6
Option Valuation under Mixture GARCH models

This chapter is based in a working paper together with the Professors Ken Siu, John Lau and Álvaro Veiga.

It has been documented that normal mixture GARCH models can provide a better description for the leptokurtosis behavior in financial returns data compared with the GARCH models with normal innovations and the student's t-GARCH models (see Alexander and Lazar (2006) (1)). In this paper, we shall consider the pricing of options under the class of discrete-time mixture of GARCH models with innovations having a finite mixture of infinitely divisible distributions. The option valuation model can provide market practitioners with a convenient and flexible way to price options under various forms of mixture GARCH models, which can incorporate different degrees of conditional skewness and conditional leptokurtosis of the distribution of the asset returns. The market described by the discrete-time mixture GARCH models is incomplete and, hence, there are infinitely many equivalent martingale measures. We shall employ the doubly stochastic Esscher transform to determine an equivalent martingale measure for pricing. The pricing result can be justified by a stochastic version of the power utility maximization. Empirical results for comparing the call and put option prices obtained from the mixture GARCH models with those from the standard Black-Scholes model based on the recent 25-year S&P 500 data will be presented and discussed.

6.1
Asset Price Dynamics and Pricing Model

We consider a discrete-time financial model consisting of one risk-free bond B and one risky stock S. We assume that the dynamics of the risky stock is governed by a mixture GARCH model with innovations having a finite mixture of infinitely divisible distributions. The mixture GARCH model can incorporate various parametric forms of the mixture GARCH models, such as the Normal-Mixture (NM) GARCH models in Alexander and Lazar (2006)(1), the GARCH models with innovations having a finite mixture of shifted gamma distributions and a finite mixture of shifted Inverse Gaussian distributions. It provides market practitioners with a great deal of flexibility in modelling various empirical "stylised" behavior of asset price dynamics, like the conditional skewed and leptokurtosis (or heavy-tailed) behaviors of the asset returns. In the following, we present the setup of the model.

First, we describe the general mixture GARCH model. Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a complete probability space, where $\mathcal{P}$ is a real-world probability. Let $\mathcal{T}$ be the time index set $\{0, 1, 2, \dots, T\}$ of the financial model. Let $S := \{S_t\}_{t \in \mathcal{T}}$ denote a stochastic process defined on $(\Omega, \mathcal{F})$ with state space $\mathcal{R}^+$, where $\mathcal{R}^+$ is the set of non-negative real numbers. For each $t \in \mathcal{T}$, S_t represents the price of the risky stock S at time t. Write $\mathcal{F}^S := \{\mathcal{F}^S_t\}_{t \in \mathcal{T}}$ for the $\mathcal{P}$-augmentation of the natural filtration generated by the process S. For each $t \in \mathcal{T}$, $\mathcal{F}^S_t$ represents the observable information about the prices of the risky stock S up to and including time t. Let $B := \{B_t\}_{t \in \mathcal{T}}$ denote the price process of the risk-free bond B, which is assumed to be a deterministic process.

Let $\{\xi_t\}_{t \in \mathcal{T}}$ denote a stochastic process defined on $(\Omega, \mathcal{F})$ taking values on the real line $\mathcal{R}$, with $\xi_t \sim D(0, h_t)$ and $\xi_0 = 0$, which represents the random fluctuations of the returns from the risky asset S. For each $t \in \mathcal{T}$, we call ξ_t the innovation of asset return at time t. For each $k = 1, 2, \dots, K$, let $h_k := \{h_{kt}\}_{t \in \mathcal{T} \setminus \{0\}}$ denote a stochastic process on $(\Omega, \mathcal{F})$ with state space $\mathcal{R}^+$, where $\mathcal{R}^+$ is the set of positive real numbers. For each $k = 1, 2, \dots, K$, we assume that the dynamics of h_k is governed by the following GARCH(p, q) structure:

$$h_{kt} = \alpha_{k0} + \sum_{j=1}^{q} \alpha_{kj} \xi^2_{t-j} + \sum_{i=1}^{K} \sum_{l=1}^{p} \beta_{kil} h_{i,t-1} \, , \tag{6-1}$$

where $p \geq 1$, $q \geq 1$ and $\alpha_{k0} > 0$, $\alpha_{kj} \geq 0$, $j \in \{1, 2, \dots, q\}$, $\beta_{kil} \geq 0$, $l \in \{1, 2, \dots, p\}$ in order to ensure the positivity of h_{kt}.

