta_{distrib.}$	μ	λ	a_0	a_1	γ	b_1	δ_{vol}
QML	0.769	-0.763	3.908	0.698	-9.343	0.000	0.072	0.980	0.910	1.231
ML	0.711	-0.54	3.779	0.675	-6.553	0.000	0.049	0.998	0.931	1.246
REC	0.642	-0.641	3.337	0.583	-7.056	0.000	0.068	0.980	0.910	1.404
From January 2, 1987 to July 20, 2011										
	α	β	$\delta_{distrib.}$	μ	λ	a_0	a_1	γ	b_1	δ_{vol}
QML	0.267	-0.266	2.308	0.251	-3.811	0.000	0.076	0.821	0.924	1.205
ML	0.265	-0.264	2.363	0.281	-3.846	0.000	0.063	0.799	0.933	1.150
REC	0.270	-0.269	2.316	0.253	-3.846	0.000	0.069	0.813	0.922	1.200

Table 8: **Estimated parameters for the APARCH-GH model using the SP500 dataset.**

The returns computed are logarithmic returns. ML stands for maximum likelihood, QML stands for Quasi Maximum Likelihood and REC stands for Recursive Likelihood. The model specifications are the following:

– For the APARCH model:

$$h_t^{\delta} = a_0 + a_1 \left(|z_{t-1}| - \gamma z_{t-1}\right)^{\delta} + b_1 h_{t-1}^{\delta}, \tag{39}$$

– The GH distribution is given by:

$$f(x) = \frac{(\sqrt{\alpha^2 - \beta^2}/\delta)^{\lambda}}{\sqrt{2\pi} K_{\lambda}(\delta\sqrt{\alpha^2 - \beta^2})} e^{\beta(x-\mu)} \frac{K_{\lambda-1/2}\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right)}{\left(\sqrt{\delta^2 + (x-\mu)^2}/\alpha\right)^{1/2-\lambda}} \tag{40}$$

					From January 2, 1987 to December 31, 1998					
	Φ	μ_1	σ_1	μ_2	σ_2	a_0	a_1	γ	b_1	δ_{vol}
QML	0.059	-0.689	2.155	0.039	0.848	0.000	0.094	0.599	0.912	1.040
ML	0.061	-0.640	2.084	0.091	0.828	0.000	0.092	0.490	0.905	1.129
REC	0.940	0.039	0.845	-0.688	2.149	0.000	0.094	0.597	0.906	1.037
					From January 2, 1999 to July 20, 2011					
	Φ	μ_1	σ_1	μ_2	σ_2	a_0	a_1	γ	b_1	δ_{vol}
QML	0.997	0.008	0.995	-2.260	0.196	0.000	0.072	0.980	0.910	1.231
ML	0.996	0.004	0.981	-2.815	4.832	0.000	0.052	0.999	0.934	1.337
REC	0.906	0.099	0.886	-0.855	1.340	0.000	0.071	0.580	0.910	1.353
					From January 2, 1987 to July 20, 2011					
	Φ	μ_1	σ_1	μ_2	σ_2	a_0	a_1	γ	b_1	δ_{vol}
QML	0.056	-0.825	1.905	0.047	0.890	0.000	0.076	0.821	0.924	1.205
ML	0.068	-0.751	2.037	0.080	0.963	0.001	0.064	0.813	0.922	1.165
REC	0.942	0.047	0.888	-0.813	1.903	0.000	0.071	0.814	0.921	1.198

Table 9: **Estimated parameters for the APARCH - MN model using the S&P 500 data set**

The returns computed are logarithmic returns. ML stands for maximum likelihood, QML stands for Quasi Maximum Likelihood and REC stands for Recursive Likelihood. The model specifications are the following:

– For the APARCH model:

$$h_t^\delta = a_0 + a_1 \left(|z_{t-1}| - \gamma z_{t-1}\right)^\delta + b_1 h_{t-1}^\delta, \tag{41}$$

– The MN density is given by:

$$f(x) = \phi f(x, \mu_1, \sigma_1) + (1 - \phi) f(x, \mu_2, \sigma_2), \tag{42}$$

where $f(., \mu_i, \sigma_i)$ is the density of a Gaussian random variable with expectation μ_i and standard deviation σ_i.

Data set	From January 2, 1987 to December 31, 1998		
Models	**ML vs. QML**	**REC vs. QML**	**REC vs. ML**
EGARCH-GH	2.711	5.790	3.079
APARCH-GH	7.235	20.33	10.06
EGARCH-MN	3.133	3.116	-1.679
APARCH-MN	4.930	16.41	16.90
Data set	From January 2, 1999 to July 20, 2011		
Models	**ML vs. QML**	**REC vs. QML**	**REC vs. ML**
EGARCH-GH	10.16	10.28	1.223
APARCH-GH	-3.175	84.35	12.21
EGARCH-MN	38.78	56.44	47.65
APARCH-MN	2.042	16.34	14.30
Data set	From January 2, 1987 to July 20, 2011		
Models	**ML vs. QML**	**REC vs. QML**	**REC vs. ML**
EGARCH-GH	15.64	20.75	5.105
APARCH-GH	-2.028	14.91	4.813
EGARCH-MN	18.76	19.66	0.904
APARCH-MN	1.927	36.68	3.688

Table 10: **Density tests for the estimated models.**

This table presents the density tests for the estimated models. The returns computed are logarithmic returns. ML stands for maximum likelihood, QML stands for Quasi Maximum Likelihood and REC stands for Recursive Likelihood. When comparing the accuracy of model 1 (with parameters θ_1) vs. model 2 (with parameters θ_2) to fit the joint distribution of a given sample, the test statistic is computed as follows:

$$t_{1,2} = \frac{1}{n} \sum_{i=1}^{n} \Big(\log f_1(Y_t | \underline{Y_{t-1}}, \theta_1) - \log f_2(Y_t | \underline{Y_{t-1}}, \theta_2) \Big), \qquad (43)$$

with $f_{.}(.)$ the selected conditional density. The test reads as follows: in the data set starting from January 2, 1987 to December 31, 1998, in the EGARCH-GH case, the column "ML vs. QML" uses the estimated parameters by ML as model 1 and the parameters estimated by QML as model 2. The test statistics value is 2.711: this value being outside the [-1.96 : 1.96] 5% interval confidence, the null hypothesis that both models are equivalent is strongly rejected. The positivity of this statistics indicates that model 1 is favored over model 2.

