CHAPITRE 3

Evaluation d'options avec un modèle GARCH

Michel LUBRANO

Avril 2011

Contents

Luc Bauwens and Michel Lubrano (2002) Bayesian option pricing using asymmetric GARCH models, *Journal of Empirical Finance* Volume 9, Issue 3, August 2002, Pages 321-342.

1 Introduction

1. The Black and Scholes (hereafter BS) formula for option pricing is a function of the parameters of the diffusion process describing the dynamics of the underlying asset price. The BS formula is derived for a geometric Brownian motion with a constant volatility parameter σ (corresponding to an hypothesis of homoskedasticity in an econometric model).

2. In order to apply the BS formula, the econometrician has to estimate the parameter σ of the diffusion process. Hull and White (1987) studied option pricing in the case of stochastic volatility where σ^2 follows an independent geometric Brownian motion. Provided the increments of the volatility process are independent of the increments of the process governing the asset price, they showed that options can be priced by averaging the BS formula over the stochastic volatility, i.e. that there is no need to price the risk attached to the stochastic volatility.

3. When relaxing the hypothesis of constant volatility, it is very difficult, if not impossible, to solve the partial differential equation which leads to the BS formula (see however the restrictive case of Hull and White 1987 cited above and also the paper of Heston and Nandi 2000).

4. One has to apply the risk-neutral pricing methodology of Cox and Ross (1976), that is to say, for evaluating the price of a call option at time t with maturity at time $T > t$, compute
$$C_t^T = e^{-r(T-t)}\,\mathrm{E}[\max(S_T - K, 0)] \tag{1}$$
where K is the exercise price, r the riskless interest rate and S_T the predicted price of the underlying asset at time T. This procedure requires to find **an equivalent martingale measure** for the returns, which leaves unchanged the volatility (in a way to be precised later) and for which the returns behave like a martingale. Duan (1995) derived the equivalent martingale measure for the case where the returns follow a discrete GARCH process with normal errors. Duan (1999) provides extensions to the case of leptokurtik errors.

The object of this paper is to investigate the interesting features that the Bayesian approach may bring in the risk-neutral methodology for option pricing. We have first to select an econometric model which captures as best as possible the well-known stylised facts of financial return series: volatility clustering, strong kurtosis, and often in the case of stocks, the leverage effect.

We have basically the choice between two families of models.

1. The stochastic volatility model in discrete time is certainly the model which is the closest to the continuous time diffusion processes considered in Hull and White (1987). Mahieu and Schotman (1998) estimated various specifications of the stochastic volatility model and applied directly the solution of Hull and White (1987) for option pricing.

2. However, when we introduce a leverage effect in this model, we allow for possible correlation between volatility increments and the increments of the process and thus violate Hull and White's assumption.

3. Moreover, Mahieu and Schotman concluded that the obtained estimates for the volatility were very imprecise. The GARCH model, which is much simpler to estimate, recovers thus all of its interest and we make it our baseline model.

Once risk neutralisation is taken into account, we have to simulate the chosen model in a way which takes into account all the available information contained in the data. The Bayesian viewpoint is particularly well suited for this requirement.

1. **The predictive density** is a natural by-product of inference that can be used for prediction as it represents the density of future observations. The risk neutral valuation requires the use of the predictive expectation of the payoff function of the option given in (1), which we call the predictive option price. We can even compute the predictive density of the payoff function.

2. In the classical framework, a point estimate of the option price can also be computed, but the Bayesian approach delivers naturally a probability distribution. This distribution integrates the uncertainty both over the parameter values of the underlying econometric model (contained in the posterior distribution of the parameters) and over the future stock price.

3. Therefore an interval estimate of any confidence level can be formed for an option price. If an agent wants to buy (or sell) an option on the market, he can gauge the market price of the option with the predictive distribution of the option price: e.g. he could decide not to buy (sell) if the market price is too far in the right (left) tail of his predictive distribution.

Let us define returns in discrete time as

$$y_t = (S_t - S_{t-1})/S_{t-1}. \tag{2}$$

The baseline GARCH(1,1) model with Gaussian errors[1]

$$\begin{cases} y_t = \mu_t + u_t \\ u_t | I_{t-1} \sim \mathrm{N}(0, h_t) \\ h_t = \omega + \alpha u_{t-1}^2 + \beta h_{t-1} \end{cases} \tag{3}$$

where μ_t represents the conditional expectation of the returns.

This simple model does not succeed completely in reproducing the stylised facts present in the data and consequently may give a wrong account of the volatility persistence.

1. The degree of persistence in the volatility process is important for option pricing since a higher persistence results in a longer delay for the predictive conditional variance to converge to its unconditional value.

2. As shown in detail by Terasvirta (1996), the introduction of Student errors improves greatly the possibility that the GARCH(1,1) model reproduces the behaviour of data with a high empirical kurtosis.

3. A second possible improvement of GARCH models for stock returns consists in taking into account the leverage effect, that is the negative correlation between current stock returns and future volatility. In the model of Glosten, Jagannathan and Runkle (1993), the conditional variance can react asymmetrically to past squared shocks in the mean process, a stronger effect for negative shocks than for positive ones corresponding to the leverage effect.

4. Finally, besides asymmetry, the news impact curve of the GARCH model may saturate for large returns. Lubrano (2001) introduces a smooth transition GARCH model which combines both asymmetry and saturation.

[1] The parameters α, β and ω are restricted to be positive. The starting value h_0 is treated as a known constant. The $\{\epsilon_t\}$ sequence is conditionally independent. I_{t-1} is the past information set.

2 Predictive option pricing: principles

2.1 Introduction

1. The terminal payoff at time T of a European call option is given by

$$P_T = \max(S_T - K, 0). \tag{4}$$

If we manage to find a risk neutral equivalent martingale measure Q to the measure P of the empirical process of the underlying security return, then an option can be priced as the discounted expected value of its future terminal payoff.

2. When the empirical process follows a geometric Brownian motion, an analytical solution to this problem is provided by the celebrated Black-Scholes formula

$$\begin{aligned} C_t^T &= S_t \Phi(d_1) - K e^{-r(T-t)} \Phi(d_2) \\ d_1 &= \frac{\ln(S_t / K e^{-r(T-t)})}{\sigma \sqrt{T-t}} + 0.5\,\sigma \sqrt{T-t} \\ d_2 &= \frac{\ln(S_t / K e^{-r(T-t)})}{\sigma \sqrt{T-t}} - 0.5\,\sigma \sqrt{T-t}, \end{aligned} \tag{5}$$

where $\Phi(.)$ is the cumulative Gaussian distribution function, S_t the observed price at time t of the underlying security, and K the exercise price or strike. The arguments depend on the volatility σ which is the standard deviation (per time unit) of the process of the return of the underlying asset.