In the particular case of the GARCH(1, 1) that we are going to deal with, for ensuring covariance stationarity of the GARCH(1, 1) structure for each $k = 1, 2, \dots, K$, we further impose the condition that the matrix

$$\alpha_1 p^T + B$$

has all the eigenvalues smaller than 1, where $\alpha_1 = [\alpha_{11}, \alpha_{21}, ..., \alpha_{K1}]^T, p = [p_1, p_2 ..., p_K]^T$ and $B = \beta_{ki}; k, i = 1, ..., K$.

Write $U := \{U_t\}_{t \in \mathcal{T}}$ for a sequence of independent and identically distributed (i.i.d.) K-dimensional random vectors, which take values from the state space $\mathcal{U} := \{e_1, e_2, \dots, e_K\}$, where $e_k := (0, 0, \dots, 1, \dots, 0, 0) \in \mathcal{R}^K$ is a unit vector with one in the k^{th} component and zero otherwise. We suppose that the common probability distribution of U_t is specified by:

$$\mathcal{P}(U_t = e_k) = p_k \ , \ k = 1, 2, \ldots, K \ , \tag{6-2}$$

where $p_k \geq 0$ and $\sum_{k=1}^{K} p_k = 1$.

Let $H := \{H_t\}_{t \in \mathcal{T}\setminus\{0\}}$ denote a stochastic process on $(\Omega, \mathcal{F})$ with state space $\mathcal{H} := \{h_{1t}, h_{2t}, \ldots, h_{Kt}\}$. For each $t \in \mathcal{T}\setminus\{0\}$, let $\mathcal{H}_t := (h_{1t}, h_{2t}, \ldots, h_{Kt})$. Then, we can write H_t as follows:

$$H_t = < \mathcal{H}_t, U_t > = \sum_{k=1}^{K} < \mathcal{H}_t, e_k >< U_t, e_k > \ . \tag{6-3}$$

Write $\mathcal{F}_t^U$ for the information set generated by the values of the process U up to and including time t. Write $\mathcal{F}_t$ for the enlarged information set $\mathcal{F}_{t-1}^S \vee \mathcal{F}_t^U$ generated by $\mathcal{F}_{t-1}^S$ and $\mathcal{F}_t^U$. We shall specify the distributional structure of the process $\{\xi_t\}_{t \in \mathcal{T}}$.

$$F(x|0, H_t) = \sum_{k=1}^{K} F(x|0, < \mathcal{H}_t, e_k >) < U_t, e_k > = \sum_{k=1}^{K} F(x|0, h_{kt}) < U_t, e_k > \tag{6-4}$$

Hence, the conditional distribution of ξ_t given the observable information set $\mathcal{F}_{t-1}^S$ is given by the following finite mixture of infinitely divisible distributions:

$$F_\xi(x|\mathcal{F}_{t-1}^S) = \sum_{k=1}^{K} p_{kt} F(x|0, h_{kt}) \ , \tag{6-5}$$

where $F(\cdots|0, h_{kt})$ is the mixing kernel of the mixture distribution.

We write $\xi_t|\mathcal{F}_{t-1}^S \sim MID(p_1, \ldots, p_K; h_{1t}, \ldots, h_{Kt})$, which represents the finite mixture of infinitely divisible distributions. We shall introduce the price dynamics of the risk-free bond B and the underlying risky asset S. Let r_t be the continuously compounded risk-free interest rate of the bond B over the time interval $[t-1, t]$, where $t \in \mathcal{T}\setminus\{0\}$; λ_t the unit risk premium over the time interval $[t-1, t]$. We suppose that both r_t and λ_t are deterministic functions of time t. Then, we assume that, under $\mathcal{P}$, the dynamics of the bond-price process $\{B_t\}_{t \in \mathcal{T}}$ and the stock-price process $\{S_t\}_{t \in \mathcal{T}}$ satisfy:

$$\begin{aligned} B_t &= B_{t-1} e^{r_t} \ , \quad B_0 = 1 \ , \\ S_t &= S_{t-1} \exp(r_t + \lambda_t \sqrt{< \mathcal{H}_t, U_t >} - \frac{1}{2} < \mathcal{H}_t, U_t > + \xi_t) \ , \quad S_0 = s \ , \end{aligned} \tag{6-6}$$

for each $t \in \mathcal{T}\setminus\{0\}$.

For each $t \in \mathcal{T}\setminus\{0\}$, Y_t denotes the continuously compounded one-period rate of return $\ln(\frac{S_t}{S_{t-1}})$ of the stock S. Clearly, Y_t is measurable with respect to the observable information set $\mathcal{F}_t^S$. Then, under $\mathcal{P}$, the dynamics of Y_t is governed by the following GARCH(p, q) model with innovations having a finite mixture of infinitely divisible distributions:

$$\begin{aligned} Y_t &= r_t + \lambda_t \sqrt{<\mathcal{H}_t, U_t>} - \frac{1}{2} <\mathcal{H}_t, U_t> + \xi_t \ , \\ \xi_t | \mathcal{F}_{t-1}^S &\sim MID(p_1, \ldots, p_K; h_{1t}, \ldots, h_{Kt}) \ , \\ h_{kt} &= \alpha_{k0} + \sum_{j=1}^{q} \alpha_{kj} \xi_{t-j}^2 + \sum_{i=1}^{K} \sum_{l=1}^{p} \beta_{kil} h_{i,t-1} \ . \end{aligned} \tag{6-7}$$

The above model is called MID(K)-GARCH (p, q) model. It can incorporate the general form NM(K)-GARCH(p, q) specified by Haas, Mittnik and Paolella (2002) (32) without explanatory variables and the NM(K)-GARCH(1,1) examined by Alexander and Lazar (2006)(1). The model by Haas, Mittnik and Paolella (2002)(32) can incorporate the cross dependence of individual conditional variances by assuming the inter-dependent autoregressive evolution of the time series of conditional variances. Alexander and Lazar (2006)(1) mentioned that the mixture GARCH model can be related to other important GARCH models with non-normal innovations and the class of Markov-Switching GARCH models. In fact, the MID(K)-GARCH (p, q) model can be considered a particular case of a general class of regime-switching GARCH(p, q) models with innovations having a finite mixture of infinitely divisible distributions and the GARCH dynamics driven by a hidden Markov chain model. The MID(K)-GARCH (p, q) model is different from the one in Alexander and Lazar (2006)(1) in two aspects. First, the proposed mixture GARCH model here has a time-varying drift depending on the conditional volatility of the return's process while the mixture GARCH model in Alexander and Lazar (2006)(1) has a constant drift. Second, the proposed mixture GARCH model has a general mixing kernel, which is specified by an infinite divisible distribution with a finite moment generating function. Hence, it is flexible enough to incorporate the Normal-Mixture (NM) GARCH models in Alexander and Lazar (2006)(1), the GARCH models with innovations having a finite mixture of shifted gamma distributions and a finite mixture of shifted Inverse Gaussian distributions, and others. This provides market practitioners with a great deal of flexibility in modelling different empirical "stylised" behaviors of asset price dynamics, such the skewed behavior and the leptokurtosis behavior of the distribution of asset's returns.

We shall present the discrete-time doubly stochastic Esscher transform in the sequel. The doubly stochastic Esscher transform is defined by the product of two stochastic processes. The idea is similar with the conditional Esscher transform introduced in Bühlmann et al. (1996)(9). Elliott, Siu and Chan (2006)(23) considered a similar type of doubly stochastic Esscher transform, namely the regime-switching Esscher transform, in the context of a continuous-time regime-switching Geometric Brownian Motion model. The discrete-time version of the regime-switching Esscher transform has been adopted in Elliott, Siu and Chan (2006)(23) for pricing options under a discrete-time Markov-switching GARCH models.