Data set	From January 2, 1987 to December 31, 1998					
Models	EGARCH-GH vs. APARCH-GH	EGARCH-GH vs. EGARCH-MN	EGARCH-GH vs. APARCH-MN	APARCH-GH vs. EGARCH-MN	APARCH-GH vs. APARCH-MN	EGARCH-MN vs. APARCH-MN
REC	0.848	0.407	0.424	-0.440	-0.423	0.017
ML	2.067	0.808	0.064	-5.102	-1.990	0.056
QML	1.331	0.411	0.201	-1.330	-1.329	2.212
Data set	From January 2, 1999 to July 20, 2011					
Models	EGARCH-GH vs. APARCH-GH	EGARCH-GH vs. EGARCH-MN	EGARCH-GH vs. APARCH-MN	APARCH-GH vs. EGARCH-MN	APARCH-GH vs. APARCH-MN	EGARCH-MN vs. APARCH-MN
REC	0.413	0.065	0.068	-0.708	-0.065	0.309
ML	2.562	0.691	0.242	-3.403	-4.290	0.041
QML	1.977	0.099	0.411	-1.977	-1.990	0.047
Data set	From January 2, 1987 to July 20, 2011					
Models	EGARCH-GH vs. APARCH-GH	EGARCH-GH vs. EGARCH-MN	EGARCH-GH vs. APARCH-MN	APARCH-GH vs. EGARCH-MN	APARCH-GH vs. APARCH-MN	EGARCH-MN vs. APARCH-MN
REC	0.113	0.042	0.032	-0.109	-0.111	-0.071
ML	6.480	0.048	0.095	-3.198	-4.205	0.090
QML	1.068	0.029	-4.004	-1.067	-1.068	-0.335

Table 11: **Density tests testing model domination**

This table presents the density tests testing model domination. The returns computed are logarithmic returns. ML stands for maximum likelihood, QML stands for Quasi Maximum Likelihood and REC stands for Recursive Likelihood. When comparing the accuracy of model 1 (with parameters θ_1) vs. model 2 (with parameters θ_2) to fit the joint distribution of a given sample, the test statistic is computed as follows:

$$t_{1,2} = \frac{1}{n}\sum_{t=1}^{n}\left(\log f_1(Y_t|\underline{Y_{t-1}}, \theta_1) - \log f_2(Y_t|\underline{Y_{t-1}}, \theta_2)\right), \tag{44}$$

with $f_.(.)$ the selected conditional density. The test reads as follows: in the data set starting from January 2, 1987 to December 31, 1998, using the REC estimation approach, when the EGARCH-GH is model 1 and the EGARCH-MN is model 2, the test statistics value is 0.407: this value being inside the [-1.96 : 1.96] 5% interval confidence, the null hypothesis that both models are equivalent is accepted.

EGARCH-GH

α	β	δ	μ	λ	α_0	γ	α_1	β_1	s	r	
0.164	-0.163	1.836	0.142	-2.885	-0.222	-0.058	0.118	0.985	-1.112		
(0.075)	(0.053)	(0.267)	(0.048)	(0.350)	(0.062)	(0.013)	(0.017)	(0.006)			
0.259	-0.136	1.719	0.128	-2.484	-0.174	0	0.113	0.990	-1.481	0.249	
(0.046)	(0.032)	(0.053)	(0.025)	(0.662)	(0.083)		(0.024)	(0.007)			

EGARCH-MN

Φ	μ_1	σ_1	μ_2	σ_2	α_0	γ	α_1	β_1	s	r	
0.047	-0.863	2.228	0.038	0.836	-0.275	-0.068	0.138	0.981	-1.034		
(0.009)	(0.251)	(1.012)	(0.012)	(0.325)	(0.061)	(0.013)	(0.019)	(0.005)			
0.051	-0.852	2.371	0.042	0.851	-0.218	0	0.139	0.987	-1.328	0.221	
(0.012)	(0.247)	(1.322)	(0.015)	(0.421)	(0.046)		(0.015)	(0.006)			

APARCH-GH

α	β	δ	μ	λ	α_0	α_1	γ	β_1	δ_{vol}	s	r
0.137	-0.136	1.824	0.125	-2.758	0.000	0.091	0.597	0.907	1.037	-0.884	
(0.062)	(0.035)	(0.552)	(0.046)	(0.163)	(0.0001)	(0.018)	(0.103)	(0.022)	(0.080)		
0.158	-0.157	1.807	0.146	-2.721	0.000	0.161	0	0.944	1.252	-1.417	0.376
(0.022)	(0.061)	(0.609)	(0.039)	(0.745)	(0.0001)	(0.009)		(0.008)	(0.067)		

APARCH-MN

Φ	μ_1	σ_1	μ_2	σ_2	α_0	α_1	γ	β_1	δ_{vol}	s	r
0.940	0.039	0.845	-0.688	2.149	0.000	0.094	0.597	0.906	1.037	-0.873	
(0.021)	(0.008)	(0.213)	(0.087)	(1.052)	(0.000)	(0.011)	(0.082)	(0.012)	(0.120)		
0.933	0.048	0.815	-0.680	2.285	0.000	0.119	0	0.890	1.438	-1.038	0.159
(0.028)	(0.016)	(0.201)	(0.085)	(1.112)	(0.000)	(0.013)		(0.014)	(0.173)		

Table 12: **Recursively estimated parameters with and without leverage effect using the S&P 500 data set from January 2, 1987 to December 31, 1998 and their standard errors.** $r = \frac{s_{\gamma=0} - s_{\gamma\neq 0}}{s_{\gamma=0}}$ **where** $s_{\gamma\neq 0}$ **and** $s_{\gamma=0}$ **are the skewness of the residual with and without leverage effect respectively.**