3. Methods to estimate this parameter can be found for instance in Campbell, Lo and MacKinley (1997) as well as an extension for the case where the underlying price follows a trending Ornstein-Uhlenbeck process.

2.2 Stochastic volatility

In the case of stochastic volatility (σ indexed by t), provided volatility is uncorrelated with the security price (which excludes GARCH processes), one can follow Hull and White (1987) and average the BS formula over future volatility in a Monte Carlo simulation of N draws so as to obtain

$$C_t^T = \frac{1}{N}\sum_{j=1}^{N} BS_j(S_t, K, \sigma_j) \tag{6}$$

where

$$\sigma_j^2 = \frac{1}{T-t}\sum_{i=t+1}^{T} \sigma_{j,i}^2. \tag{7}$$

1. The main advantage of this approach is that we do not have to predict the future price S_T of the underlying asset. We have only to predict future volatility under measure P. Consequently, there is no need to find an equivalent risk neutral process to the data generating process.

2. But this approach is limited in its application as first it precludes any correlation between the price of the underlying asset and the volatility (leverage effect) and second it cannot price the risk attached to stochastic volatility.

2.3 GARCH option pricing and risk neutralisation

Duan (1995) provides a general method for option pricing in the case of GARCH processes.

1. Risk neutralisation should leave the variance unchanged and should transform the conditional expectation so that the discounted expected price of the underlying asset follows a martingale.

2. In GARCH processes, it is not possible to find a risk neutralisation procedure that leaves unchanged the marginal variance of the process or the conditional variance beyond one period.

3. Duan (1995) introduces the *Local Risk Neutral Valuation Relationship* which leaves unchanged the one period ahead conditional variance and implies that the expected future return is equal to the risk free interest rate r. Adapted to discrete time, this means that
$$\mathrm{E}[(S_T - S_{T-1})/S_{T-1}] = r. \tag{8}$$

4. The econometric model (3) specifies the empirical distribution of y_t defined in (2), conditionally on the past. The pricing measure Q shifts the error term u_t so that the conditional expectation of y_t becomes equal to r. The new error term is $v_t = u_t + \mu_t - r$ and we have
$$\begin{cases} y_t = r + v_t \\ v_t | I_{t-1} \sim \mathrm{N}(0, h_t) \\ h_t = \omega + \alpha(v_{t-1} - \mu_{t-1} + r)^2 + \beta h_{t-1}. \end{cases} \tag{9}$$

5. The functional form of the conditional variance remains unchanged. However h_t is no longer governed by a χ^2 variable as under measure P but by a non-central χ^2.

6. This type of shifting can be applied to many other GARCH processes, in particular those reviewed in the introduction which take into account asymmetry, leverage effect and saturation.

Duan (1995) used the GARCH-M model of Engle, Lilien and Robin (1987) which implies that

$$\mu_t = \mu + \lambda\sqrt{h_t}.$$

This model, which may find some justification in the financial literature, generally has not a very good fit and provides poor estimates for λ.

As noted for instance by Campbell, Lo, and Mac Kinlay (1997, Ch. 2), financial series such as stock indexes often present positive autocorrelation of the returns. Positive autocorrelation may be due to various phenomena such as infrequent and nonsynchronous trading of individual stocks entering the index. But it also may be due to the risk premium attached to nonconstant volatility. Consequently, the baseline GARCH model with normal errors we selected is

$$\begin{cases} y_t & = & \mu + \rho y_{t-1} + u_t \\ u_t|I_{t-1} & \sim & \mathrm{N}(0, h_t) \\ h_t & = & \omega + \alpha u_{t-1}^2 + \beta h_{t-1}. \end{cases} \tag{10}$$

We follow the same approach as Hafner and Herwatz (2001) who found on German securities that incorporating ρy_{t-1} in the conditional expectation gave a higher likelihood value than incorporating λh_t.

The variance of y_t is as usual equal to

$$\text{Var}(y_t) = \omega/(1-\alpha-\beta)$$

and the stationarity condition is given by $\alpha + \beta < 1$. For the risk neutralised process, we show in the Appendix that the variance is

$$\text{Var}(y_t) = \frac{\omega + \alpha[(1-\rho)r - \mu]^2}{1 - \alpha(1+\rho^2) - \beta} \tag{11}$$

The process is stationary if $\alpha(1+\rho^2) + \beta < 1$. Consequently, risk neutralisation increases the marginal variance of y, fragilises the stationarity condition and increases volatility persistence as the predictive variance will take a longer time to converge to its limiting value given in (11).

2.4 Predictive option pricing

The predictive density of the terminal payoff P_T under measure Q is defined by

$$f_Q(P_T|y) = \int f_Q(P_T|\theta, y)\, \varphi(\theta|y)\, d\theta, \tag{12}$$

where $\varphi(\theta|y)$ is the posterior density of the parameters of the econometric model (10) and $f_Q(P_T|\theta, y)$ is the density of a future payoff after translation of the error term. The predictive density (12) gives us all the information we need to compute the predictive option price which is just the predictive expectation of P_T multiplied by $(1+r)^{-(T-t)}$ as we are in discrete time for discounting. The Bayesian approach provides a complete density so that a measure of uncertainty can be attached to the option price by computing a confidence interval.

The predictive density of either P_T, S_T or y_T is not straightforward to compute in a GARCH model. Details can be found in Bauwens, Lubrano and Richard (1999, Ch. 7) for the usual GARCH model.

When $T = t+1$, simple case. ns the number of predictions ahead is equal to one, the conditional density of a future return y_T under measure Q is a normal density given by

$$f_Q(y_T|\theta, y_t) = f_N(r, \omega + \alpha[y_t - \mu - \rho y_{t-1} + r]^2 + \beta\, h_t). \tag{13}$$

Conditionally on θ, all the parameters of this distribution are known as y_t and h_t are observed or already computed.

General case is complex. When $ns > 1$, the complete predictive is obtained by the multiple integral

$$\begin{aligned} f_Q(y_T|y_t) \;=\; & \int_\theta \int_{\mathbf{R^{ns-1}}} f_Q(y_T|\theta, y_{T-1}) \\ & f_Q(y_{T-1}|\theta, y_{T-2}) \;\cdots\; f_Q(y_{t+1}|\theta, y_t)\varphi(\theta|y)\, dy_{T-1} \cdots dy_{t+1}\, d\theta. \end{aligned} \tag{14}$$

We have to integrate θ as in (12), but also $y_{t+1}, \ldots, y_{T-1}$ which are unobserved, but can be simulated sequentially. The dimension of the integration over future y_{t+j} is thus $ns-1$. Each element in (14) is a normal density with mean r and variance h_{t+j}, but the resulting density in the inner integral is not of a known form. We propose in section 4 **a simulation algorithm which draws random numbers in the predictive density** (14) and which is inspired from Geweke (1989). Once a simulated sequence y_i $(i = t+1, \ldots, T)$ is obtained, it is transformed into a draw from the predictive density of P_T through the following transformation

$$\begin{aligned} S_T &= S_t \textstyle\prod_{i=t+1}^{T}(1+y_i) \\ P_T &= \max(S_T - K, 0). \end{aligned} \tag{15}$$