First, we consider a real-valued stochastic process $\{\Theta_t\}_{t\in\mathcal{T}}\backslash\{0\}$ defined on $(\Omega,\mathcal{F})$. We assume that for each $t\in\mathcal{T}\backslash\{0\}$, Θ_t is measurable with respect to $\mathcal{F}_t^U$.

Let $M_Y(t,\Theta_t)$ denote the moment generating function of Y_t given $\mathcal{F}_{t-1}^S$ under $\mathcal{P}$; that is,

$$M_{Y|\mathcal{F}_{t-1}^S}(t,\Theta_t) := E_{\mathcal{P}}(e^{\Theta_t Y_t}|\mathcal{F}_{t-1}^S)\ . \tag{6-8}$$

We assume that $E_{\mathcal{P}}(e^{\Theta_t Y_t}|\mathcal{F}_{t-1}^S)<\infty$, for $t\in\mathcal{T}\backslash\{0\}$. As in Bühlmann et al. (1996)(9), we define a sequence $\{\Lambda_t\}_{t\in\mathcal{T}}$ with $\Lambda_0=1$ as follows:

$$\Lambda_t=\prod_{u=1}^{t}\frac{e^{\Theta_u Y_u}}{M_Y(u,\Theta_u)}\ ,\ t\in\mathcal{T}\backslash\{0\}\ , \tag{6-9}$$

where Λ_t is specified by a product of two stochastic processes $\{\Theta_t\}_{t\in\mathcal{T}\backslash\{0\}}$ and $\{Y_t\}_{t\in\mathcal{T}\backslash\{0\}}$.

Lemma 36 $\{\Lambda_t\}_{t\in\mathcal{T}}$ *is a* $(\mathcal{F},\mathcal{P})$*-martingale.*

Proof: From its definition, Λ_t is $\mathcal{F}_{t+1}$-measurable, for each $t\in\mathcal{T}$ and that

$$\begin{aligned} & E_{\mathcal{P}}\left(\frac{\Lambda_t}{\Lambda_{t-1}}\middle|\mathcal{F}_{t-1}^S\right)\\ =\ & E_{\mathcal{P}}\left[\frac{e^{\Theta_t Y_t}}{E_{\mathcal{P}}(e^{\Theta_t Y_t}|\mathcal{F}_{t-1}^S)}\middle|\mathcal{F}_{t-1}^S\right]\\ =\ & 1\ ,\ \mathcal{P}-\text{a.s.} \end{aligned} \tag{6.1}$$

Hence, the result follows. □

Then, the doubly stochastic Esscher transform $\mathcal{P}_\Theta\sim\mathcal{P}$ on $\mathcal{F}_t$ with respect to $\{\Theta_1,\Theta_2,\ldots,\Theta_{t-1}\}$ is given by:

$$\left.\frac{d\mathcal{P}_\Theta}{d\mathcal{P}}\right|_{\mathcal{F}_t} := \Lambda_{t-1}\ ,\qquad t=2,3,\ldots,T+1\ . \tag{6-2}$$

Let $M_Y(z,t;\Theta)$ be the moment generating function of Y_t given $\mathcal{F}_t$ under $\mathcal{P}_\Theta$. Then, we have the following lemma:

Lemma 37

$$M_Y(z,t;\Theta) = \frac{M_Y(t, z+\Theta_t)}{M_Y(t,\Theta_t)} \ . \tag{6-3}$$

Proof: The proof is adapted to the argument in Elliott, Siu and Chan (2006)(23) in the case of mixture GARCH models. By Baye's rule,

$$\begin{aligned} M_Y(z,t;\Theta) &= E_{\mathcal{P}_\Theta}(e^{zY_t}|\mathcal{F}^S_{t-1}) \\ &= \frac{E_{\mathcal{P}}(\Lambda_t e^{zY_t}|\mathcal{F}^S_{t-1})}{E_{\mathcal{P}}(\Lambda_t|\mathcal{F}^S_{t-1})} \\ &= E_{\mathcal{P}}\left(\frac{\Lambda_t}{\Lambda_{t-1}} e^{zY_t}\middle|\mathcal{F}^S_{t-1}\right) \\ &= \frac{E_{\mathcal{P}}(e^{(z+\Theta_t)Y_t}|\mathcal{F}^S_{t-1})}{E_{\mathcal{P}}(e^{\Theta_t Y_t}|\mathcal{F}^S_{t-1})} \\ &= \frac{M_Y(t,z+\Theta_t)}{M_Y(t,\Theta_t)} \ . \end{aligned} \tag{6-4}$$