EGARCH-GH

α	β	δ	μ	λ	α_0	γ	α_1	β_1	s	r	
1.026	-1.025	3.969	0.914	-9.824	-0.198	-0.128	0.090	0.986	-0.380		
(0.226)	(0.165)	(0.387)	(0.029)	(1.471)	(0.025)	(0.011)	(0.012)	(0.002)			
1.834	-0.299	1.536	0.191	-0.006	-0.221	0	0.172	0.990	-0.453	0.161	
(0.172)	(0.069)	(0.006)	(0.113)	(0.054)	(0.035)		(0.018)	(0.003)			

EGARCH-MN

Φ	μ_1	σ_1	μ_2	σ_2	α_0	γ	α_1	β_1	s	r	
0.922	0.008	0.912	-0.998	1.474	-0.200	-0.130	0.088	0.985	-0.381		
(0.006)	(0.002)	(0.251)	(0.091)	(0.281)	(0.056)	(0.018)	(0.023)	(0.004)			
0.908	0.0003	0.914	-0.729	1.514	-0.221	0	0.171	0.990	-0.453	0.159	
(0.015)	(0.000)	(0.321)	(0.109)	(0.360)	(0.041)		(0.019)	(0.006)			

APARCH-GH

α	β	δ	μ	λ	α_0	α_1	γ	β_1	δ_{vol}	s	r
0.642	-0.641	3.337	0.583	-7.056	0.000	0.068	0.980	0.910	1.404	-0.345	
(0.054)	(0.029)	(0.954)	(0.096)	(1.203)	(0.000)	(0.003)	(0.008)	(0.019)	(0.061)		
1.881	-0.307	1.559	0.278	-0.006	0.000	0.072	0	0.918	2.180	-0.402	0.140
(0.254)	(0.075)	(0.421)	(0.032)	(0.001)	(0.000)	(0.011)		(0.022)	(0.059)		

APARCH-MN

Φ	μ_1	σ_1	μ_2	σ_2	α_0	α_1	γ	β_1	δ_{vol}	s	r
0.906	0.099	0.886	-0.855	1.340	0.000	0.071	0.580	0.910	1.353	-0.315	
(0.016)	(0.009)	(0.213)	(0.021)	(0.213)	(0.000)	(0.007)	(0.027)	(0.007)	(0.165)		
0.989	0.003	0.977	-2.33	0.171	0.000	0.059	0	0.918	2.511	-0.397	0.206
(0.008)	(0.000)	(0.139)	(1.029)	(0.249)	(0.000)	(0.011)		(0.009)	(0.310)		

Table 13: **Recursively estimated parameters with and without leverage effect using S&P 500 data set from January 2, 1999 to July 20, 2011 and their standard errors.** $r = \frac{s_{\gamma=0} - s_{\gamma\neq0}}{s_{\gamma=0}}$ **where** $s_{\gamma\neq0}$ **and** $s_{\gamma=0}$ **are the skewness of the residual with and without leverage effect respectively.**

EGARCH-GH

α	β	δ	μ	λ	α_0	γ	α_1	β_1	s	r	
0.265	-0.264	2.246	0.243	-3.744	-0.215	-0.094	0.118	0.986	-0.690		
(0.061)	(0.051)	(0.319)	(0.070)	(0.251)	(0.026)	(0.009)	(0.011)	(0.002)			
1.643	-0.183	0.937	0.179	0.537	-0.192	0	0.147	0.991	-0.818	0.156	
(0.215)	(0.096)	(0.025)	(0.127)	(0.056)	(0.028)		(0.013)	(0.002)			

EGARCH-MN

Φ	μ_1	σ_1	μ_2	σ_2	α_0	γ	α_1	β_1	s	r	
0.046	-0.932	1.954	0.044	0.882	-0.232	-0.100	0.124	0.984	-0.678		
(0.011)	(0.451)	(0.652)	(0.008)	(0.106)	(0.028)	(0.015)	(0.017)	(0.003)			
0.063	-0.715	1.962	0.050	0.876	-0.195	0	0.155	0.991	-0.806	0.159	
(0.020)	(0.328)	(0.649)	(0.007)	(0.112)	(0.021)		(0.023)	(0.004)			

APARCH-GH

α	β	δ	μ	λ	α_0	α_1	γ	β_1	δ_{vol}	s	r
0.270	-0.269	2.316	0.253	-3.846	0.000	0.069	0.813	0.922	1.200	-0.637	
(0.114)	(0.052)	(0.297)	(0.094)	(0.694)	(0.0001)	(0.007)	(0.100)	(0.007)	(0.028)		
0.219	-0.218	2.121	0.212	-3.340	0.000	0.073	0	0.934	1.514	-0.797	0.201
(0.026)	(0.171)	(0.382)	(0.059)	(0.612)	(0.000)	(0.007)		(0.006)	(0.051)		

APARCH-MN

Φ	μ_1	σ_1	μ_2	σ_2	α_0	α_1	γ	β_1	δ_{vol}	s	r
0.942	0.047	0.888	-0.813	1.903	0.000	0.071	0.814	0.921	1.198	-0.634	
(0.003)	(0.017)	(0.524)	(0.126)	(0.692)	(0.000)	(0.007)	(0.091)	(0.006)	(0.096)		
0.924	0.056	0.863	-0.687	1.898	0.000	0.095	0	0.907	1.659	-0.719	0.118
(0.007)	(0.029)	(0.561)	(0.123)	(0.536)	(0.000)	(0.008)		(0.007)	(0.155)		

Table 14: **Recursively estimated parameters with and without leverage effect using the S&P 500 data set from January 2, 1987 to July 20, 2011 and their standard errors.** $r = \frac{s_{\gamma=0} - s_{\gamma\neq 0}}{s_{\gamma=0}}$ **where** $s_{\gamma\neq 0}$ **and** $s_{\gamma=0}$ **are the skewness of the residual with and without leverage effect respectively.**