The predictive option price is defined as the discounted expectation of P_T. When we have obtained N draws P_T^j $(j = 1, \dots, N)$ of P_T, the predictive option price can be approximated by the empirical mean

$$C_t^T \simeq (1+r)^{-(T-t)} \frac{1}{N} \sum_{j=1}^{N} P_T^j. \tag{16}$$

With the same draws, we can estimate the complete predictive density of P_T using a non-parametric approach.
The empirical mean (16) converges to its population value if the empirical process is stationary. As we are in a Bayesian framework, the stationarity condition may be verified at the posterior mean of θ, but not necessarily for every point drawn from the posterior density. The points for which the stationarity condition is not met have to be rejected for the predictive evaluation[2].
As (16) is a Monte Carlo estimator, we would like to qualify its precision. Let us define the empirical standard deviation of C_t^T as

$$\hat{\sigma}_C = (1+r)^{-(T-t)} \sqrt{\frac{1}{N} \sum (P_T^j)^2 - (\frac{1}{N} \sum P_T^j)^2}$$

Provided the N draws are independent and N is large, the Monte Carlo error is asymptotically normal (see Bauwens et al 1999, p.75). A confidence interval for C_T^t is given by

$$C_T^t \pm z_\alpha \frac{\hat{\sigma}_C}{\sqrt{N}}$$

where z_α is the $1-\alpha$ quantile of the standard normal distribution.
If the model is estimated with a MCMC method (as this will be the case later on with the Griddy Gibbs sampler), the empirical standard deviation cannot be computed directly because the obtained draws of θ are not independent. A modified estimator is difficult to implement, but an easy solution is to compute the Monte Carlo error using only a part of the draws, say $\theta_{(j)}$, $j = 1, 11, 21, ...$ so that the resulting sequence of θ can fairly reasonably be taken as independent.

2.5 Comparison with the BS formula

The BS formula ignores the stochastic nature of volatility as it is modeled for instance by the GARCH model. It assumes incorrectly homoskedasticity when the real governing process is heteroskedastic. If we want to compare the pricing obtained by an incorrect use of the BS formula to the predictive option pricing in case of GARCH volatility, we have to plug in the BS formula the marginal variance of the empirical GARCH process. The stationary marginal variance of the neutralised GARCH process is greater than that of the empirical process. However Duan (1995) explains that this difference between the two variances cannot exhaust all the differences of results between GARCH and BS option pricing. It is well known (see for instance the references in Duan 1995) that

[2] This problem is common to all dynamic models. Bauwens, Lubrano and Richard (1999, p. 139) give details for the homoskedastic AR(1) model.

1. the BS formula underprices out-of-the-money options,
2. the BS formula underprices options on low-volatility securities,
3. the BS formula underprices short maturity options.

If we observe option prices, we can invert the BS formula to find the implied BS volatility. When plotting the result against the exercise price for a fixed maturity, the theory predict a straight horizontal line when this empirical curve is typically convex and known as the "**smile**". See Duan (1996) for an illustration.

3 GARCH models for Brussels stock index

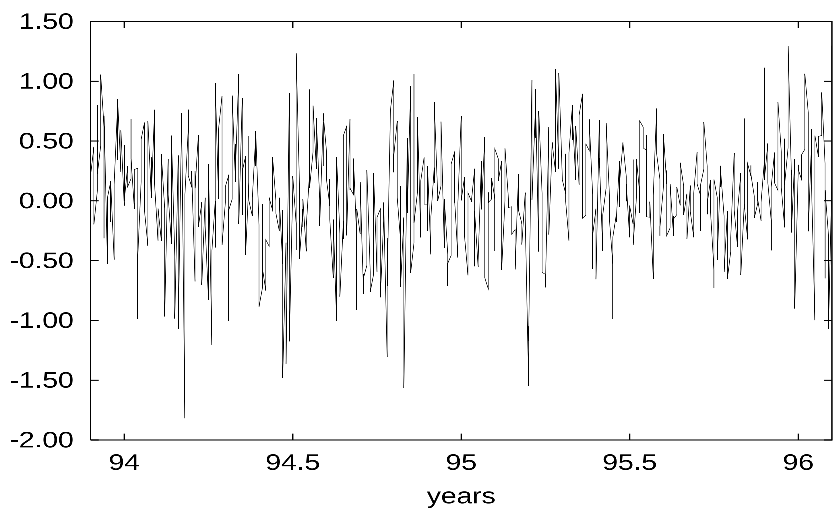

Figure 1: Brussels spot market index return
(23/11/93 - 30/01/96)

We have estimated three different GARCH models on a return series of the Brussels stock exchange. We used **daily data** on the spot market index of shares of Belgian firms, for the period 23/11/93 to 30/01/96. The data consist of **closing quotes**, providing 550 observations. The dependent variable (y_t) is the index return defined in (2) and multiplied by 100 to get percentages. A time series plot of the percentage returns is provided on Figure 1. The mean return is equal to 0.043% and the standard deviation to 0.48%. The greatest positive and negative returns are 1.30% and -1.82%, the last observation is 1.30% and the last but two observation is -1.07%. The index (not reproduced here) exhibits a downward trend over the first half of the sample period and an upward trend after.

3.1 Models and likelihood function

The general form of the three models that we estimate and their likelihood functions are given by:

$$\begin{aligned} y_t &= \rho y_{t-1} + \sqrt{h_t}\epsilon_t \qquad \epsilon_t \sim N(0,1) \\ h_t &= g(y_{t-1}, h_{t-1}, \theta) \\ l(\theta; y) &\propto \prod_{t=2}^n h_t^{-1/2} \exp(-\tfrac{1}{2}\textstyle\sum_{t=2}^T y_t^2/h_t). \end{aligned} \tag{17}$$

For the symmetric model, h_t is given in (10). We use a uniform prior on finite intervals for all the parameters. The posterior moments exist provided h_t is strictly positive and finite. This implies that the range of integration for ω starts at a strictly positive value.
The two variants of the symmetric GARCH model (10) we shall consider adopt the common formulation

$$h_t = \omega + \alpha_1 u_{t-1}^2 (1 - f_t) + \alpha_2 u_{t-1}^2 f_t + \beta h_{t-1}. \tag{18}$$

First variant: The model of Glosten, Jagannathan and Runkle (1993) assumes an asymmetry between positive and negative shocks and is obtained by defining f_t as a Heavyside function with

$$f_t = \begin{cases} 1 & \text{if } u_{t-1} < 0 \\ 0 & \text{otherwise .} \end{cases} \tag{19}$$

The leverage effect is present in (18) if α_2 is larger than α_1, so that the news impact curve is steeper for negative shocks than for positive ones. In this asymmetric model (hereafter GJR), the impact of shocks on the variance depends on the sign of the shock. This model was already used in a Bayesian framework by Bauwens and Lubrano (1998) on weekly data for the Brussels stock index. We use the same type of prior as above. Posterior moments exist under under the same condition provided there are enough observations in each regime.