Hence, the result follows. □

Harrison and Krep (1979) (33)and Harrison and Pliska (1981, 1983)(34, 35) provided a solid theoretical foundation to establish the relationship between the absence of arbitrage and the existence of an equivalent martingale measure using the modern language of probability. They introduced the fundamental theorem of asset pricing which states that the absence of arbitrage is equivalent to the existence of an equivalent martingale measure under which the discounted asset price process is a martingale. The fundamental theorem of asset pricing was then further extended by Dybyig and Ross (1987)(20), Back and Pliska (1991)(2), Schachermayer (1992)(39) and Delbaen and Schachermayer (1994)(15). Back and Pliska (1991)(2) has shown that the absence of arbitrage is equivalent to the existence of an equivalent martingale measure in a discrete-time and an infinite-state-space setting.

We shall determine an equivalent martingale measure using the doubly stochastic Esscher transform $\mathcal{P}_\Theta$. The sufficient condition on Θ_t for $\mathcal{P}_\Theta$ to be an equivalent martingale measure is presented in the following proposition.

Proposition 38 *(Martingale Condition) Suppose Θ_t satisfies the following condition:*

$$\frac{M_Y(t,\Theta_t+1)}{M_Y(t,\Theta_t)} = e^{r_t} \ . \tag{6-5}$$

Then, the discounted price process $\{\exp(-\sum_{u=1}^{t} r_u)S_t\}_{t\in\mathcal{T}}$ *is a* $(\mathcal{F},\mathcal{P}_\Theta)$-*martingale.*

Proof:

$$\begin{aligned}
& E_{\mathcal{P}_\Theta}\left(e^{-\sum_{u=1}^{t} r_u} S_t \middle| \mathcal{F}^S_{t-1}\right) \\
= \; & e^{-\sum_{u=1}^{t-1} r_u} S_{t-1} E_{\mathcal{P}_\Theta}(e^{Y_t - r_t} | \mathcal{F}^S_{t-1}) \\
= \; & e^{-\sum_{u=1}^{t-1} r_u} S_{t-1} E_{\mathcal{P}}\left[\left(\frac{\Lambda_t}{\Lambda_{t-1}}\right) e^{Y_t - r_t} \middle| \mathcal{F}^S_{t-1}\right] \\
= \; & e^{-\sum_{u=1}^{t-1} r_u} S_{t-1} e^{-r_t} \frac{E_{\mathcal{P}}(e^{(\Theta_t+1)Y_t} | \mathcal{F}^S_{t-1})}{E_{\mathcal{P}}(e^{\Theta_t Y_t} | \mathcal{F}^S_{t-1})} \\
= \; & e^{-\sum_{u=1}^{t-1} r_u} S_{t-1} e^{-r_t} \frac{M_Y(t, \Theta_t + 1)}{M_Y(t, \Theta_t)} \\
= \; & e^{-\sum_{u=1}^{t-1} r_u} S_{t-1} \;, \; \mathcal{P} - \text{a.s.}
\end{aligned} \tag{6-6}$$

Hence, the result follows.

□

Note that the existence and uniqueness of Θ_t satisfying the condition in Proposition 2.3 can be proved by following the arguments in Chan and van-der Hoek (2003) (11). The martingale condition with respect to the enlarged filtration $\{\mathcal{F}_t\}_{t\in\mathcal{T}}$ is stronger than that with respect to the observable information structure $\{\mathcal{F}^S_t\}_{t\in\mathcal{T}}$. In other words, if there exists a probability measure $\mathcal{P}_\Theta$ satisfying the martingale condition with respect to $\{\mathcal{F}_t\}_{t\in\mathcal{T}}$, $\mathcal{P}_\Theta$ also satisfies the martingale condition with respect to $\{\mathcal{F}^S_t\}_{t\in\mathcal{T}}$. This can be verified easily by the double expectation formula for conditional expectations.