Data set	From January 2, 1987 to December 31, 1998					
M	0	1	2	3	4	5
EGARCH-GH	0.177	0.212	0.181	0.157	0.138	0.137
APARCH-GH	0.199	0.211	0.238	0.231	0.224	0.234
EGARCH-MN	0.174	0.164	0.150	0.162	0.179	0.178
APARCH-MN	0.330	0.323	0.295	0.304	0.275	0.268
Data set	From January 2, 1999 to July 20, 2011					
M	0	1	2	3	4	5
EGARCH-GH	0.158	0.153	0.124	0.108	0.138	0.127
APARCH-GH	0.151	0.125	0.111	0.122	0.136	0.108
EGARCH-MN	0.143	0.131	0.136	0.156	0.142	0.130
APARCH-MN	0.108	0.101	0.082	0.094	0.092	0.098
Data set	From January 2, 1987 to July 20, 2011					
M	0	1	2	3	4	5
EGARCH-GH	0.174	0.105	0.146	0.151	0.147	0.142
APARCH-GH	0.098	0.091	0.060	0.092	0.096	0.090
EGARCH-MN	0.113	0.126	0.146	0.149	0.152	0.121
APARCH-MN	0.095	0.077	0.093	0.090	0.098	0.106

Table 15: **Hansen's test for the estimated models using the S&P 500 log returns.**

This table presents the p-value of the Hansen's test comparing the models with and without leverage effect for the S&P 500 data set. The test reads as follows: for the data set starting from January 2, 1987 to December 31, 1998, for the EGARCH-GH model and for $M = 0$, its p-value is 0.177. This p-value being greater than 5%, the hypothesis H_0 (absence of leverage effect) is not rejected.

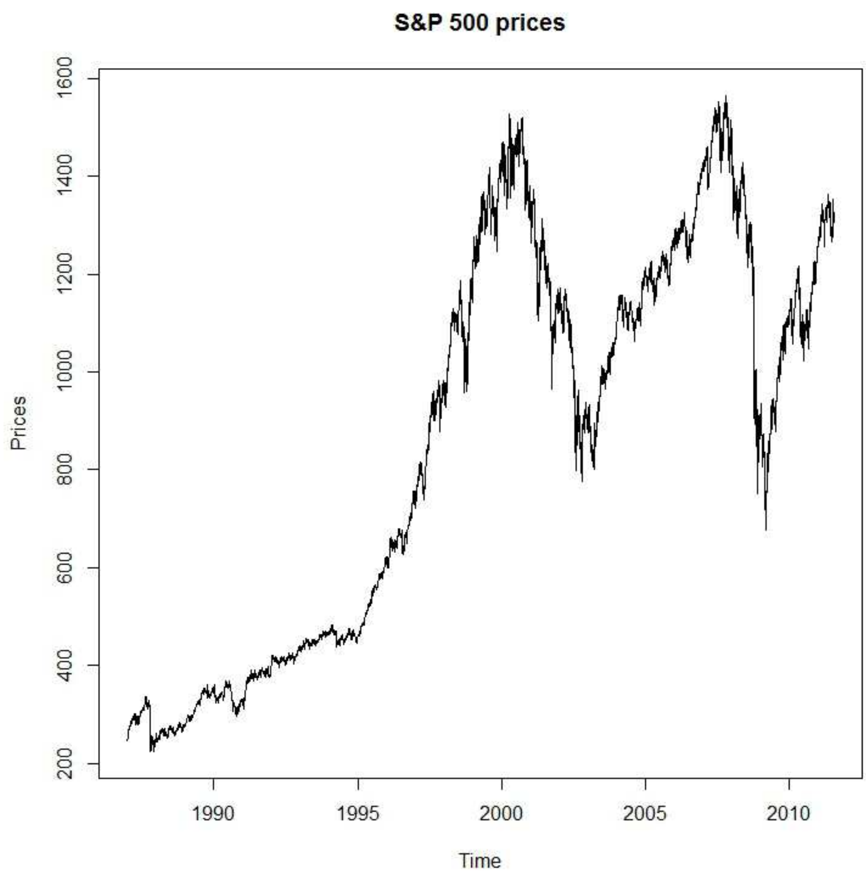

Figure 1: **S&P 500 prices. Data source: Bloomberg.**

The data set starts on January, 2^{nd} 1987 and ends on December, 31^{st} 2011. It includes 6190 data

S&P 500 index

EGARCH-GH							
Forecast Horizon	1	10	20	30	40	50	60
Unweighted ω_0	**-2.200**	**-2.245**	**-2.229**	**-2.319**	**-2.262**	**-2.264**	**-2.244**
Center ω_1	-1.306	-1.328	-1.293	-1.364	-1.318	-1.324	-1.292
Tails ω_2	**-2.594**	**-2.656**	**-2.647**	**-2.739**	**-2.678**	**-2.677**	**-2.670**
Right Tail ω_3	-1.829	-1.780	-1.806	-1.864	-1.831	-1.804	-1.854
Left Tail ω_4	-1.725	-1.787	-1.776	-1.839	-1.804	-1.821	-1.768
APARCH-GH							
Forecast Horizon	1	10	20	30	40	50	60
Unweighted ω_0	-0.133	0.097	-0.241	-0.051	0.099	0.284	-0.264
Center ω_1	-1.146	-1.097	-1.492	-1.339	-1.110	-1.139	-1.317
Tails ω_2	0.348	0.621	0.426	0.585	0.669	0.892	0.271
Right Tail ω_3	-0.417	-0.190	-0.370	-0.440	-0.030	0.269	-0.643
Left Tail ω_4	0.131	0.287	-0.044	0.251	0.162	0.172	0.111
EGARCH-MN							
Forecast Horizon	1	10	20	30	40	50	60
Unweighted ω_0	-1.276	-1.421	-1.257	-1.356	-1.436	-1.425	-1.561
Center ω_1	-1.365	-1.466	-1.237	-1.461	-1.546	-1.427	-1.710
Tails ω_2	-0.786	-0.931	-0.879	-0.814	-0.879	-0.958	-0.932
Right Tail ω_3	1.326	1.283	1.414	1.251	1.051	1.274	0.973
Left Tail ω_4	**-2.029**	**-2.125**	**-2.018**	**-2.129**	**-2.067**	**-2.192**	**-2.179**
APARCH-MN							
Forecast Horizon	1	10	20	30	40	50	60
Unweighted ω_0	0.493	1.289	0.786	1.330	1.047	1.070	0.778
Center ω_1	0.514	0.856	0.447	0.828	1.032	1.308	0.482
Tails ω_2	0.272	1.542	1.122	1.672	0.634	0.036	1.174
Right Tail ω_3	0.277	1.691	0.834	1.309	0.649	1.478	0.902
Left Tail ω_4	0.562	0.533	0.551	0.993	1.223	0.476	0.445