The second variant considers the possibility of saturation in an asymmetric model. Volatility is increased by bigger shocks, but till a certain extent beyond which it stays constant. This effect is obtained by a smooth transition model introduced in Lubrano (2001) who replaces the above f_t Heavyside function by a smooth exponential transition function with a threshold:

$$f_t(u, \gamma, c) = 1 - \exp(-\gamma(u - c)^2). \tag{20}$$

In this transition function,

1. the threshold c monitors the degree of asymmetry and

2. γ monitors the saturation.

In order to get a scale free γ, the quantity $u - c$ has to be normalised by the standard deviation of y_t. This model (hereafter STR) is known to present **identification and estimability problems** at both ends of the support of γ. For $\gamma = 0$, the transition function is zero.

Consequently, α_2 becomes not identified. When $\gamma \to \infty$, the transition function becomes a Heavyside function equal to zero for $u = c$ and equal to one otherwise. The parameter c will tend to pick up an outlier so that the model tends to be equivalent to a linear GARCH model with a dummy variable for that point. As this is another well defined model having a positive probability, the posterior density of γ is not integrable because its right tail does not tend to zero quickly enough. See Lubrano (2001) for a formal proof. **We introduce a gamma prior to solve both problems**:

$$\varphi(\gamma) \propto \gamma^{\nu-1} \exp(-\gamma/\gamma_0). \tag{21}$$

The prior is zero at $\gamma = 0$ if $\nu > 1$ and has an exponentially decaying tail which insures the existence of posterior moments[3]. Generally speaking, the problem of existence of posterior moments arises every time the likelihood function does not go to zero quickly enough or presents a non-integrable singularity. This is the case for some simultaneous equation models and for cointegration models (see Bauwens, Lubrano and Richard 1999). For GARCH models, the likelihood function based on Gaussian errors does not present this kind of pathology. The problem could have arisen with the STR model only. But the prior solves the matter. And as anyway we integrate on a finite interval and the function is finite, the resulting integrals converge.

[3]Note that with $\nu = 1$ we get as a particular case the exponential prior used for instance by Geweke (1993) in a regression model with Student errors.

3.2 Inference results

Table 1: Inference in the GARCH(1,1) model

	μ	ρ	ω	β	α_1	α_2	γ	c
				Classical				
Basic	0.031	0.24	0.0075	0.91	0.056			
	[0.019]	[0.044]	[0.0053]	[0.036]	[0.019]			
GJR	0.027	0.25	0.0098	0.90	0.021	0.083		
	[0.020]	[0.045]	[0.0069]	[0.042]	[0.021]	[0.031]		
STR	0.028	0.24	0.0081	0.90	0.0	0.088	2.17	1.16
	[0.019]	[0.045]	[0.0058]	[0.038]	[–]	[0.031]	[4.36]	[0.46]
				Bayesian				
Basic	0.032	0.24	0.028	0.80	0.082			
	[0.020]	[0.046]	[0.020]	[0.11]	[0.037]			
GJR	0.026	0.25	0.027	0.80	0.039	0.13		
	[0.020]	[0.047]	[0.020]	[0.11]	[0.033]	[0.064]		
STR	0.028	0.25	0.028	0.79	0.0	0.13	1.15	0.91
	[0.020]	[0.047]	[0.016]	[0.088]	[–]	[0.063]	[0.81]	[0.44]

5000 draws plus 750 for warming were used for the Griddy-Gibbs algorithm, see Bauwens and Lubrano (1998). Average computing time was 10 minutes on a Pentium 1Gh. All computations were done with GAUSS.

Classical and Bayesian inference results for the three models are given in Table 1. The following general comments apply to these results:

1) Negative shocks have a stronger impact than positive shocks on the next conditional variance. In the GJR model, the difference $\lambda = \alpha_2 - \alpha_1$ has a posterior mean equal to 0.090 with a standard deviation equal to 0.066. The posterior probability that λ is negative is small (less than 6%). The posterior probability that α_2 is negative in the STR model is zero.

2) For the smooth transition model, we chose $\nu = 3$ and $\gamma_0 = 0.5$ for the gamma prior, which implies $\mathrm{E}(\gamma) = 1.5$ and $\mathrm{Var}(\gamma) = 0.75$. The posterior density is dominated by the prior. The coefficient corresponding to positive shocks, α_1 was set equal to zero, as the maximum likelihood algorithm could not depart from this starting value. But the news impact curve (not reported here) is not flat for positive shocks as the transition between the two regimes is smooth. Note that the news impact curves of the GJR and the smooth transition model are fairly similar, but with a difference in the right tail which implies a further dampening influence of large positive shocks for the smooth transition model compared to the GJR model.

3) Classical estimates and Bayesian posterior moments are very close for the regression parameters μ and ρ. Note the differences for ω and β because the posterior densities of these two parameters are rather skewed. The gamma prior on γ entails a much smaller posterior mean and standard deviation than their classical counterparts for γ and c. But Bayesian posterior standard deviations are substantially larger for the GARCH parameters ω, α and β, giving certainly a better account of parameter uncertainty than that provided by classical standard errors.

3.3 Comparing estimated volatilities

We can compute the implied marginal variance for each empirical model. Formulae are given in the Appendix. These estimates can be used in the BS formula when one decides to ignore heteroskedasticity. They have to be compared to the empirical variance 0.231. Table

Table 2: Marginal volatilities
for the empirical models

	Basic		GJR		STR	
	P	Q	P	Q	P	Q
Classical σ^2	0.232	0.291	0.220	0.253	0.892	2.292
Bayesian σ^2	0.249	0.299	0.249	0.319	0.375	0.470
Prob. stat.	0.986	0.976	0.984	0.979	0.874	0.837

Moments and posterior probabilities for P and Q are computed on the same set of draws, supposing that $r = 0$ for the Q case. Prob. stat. is the posterior probability of stationarity.

2 presents the classical and the Bayesian results for the three models. They are obtained as transformations of the original parameters. Classical results are evaluated at the MLE and are thus very sensitive to the verification of the stationarity condition. When we are near to nonstationarity, any uncertainty on α and β has huge consequences on the final estimate. Bayesian marginal moments are obtained as a transformation of every draw of the posterior density, the expectation being taken at the end (instead of computing the marginal moments at the posterior expectation of the parameter). They are truncated moments as the draws which do not verify the stationarity condition are rejected. The reported posterior probabilities indicate that a small amount of draws had to be rejected and that the process can be taken as stationary. Due to the way they are computed, the Bayesian results seem to be more reliable than the classical ones, especially for the STR model.

The results indicate **a significant increase in the marginal volatility when risk neutralisation is taken into account**. But the adoption of a non-linear model does not seem to make a great difference. In other words, taking into account asymmetries with the GJR model does not seem to improve a lot over the symmetric GARCH. The smooth transition model, on the contrary seems to give a slightly different account of the importance of volatility. The correlation between a non-parametric estimate of the historical volatility and the estimated h_t series produced by the STR model is higher than for the two other models. Thus **there is some incentive to use a smooth transition GARCH.**

4 Predictive option prices: computations and results

Predictive option pricing requires the evaluation of **a high dimensional integral as we have to take the expectation both over the parameter space of θ and over future returns**. As indicated before, Geweke (1989) has proposed a simulation algorithm to compute the integrals in (14). In order to adapt this algorithm, we must take into account the fact that the original GARCH model has an autoregressive mean and that the prediction has to be done with the risk neutralised GARCH process. We first detail the predictive skedastic function for each model and then give the simulation algorithm. The rest of the section is devoted to empirical results.

4.1 Predictive skedastic functions

Let us write down **the risk neutralised predictive skedastic function** for each of the three proposed empirical models. The general formula is given in (9) and we have justified previously the choice $\mu_t = \mu + \rho y_{t-1}$. For the symmetric GARCH, we get

$$h_T = \omega + \alpha v_{T-1}^2 + \beta h_{T-1} \tag{22}$$

where

$$v_{T-1} = (1-\rho)r - \mu + \epsilon_{T-1}\sqrt{h_{T-1}} - \rho\epsilon_{T-2}\sqrt{h_{T-2}} \tag{23}$$

and $\epsilon \sim \mathrm{N}(0,1)$. For the asymmetric models, we have

$$h_T = \omega + \alpha_1 v_{T-1}^2 + (\alpha_2 - \alpha_1)v_{T-1}^2 f_T + \beta h_{T-1}. \tag{24}$$

First case: Let us first consider the case where f_T is the Heavyside function being one if v_{T-1} is negative and zero otherwise. We show in the Appendix that the predictive variance of y_T is

$$\mathrm{Var}(y_T) = \frac{\omega + [(1-\rho)r - \mu]^2(\alpha_1+\alpha_2)/2}{1-(1+\rho^2)(\alpha_1+\alpha_2)/2 - \beta} \tag{25}$$

provided

$$(1+\rho^2)(\alpha_1+\alpha_2)/2 + \beta < 1. \tag{26}$$

Second case: When f_T is the exponential smooth transition function defined in (20), the predictive variance is

$$\mathrm{Var}(y_T) = \frac{\omega + \alpha_2[(1-\rho)r - \mu]^2}{1-(1+\rho^2)\alpha_2 - \beta} \tag{27}$$

provided

$$(1+\rho^2)\alpha_2 + \beta < 1. \tag{28}$$

These two stationarity conditions together with $\alpha + \beta < 1$ for the symmetric case, serve to **reject draws of θ for which stationarity is not verified.**

Table 3: Simulation algorithm for option prices
with maturity ns periods ahead of t

$i = 1, N$
  $\theta_i \sim \varphi(\theta|y)$
  $k = 1, M$
    $\epsilon \sim \mathrm{N}_{ns}(0, I_{ns})$
    $hs_1 = \omega_i + \alpha_i(y_t - \mu_i - \rho_i y_{t-1})^2 + \beta_i h_t$
    $ys_1 = r + \epsilon_1\sqrt{hs_1}$
    $hs_2 = \omega_i + \alpha_i(ys_1 - \mu_i - \rho_i y_t)^2 + \beta_i hs_1$
    $ys_2 = r + \epsilon_2\sqrt{hs_2}$
    $j = 3, ns$
      $hs_j = \omega_i + \alpha_i(ys_{j-1} - \mu_i - \rho_i ys_{j-2})^2 + \beta_i hs_{j-1}$
      $ys_j = r + \epsilon_j\sqrt{hs_j}$
    $j = j + 1$
    $S_T = S_t \prod_{j=1}^{ns}(1 + ys_j/100)$
    $P_T = (1 + r)^{-ns} \max(S_T - K, 0)$
  k = k+1
$i = i + 1$

4.2 Computational aspects

The algorithm given in Table 3 produces $N \times M$ draws of the predictive density of P_T. It uses draws from the posterior density of θ which are easily obtained provided a Monte Carlo method has been used to make inference and that the draws have been stored. The draws that do not satisfy the stationarity conditions stated in the previous subsection have to be rejected. **The inner loop on k approximates the inner integral in y while the outer loop on i approximates the outer integral in θ.**
Geweke (1989) shows that when $N \to \infty$ the value of M no longer matters to obtain consistent results. However, as θ is typically costly to obtain, there is an incentive to take M greater than 1. In the example treated below, **we took $N = 5000$ (minus the number of rejections) and $M = 50$. In a classical approach, N would be equal to one as θ is fixed at the maximum likelihood estimate and M would be chosen large.** Duan (1996) for instance uses $M = 5000$.

The predictive density of the payoff function of the option can be approximated by a kernel estimate of the $N \times M$ simulated values of that function. Option prices are computed as the discounted sample mean of P_T as given in (16) with $r = 0$.

4.3 Option pricing for the Brussels spot index

We report in Table 4 the **predictive option prices for the Brussels spot index for 3 different maturities**: 15 days, 30 days and 60 days, starting at the end of the observed sample minus 2 observations (because this $n-2$ observation is largely negative and thus constitutes a good example for the asymmetric models). A 15 day maturity is certainly very short for an option, but one should remember that its price is also the price of any option 15 days before its maturity. When we observe a real option market for say 3 month options, the market is very quiet at the beginning of the option life, but gets very active a few weeks before the maturity. We have supposed that the riskless interest rate r was equal to zero. We took the normalisation rule $S_t = 1$ and considered a range of moneyness (S_T/K) between 0.959 and 1.045.

We have computed the **Monte Carlo standard errors**. We do not give them in full detail, but the following indications are useful. For out-of-the-money options, t statistics are slightly above the maturity (20 for a 15 day maturity, etc). For at-the-money options, they are all of the order of 200. For in-the-money options, they are between 1000 and 15000. Translated in term of precision for the estimates, a 1% precision needs a t of 200, and 10% precision a t of 20 (if the mean is 0.010, then 0.011 is at the border of the 95% confidence interval in the case of a 10% precision). Every time we shall say that two option prices are different, we mean that the second is not contained within a 95% confidence interval of the first.