Table 16: **Amisano Giacomini Test Statistics using S&P 500 data set.**

This table presents Amisano Giacomini's test comparing the models with and without leverage effect using the S&P 500 data set covering the 1987-2011 period. We use rolling windows of size $\tau = 2000$ and a maximal forecast horizon k_{max} of 60 days. The test reads as follows: for the EGARCH-GH model, for horizon $k = 1$ and for unweighted ω_0, the test statistic is -2.200: this value being outside the [-1.96 : 1.96] 5% interval confidence, the null hypothesis that both models (with or without leverage) are equivalent is rejected in favor of the non leveraged one (negative value).

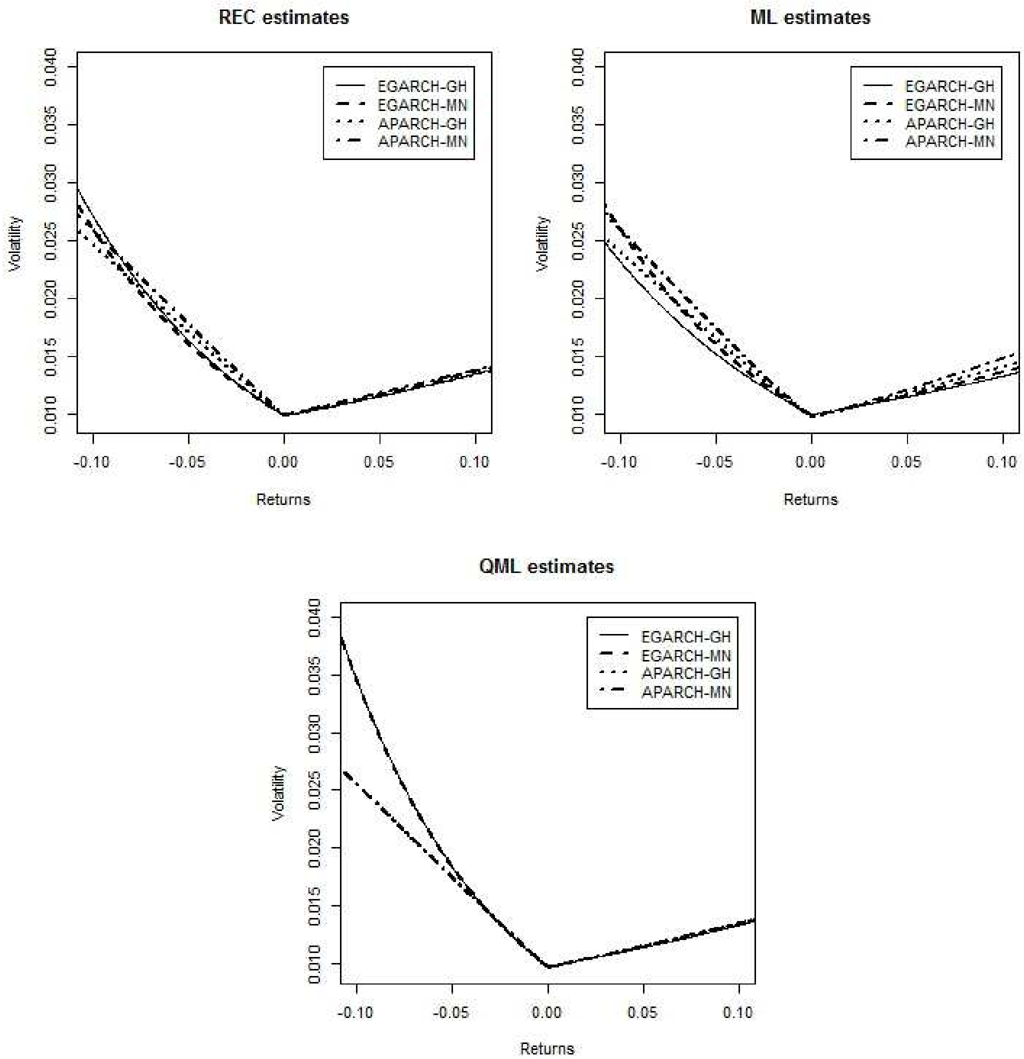

Figure 2: **Volatility News Impact curve.**

The data set used starts from January 2, 1987 to December 31, 1998. The figure presents the News Impact curve when the initial level of volatility is set to the average of the sample at hand. The parameters used are the parameters estimated with the three maximum likelihood approaches. The top-left figure is for the REC estimates, the top-right is for the ML estimates and the bottom figure is for the QML estimates.