We computed but do not report classical results. Broadly speaking, **classical estimates of the option prices are close the Bayesian predictive means.** They are slightly different for out-of-the-money options. The differences are weaker for at-the-money options. We can say that there are no noticeable differences for in-the-money options, except when using the STR model. This is a sign of the robustness of the risk-neutral pricing methodology.

Table 4: Predictive call option prices
on Brussels index

Moneyness	0.959 (out)		1.000 (at)		1.045 (in)	
Basic						
Maturity	C_t^T	BS_t^T	C_t^T	BS_t^T	C_t^T	BS_t^T
15 days	0.00021	0.00013	0.0081	0.0076	0.0430	0.0430
	(0.02)		(0.50)		(0.98)	
30 days	0.00099	0.00079	0.0113	0.0108	0.0437	0.0435
	(0.07)		(0.50)		(0.94)	
60 days	0.00315	0.00278	0.0159	0.0153	0.0457	0.0453
	(0.14)		(0.49)		(0.86)	
GJR						
Maturity	C_t^T	BS_t^T	C_t^T	BS_t^T	C_t^T	BS_t^T
15 days	0.00018	0.00015	0.0084	0.0077	0.0431	0.0430
	(0.02)		(0.51)		(0.91)	
30 days	0.00088	0.00081	0.0116	0.0109	0.0440	0.0435
	(0.07)		(0.51)		(0.93)	
60 days	0.00306	0.00283	0.0162	0.0153	0.0461	0.0453
	(0.14)		(0.50)		(0.86)	
STR						
Maturity	C_t^T	BS_t^T	C_t^T	BS_t^T	C_t^T	BS_t^T
15 days	0.00020	0.00038	0.0082	0.0090	0.0430	0.0432
	(0.02)		(0.50)		(0.98)	
30 days	0.00089	0.00158	0.0113	0.0128	0.0438	0.0442
	(0.06)		(0.50)		(0.94)	
60 days	0.00283	0.00455	0.0159	0.0180	0.0457	0.0469
	(0.14)		(0.50)		(0.86)	

C_t^T is the predictive option price defined in (16). Numbers in parentheses are the probabilities of exercise of the option. BS_t^T is the mean of (5) evaluated at the marginal variance under measure P, the mean being taken with respect to θ. Moneyness is S_T/K. It is equal to 1 for options at the money, smaller than 1 for out-of-the-money options, and larger than 1 for in-the-money options (at time t). The computing time (with GAUSS on an Pentium 350Mh) using 500 (N) times 50 (M) repetitions is 77 seconds with the basic model for a maturity of 60 days, shorter maturities being obtained as a by-product. This time goes up to 102 seconds for the GJR model and 109 seconds for the STR model.

Let us now **compare** the prices obtained using the predictive method and those obtained with the expectation of **the BS formula**[4] , both reported in Table 4.

1. For in-the-money options, there are no economically significant differences.

2. For at-the-money options, significant differences are appearing (more than a 5% difference in price when the numerical precision is of the order of 1%), but these differences are decreasing with maturity. The BS formula underprices options for the Basic and GJR models, but overprices it for the STR model.

3. For out-of-the-money options, the BS formula underprices strongly short maturities (38% for 15 days) for the Basic model. This underpricing attenuates with maturity and when the GJR model is used. This means that the GJR model gives a better account of the marginal volatility of the underlying security. With the STR model, the BS formula strongly overprices, showing that the marginal volatility is not estimated with a satisfactory accuracy in this model.

All these results amplify when we plug in the BS formula the classical estimate of the marginal volatility.
We report in Table 5 the implied mean predictive volatility (the mean of the predicted h_{T+j} under process Q). For the three models, the conditional mean predictive volatility converges

Table 5: Mean predictive volatilities

	15 days	30 days	60 days
Basic	0.282	0.275	0.270
GJR	0.304	0.290	0.282
STR	0.291	0.278	0.269

with maturity to a smaller value. The decrease is due to the configuration of the sample and convergence results from stationarity. These mean conditional volatilities converge also to smaller values than the computed marginal volatilities reported in Table 2. The latter were computed as truncated moments. Table 5 reveals no serious differences (not more than 5%) between the models. But Table 2 reveals a similar picture for the Bayesian estimates of the Basic and the GJR models. **So the differences we have found between the BS and the predictive pricing methods for out-of-the-money options come just from the fact that the BS formula is too sensitive to variations of the volatility.**

[4]We computed $\mathrm{E}_{\theta|y}(BS[\sigma(\theta)])$. This is the posterior expectation of the BS formula of section 2.1. It is different from the Hull and White (1987) approach of section 2.2.

Looking at the results given in Table 4, we note that the expected option prices are not much different between the Basic and the two asymmetric models (within the Monte Carlo precision).

1. On one side, this gives confidence about the robustness of the predictive pricing method compared to a misuse of the BS formula as reported above.

2. On the other side, this is disrupting as one may think that asymmetric models would have produced a different pricing in the presence of negative shocks and short maturities.

However, we must not forget that the Bayesian approach gives not only an expected option price, but also a complete density. Considering the complete density may shed some more light on the question.

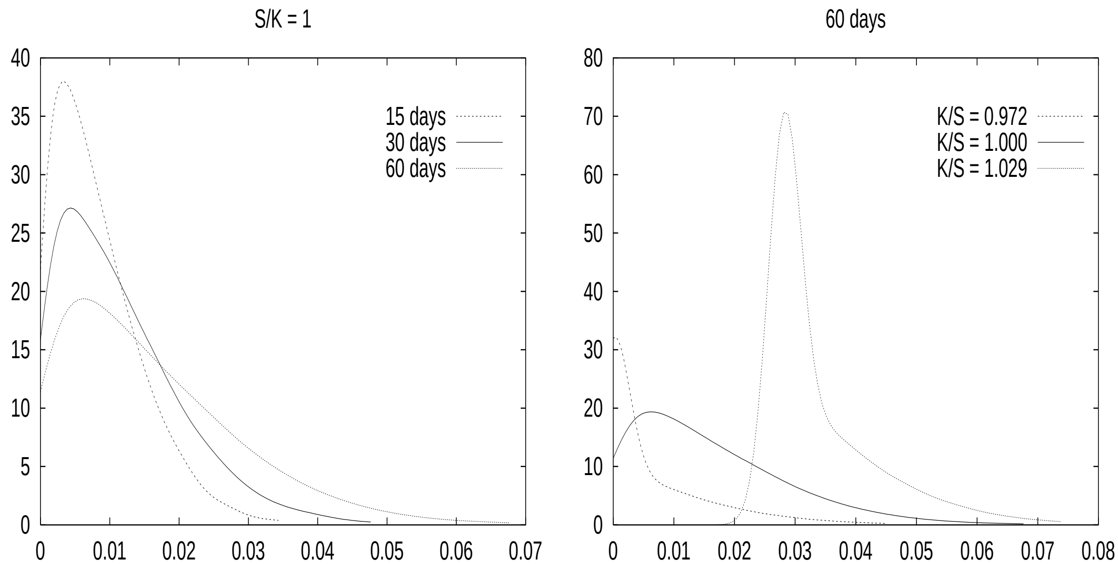

Figure 2: Predictive density for the payoff function
for the Basic model

The interest of our approach is to provide a probability distribution with respect to which (observed) option prices can be assessed, namely the predictive distribution of the option payoff function (the expected value of this distribution being the predictive option price).