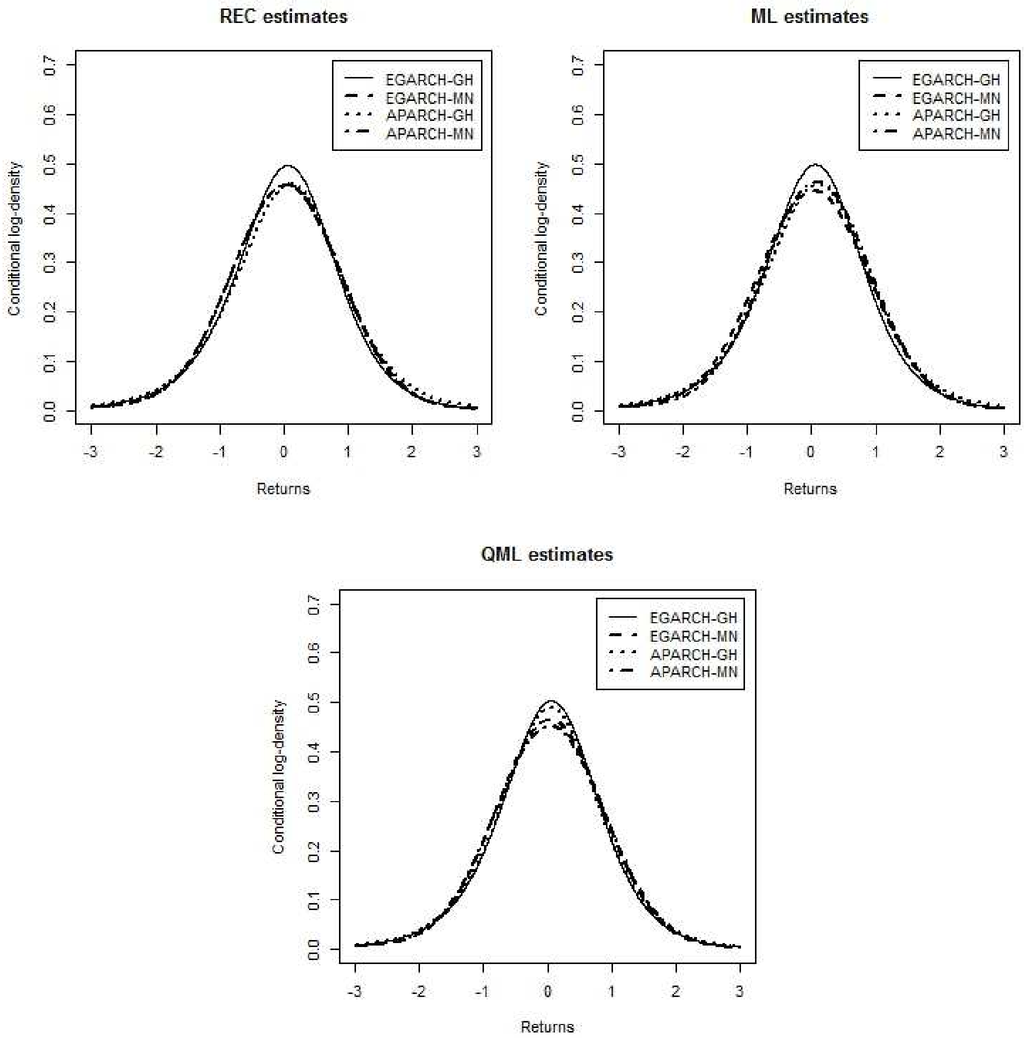

Figure 3: **Estimated conditional log-densities.**
The data set used starts from January 2, 1987 to December 31, 1998. The figure presents the estimated conditional densities. The parameters used are the parameters estimated with the three maximum likelihood approaches. The top-left figure is for the REC estimates, the top-right is for the ML estimates and the bottom figure is for the QML estimates.

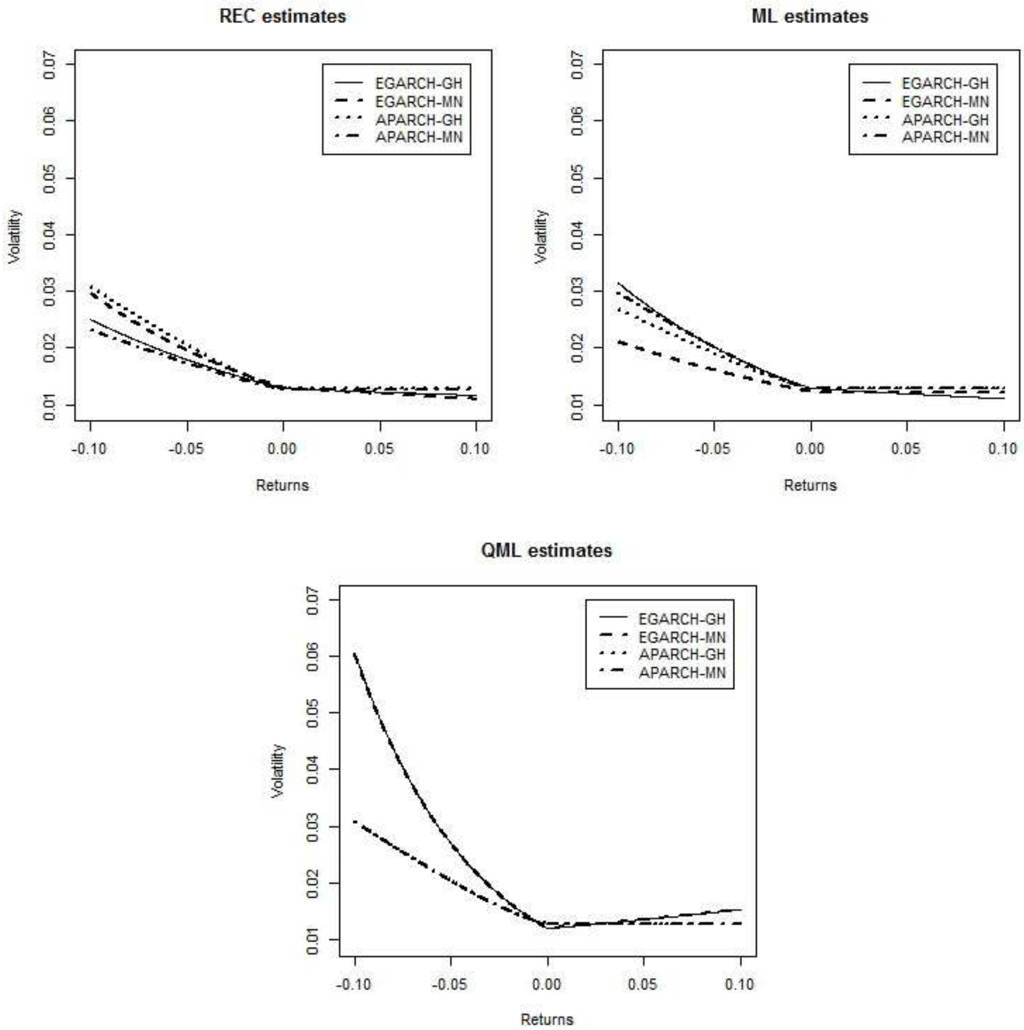

Figure 4: **Volatility News Impact curve.**

The data set used starts from January 2, 1999 to July 20, 2011. The figure presents the News Impact curve when the initial level of volatility is set to the average of the sample at hand. The parameters used are the parameters estimated with the three maximum likelihood approaches. The top-left figure is for the REC estimates, the top-right is for the ML estimates and the bottom figure is for the QML estimates.