1. **Left panel**: the computed predictive densities for the three maturities and a strike of 1. Actually this type of density has a discrete component and a continuous one: the discrete component is at zero and corresponds to the fact that the payoff function is the maximum of zero and the difference $S_T - K$, and the continuous part corresponds to the positive values of this difference. Only the continuous part is drawn on the figures, so that the densities integrate to the probability of being positive (e.g. about equal to 0.5 for at-the-money options). **The probability at zero is the predictive probability that the option will not be exercised.** It is evaluated by the proportion of null values of the payoff function in the total number of simulations ($N \times M$). The complementary probabilities are reported in Table 4 (see the numbers in parentheses). The message which is apparent from the left panel of Figure 2 is that, at least for the observed sample, **uncertainty increases with maturity while the predictive mode does not increase significantly.**

2. On the contrary, as shown in the right panel of Figure 2, the option price is very sensitive to the moneyness (given there for a maturity of 60 days and a smaller range of moneyness than that displayed in Table 4).

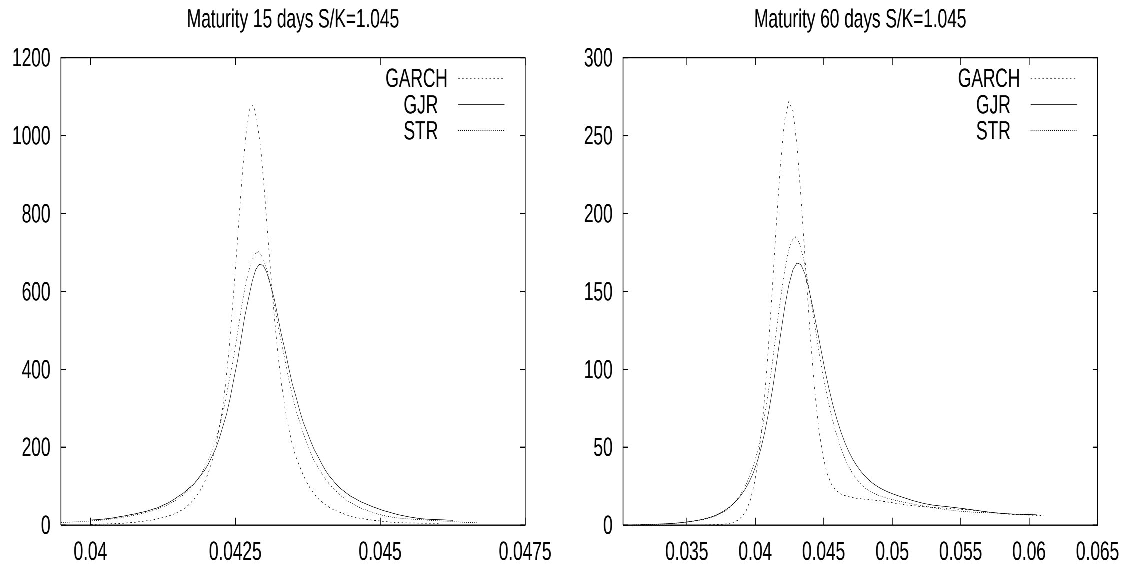

Figure 3: Impact of the model choice
on the Predictive density for the payoff function

It is difficult to detect an influence of modeling asymmetry on option prices. For out-of-the-money options and at-the-money options, there is no significant differences in the graphs of the predictive payoffs. But surprisingly **for options largely out-of-the-money** as those reported here, a difference appears as shown in Figure 3. The two asymmetric models perform similarly (which is not so surprising). But **ignoring asymmetry apparently gives a too high precision about the mean option price** for our particular sample. This effect seems to be valid independently of the maturity.

5 Conclusion

This paper shows how option prices can be evaluated from a Bayesian viewpoint using an econometric model to predict the future stock price and volatility. The option price provided by our method is the predictive expectation of the payoff function of the option. Other characteristics of options (delta, gamma, ...) could be predicted. Our method delivers the predictive distribution of the payoff function for a given econometric model. This probability distribution could be useful to market participants who wish to compare the model prediction to potential prices on the market or to other predictions.

This paper also shows that in a Bayesian approach an additional source of uncertainty is taken into account which has consequences on the measure of volatility and as a result on option pricing. It shows how the predictive method finally produces a better evaluation of the actual volatility contained in the data when there is a large uncertainty on marginal measures due to nearly nonstationarity. Because of this nearly nonstationarity, using a marginal measure to be plugged in the BS formula can be very dangerous.

Finally, this paper shows that modelling asymmetry does not seem essential for pricing options nearly at the money, but may have a significant impact on the dispersion of the predicted prices for in-the-money options.

APPENDIX
VARIANCE AND STATIONARITY CONDITION
FOR RISK NEUTRALISED GARCH PROCESSES

The marginal variance of y in GARCH processes is computed using the law of iterated expectations and the fact that the process is supposed to be stationary. See Bauwens, Lubrano and Richard (1999, Ch. 7) for details.

Let us write down the risk neutralised predictive process for each of the three proposed empirical models. We have first

$$y_T = r + \epsilon_T \sqrt{h_T} \quad \epsilon_T \sim \mathrm{N}(0,1) \tag{29}$$

For the skedastic function, the general formula is given in (9) and we have justified previously the choice $\mu_t = \mu + \rho y_{t-1}$. For the symmetric GARCH, the predictive skedastic function is

$$h_T = \omega + \alpha v_{T-1}^2 + \beta h_{T-1} \tag{30}$$

where

$$v_{T-1} = (1-\rho)r - \mu + \epsilon_{T-1}\sqrt{h_{T-1}} - \rho\epsilon_{T-2}\sqrt{h_{T-2}} \tag{31}$$

The predictive expectation of h_T is given by

$$\mathrm{E}(h_T) = \omega + \alpha \mathrm{E}(v_{T-1}^2) + \beta \mathrm{E}(h_{T-1}) \tag{32}$$

We have essentially to compute

$$\begin{aligned}
\mathrm{E}(v_{T-1}^2) &= [(1-\rho)r-\mu]^2 + \mathrm{E}[(\epsilon_{T-1}\sqrt{h_{T-1}} - \rho\epsilon_{T-2}\sqrt{h_{T-2}})^2] \\
&= [(1-\rho)r-\mu]^2 + \mathrm{E}[\epsilon_{T-1}^2 h_{T-1} + \rho^2\epsilon_{T-2}^2 h_{T-2}] \\
&= [(1-\rho)r-\mu]^2 + (1+\rho^2)\mathrm{E}[h]
\end{aligned} \tag{33}$$

Solving for $\mathrm{E}(h)$, we get

$$\mathrm{E}(h_T) = \frac{\omega + \alpha[(1-\rho)r-\mu]^2}{1-\alpha(1+\rho^2)-\beta} \tag{34}$$

The process is stationary if $1-\alpha(1+\rho^2)-\beta > 0$.