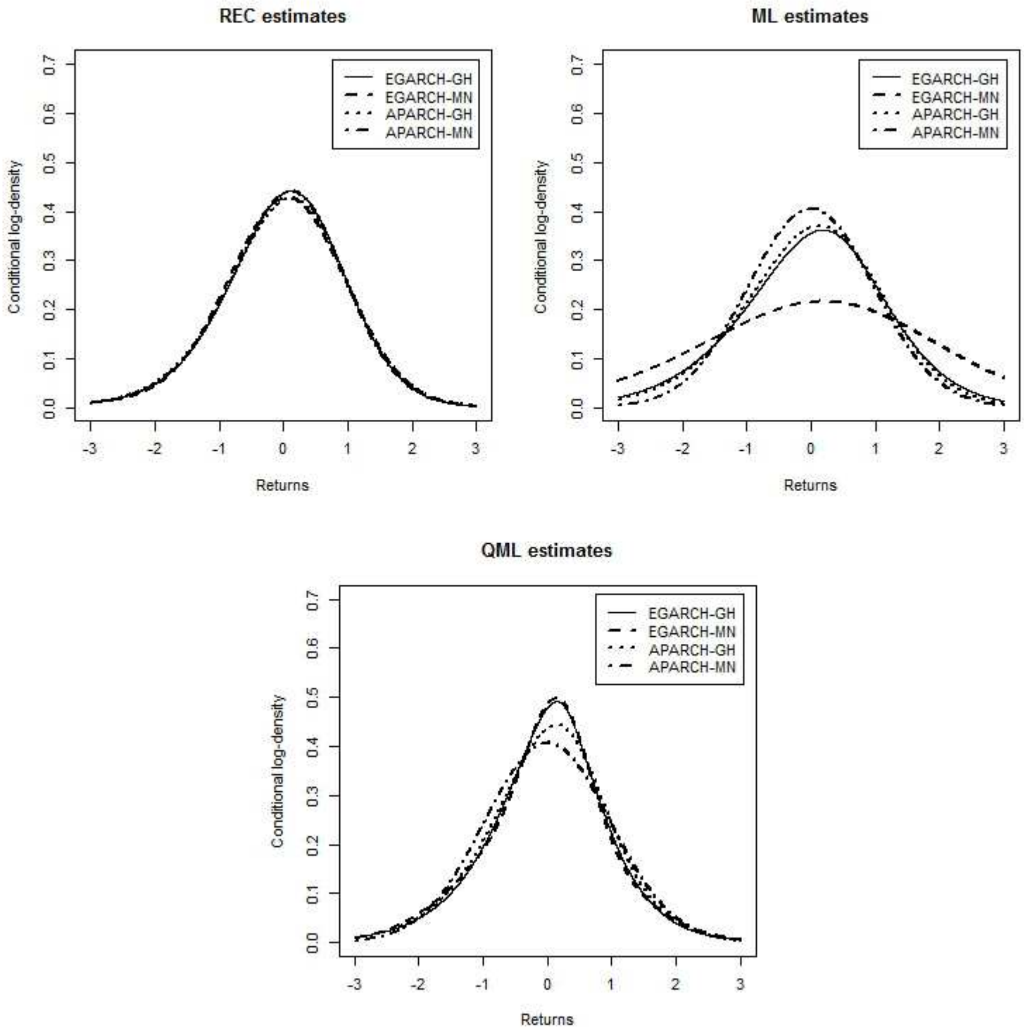

Figure 5: **Estimated conditional log-densities.**
The data set used starts from January 2, 1999 to July 20, 2011. The figure presents the estimated conditional densities. The parameters used are the parameters estimated with the three maximum likelihood approaches. The top-left figure is for the REC estimates, the top-right is for the ML estimates and the bottom figure is for the QML estimates.

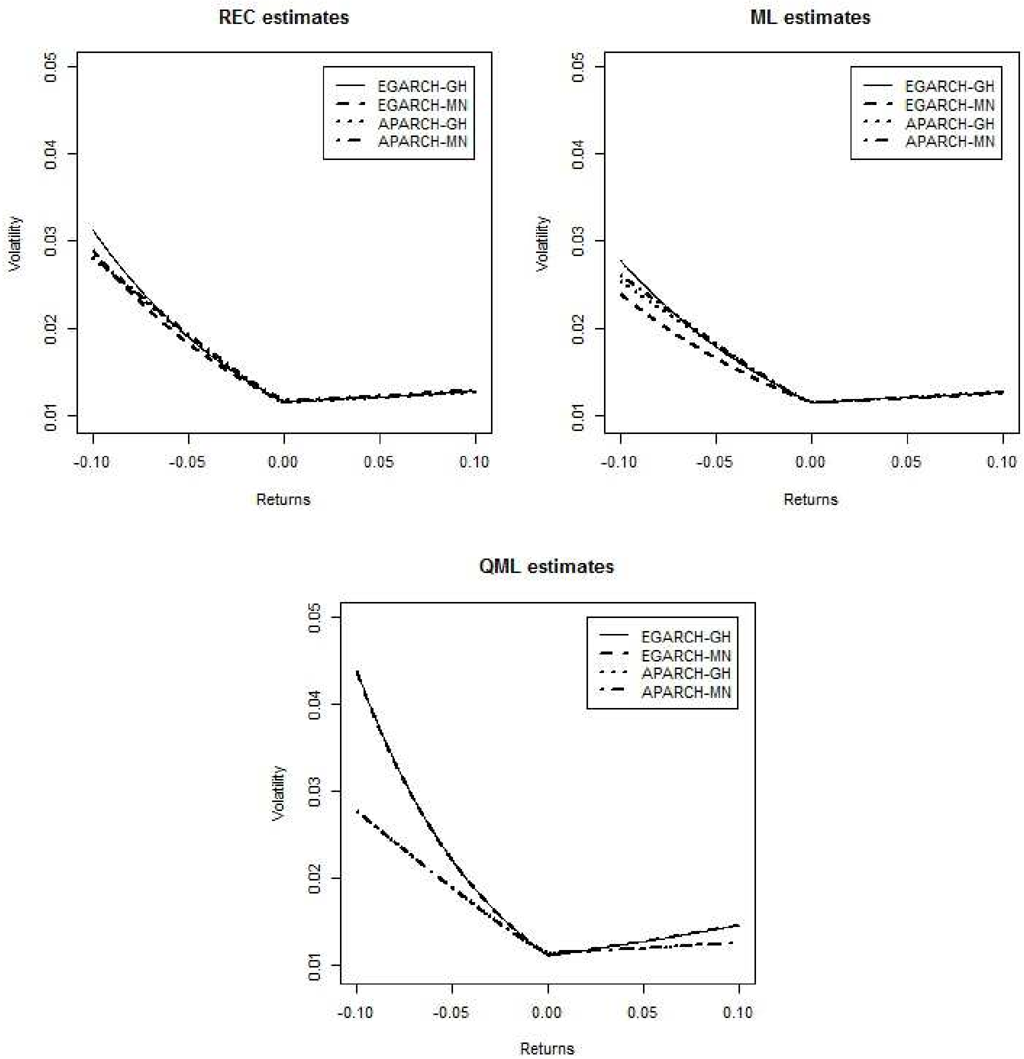

Figure 6: **Volatility News Impact curve.**

The data set used starts from January 2, 1987 to July 20, 2011. The figure presents the News Impact curve when the initial level of volatility is set to the average of the sample at hand. The parameters used are the parameters estimated with the three maximum likelihood approaches. The top-left figure is for the REC estimates, the top-right is for the ML estimates and the bottom figure is for the QML estimates.

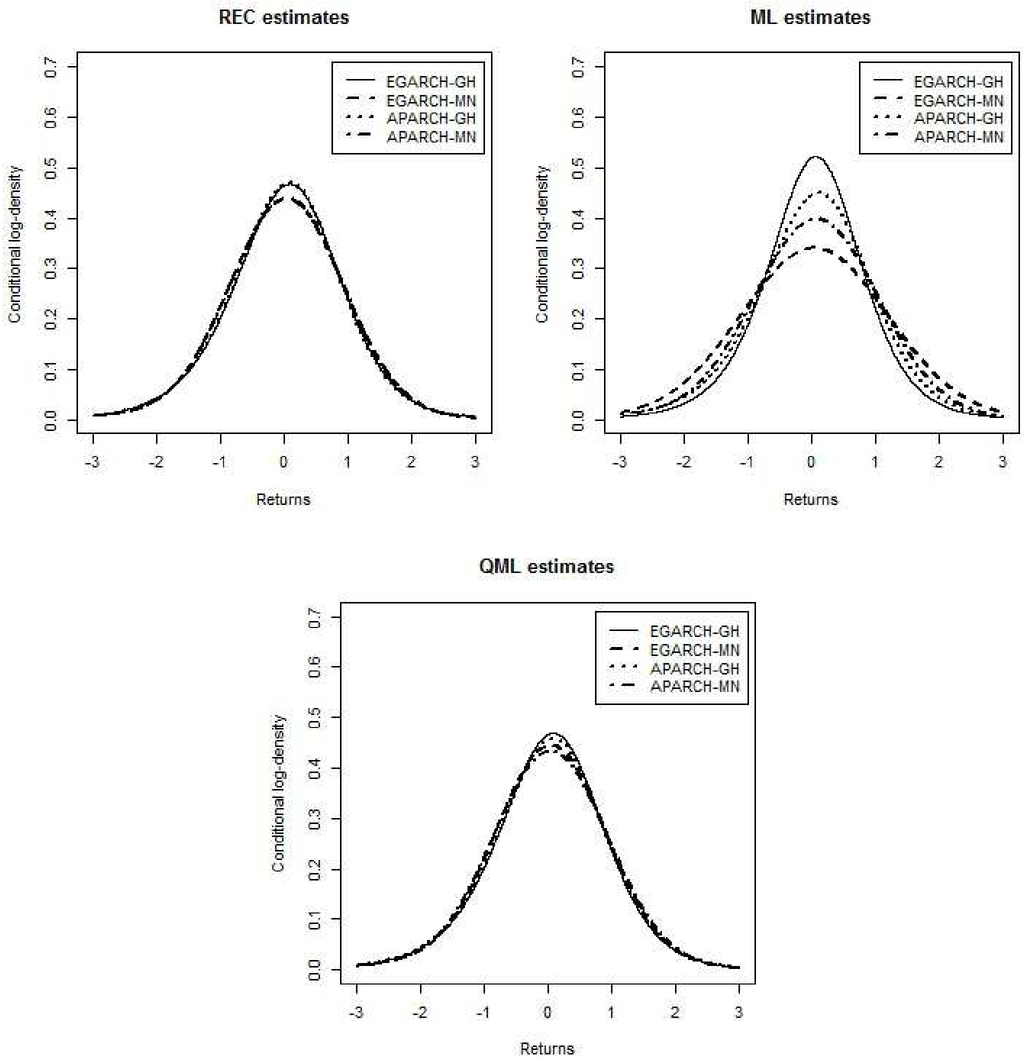

Figure 7: **Estimated conditional log-densities.**
The data set used starts from January 2, 1987 to July 20, 2011. The figure presents the estimated conditional densities. The parameters used are the parameters estimated with the three maximum likelihood approaches. The top-left figure is for the REC estimates, the top-right is for the ML estimates and the bottom figure is for the QML estimates.