For the GJR asymmetric model, the skedastic function is

$$h_T = \omega + \alpha_1 v_{T-1}^2 + (\alpha_2-\alpha_1)v_{T-1}^2 f_T + \beta h_{T-1} \tag{35}$$

where f_T is the Heavyside function which is one if $v_{T-1} < 0$ and zero otherwise. We have to compute

$$\mathrm{E}(h_T) = \omega + \alpha_1\mathrm{E}(v_{T-1}^2) + (\alpha_2-\alpha_1)\mathrm{E}(v_{T-1}^2|v_{t-1}<0) + \beta\mathrm{E}(h_{T-1}) \tag{36}$$

We have already got $\mathrm{E}(v_{T-1}^2)$ from the symmetric case. It thus remains to compute the truncated conditional expectation $\mathrm{E}(v_{T-1}^2|v_{t-1}<0)$

$$\mathrm{E}(v_{T-1}^2|v_{t-1}<0) = ((1-\rho)r-\mu)^2 + \mathrm{E}[(\epsilon_{T-1}\sqrt{h_{T-1}} - \rho\epsilon_{T-2}\sqrt{h_{T-2}})^2|v_{t-1}<0]$$

Supposing that the distribution of v_{t-1} is symmetric and that the function $\epsilon_{T-1}\sqrt{h_{T-1}} - \rho\epsilon_{T-2}\sqrt{h_{T-2}}$ is also symmetric, we have

$$\mathrm{E}[(\epsilon_{T-1}\sqrt{h_{T-1}} - \rho\epsilon_{T-2}\sqrt{h_{T-2}})^2|v_{t-1}<0] = (1+\rho^2)\mathrm{E}(h)/2$$

which is just half the quantity obtained in the symmetric case. Regrouping partial results, we get

$$\mathrm{E}(h_T) = \frac{\omega + [(1-\rho)r-\mu]^2(\alpha_1+\alpha_2)/2}{1-(1+\rho^2)(\alpha_1+\alpha_2)/2-\beta}$$

The condition for stationarity this time is

$$1-(1+\rho^2)(\alpha_1+\alpha_2)/2-\beta > 0. \tag{37}$$

When f_T is a smooth transition function, it seems natural to study the stationarity condition for the limiting case when $\gamma \to \infty$. With an even smooth transition function, the limiting case is the GJR model. When the smooth transition function is the asymmetric exponential function defined in (20), it must be noted that the model becomes linear when $\gamma \to \infty$. Consequently, the stationarity condition in this case is identical to that derived for the standard GARCH(1,1) model, just replacing α by α_2. Consequently

$$\mathrm{E}(h_T) = \frac{\omega + \alpha_2[(1-\rho)r-\mu]^2}{1-(1+\rho^2)\alpha_2-\beta} \tag{38}$$

with the associated stationarity condition

$$1-(1+\rho^2)\alpha_2-\beta > 0. \tag{39}$$

References

Bauwens, L. and M. Lubrano, 1998, Bayesian inference on GARCH models using the Gibbs sampler, *The Econometrics Journal* 1, C23-C46.

Bauwens, L., M. Lubrano and J.F. Richard, 1999, *Bayesian Inference in Dynamic Econometric Models* (Oxford University Press, Oxford).

Campbell, J., A.W. Lo and A.C. MacKinley, 1997, *The Econometrics of Financial Markets* (Princeton University Press, Princeton).

Cox, J.C. and S.A. Ross, 1976, The valuation of options for alternative stochastic processes., *Journal of Financial Economics* 3, 145-166.

Davidson, R. and J.G. MacKinnon, 1993, *Estimation and inference in econometrics* (Oxford University Press, Oxford).

Duan, J.C., 1995, The GARCH option pricing model, *Mathematical Finance* 5, 13-32.

Duan, J.C., 1996, Craking the smile, *Risk* 9 (12), 55-59.

Duan, J.C., 1999, Conditionally fat-tailed distributions and the volatility smile in options, Working Paper, Department of Finance, Hong-Kong University.

Engle, R.F., D.M. Lilen and R.P. Robin, 1987, Estimating time varying risk premia in the term structure: the ARCH-M model, *Econometrica* 55, 391-407.

Garcia, R. and E. Renault, 1997, A note on hedging in ARCH and stochastic volatility option pricing models, Working Paper No 97s-13, CIRANO, Université de Montréal.

Geweke, J., 1989, Exact predictive densities for linear models with ARCH disturbances, *Journal of Econometrics* 40, 63-86.

Geweke, J., 1993, Bayesian treatment of the independent Student-t linear model, *Journal of Applied Econometrics* 8 (Supplement), 19-40.

Glosten, L.R., R. Jagannathan and D.E. Runkle, 1993, On the relation between the expected value and the volatility of the nominal excess return on stocks, *Journal of Finance* 48, 1779-1801.

Hafner, C.M. and H. Herwartz, 2001, Option pricing under linear autoregressive dynamics, heteroskedasticity and conditional leptokurtosis, *Journal of Empirical Finance* 8(1), 1-34.

Heston, S.L. and S. Nandi, 2000, A closed-form GARCH option valuation model, *The Review of Financial Studies* 13, 585-625.

Hull, J. and A. White, 1987, The pricing of options with stochastic volatilities, *Journal of Finance* 42, 281-300.

Kallsen, J. M.S. and Taqqu, 1998, Option pricing in ARCH-type models, *Mathematical Finance* 8, 13-26.

Lubrano, M., 2001, Smooth transition GARCH models: a Bayesian perspective. *Recherches Economiques de Louvain*, forthcoming.

Mahieu R.J. P.C. and Schotman, 1998, An empirical application of stochastic volatility models, *Journal of Applied Econometrics* 13, 333-360.

Noh, J., R.F. Engle and A.Kane, 1994, Forecasting volatility and option prices of the S & P 500 index, *Journal of Derivatives* 17-30.

Pagan, A., 1996, The econometrics of financial markets, *Journal of Empirical Finance* 3, 15-102.

Terasvirta, T., 1996, Two stylised facts and the GARCH(1,1) model, *Working Paper No. 96*, Working Paper Series in Economics and Finance, Stockholm School of Economics.