Then, given $\mathcal{F}_t := \mathcal{F}^S_{t-1} \vee \mathcal{F}^U_t$, the price of a European-style contingent claim written on the underlying stock S with payoff $V(S_T)$ at maturity T is given by:

$$V(t-1, T|\mathcal{F}_t) = E_{\mathcal{P}_\Theta}\left[\exp\left(-\sum_{u=t}^{T} r_u\right) V(S_T) \middle| \mathcal{F}_t\right] . \tag{6-7}$$

Given the observable information $\mathcal{F}^S_{t-1}$, the price of the option can be determined as follows:

$$\begin{aligned}
V(t-1, T|\mathcal{F}^S_{t-1}) &= E_{\mathcal{P}_\Theta}[V(t-1, T|\mathcal{F}_t)|\mathcal{F}^S_{t-1}] \\
&= E_{\mathcal{P}_\Theta}\left[\exp\left(-\sum_{u=t}^{T} r_u\right) V(S_T) \middle| \mathcal{F}^S_{t-1}\right] .
\end{aligned}$$

In order to compute the option price $V(t-1, T|\mathcal{F}^S_{t-1})$ under a given parametric distribution for the innovations, we first need to estimate the unknown parameters of the mixture GARCH model using the observed market price data for S under $\mathcal{P}$ and then simulate the terminal stock prices from the mixture GARCH model under $\mathcal{P}_\Theta$ for approximating $V(t-1,T|\mathcal{F}^S_{t-1})$ via Monte Carlo simulation coupled with its control variates technique. Due to the structure of the mixture GARCH model, its likelihood function can be determined completely given the observed market price data $\mathcal{F}^S_{t-1}$. Hence, we do not need to use information from the values of the hidden mixing process $\{U_t\}_{t\in\mathcal{T}\backslash\{0\}}$ for the estimation part. However, in the Monte Carlo simulation, we need to generate the values of the process $\{U_t\}_{t\in\mathcal{T}\backslash\{0\}}$ and the values of the innovations process $\{\xi_t\}_{t\in\mathcal{T}\backslash\{0\}}$ for simulating the terminal stock prices. In section 4, we shall discuss in some detail the estimation and simulation procedures for computing the prices of a European call option under two parametric mixture GARCH models described in Section 3.

We shall justify the pricing result by a power utility maximization problem with respect to $\mathcal{F}_t$. For each $t\in\mathcal{T}\backslash\{0\}$, let ζ_t denote a random variable measurable with respect to $\mathcal{F}_t$. Then, for each $t\in\mathcal{T}\backslash\{0\}$, we consider a power utility function u_t with the stochastic risk-averse parameter ζ_t; that is, for each $t\in\mathcal{T}\backslash\{0\}$,

$$u_t(x)=\begin{cases}\frac{x^{1-\zeta_t}}{1-\zeta_t} & \text{if } \zeta_t\neq 1,\\ \ln x & \text{if } \zeta_t=1.\end{cases}$$

Equivalently, the power utility function can be written as:

$$u_t(x)=\frac{x^{1-\zeta_t}}{1-\zeta_t}I_{\{\zeta_t\neq 1\}}+(\ln x)I_{\{\zeta_t=1\}}\ . \tag{6-8}$$

We call the above power utility function a stochastic power utility function. We assume that an economic agent can adjust the stochastic risk-averse parameter ζ_t based on $\mathcal{F}_t$. For each $t\in\mathcal{T}\backslash\{0\}$, following Gerber and Shiu (1994) (31), we impose the following assumptions:

1. The economic agent has m_t units of stock S and η_t units of the option from time $t-1$ to time t, where $m_t,\eta_t\in\mathcal{F}_t$.

2. $\tilde{V}(t-1,T|\mathcal{F}_t)$ represents the economic agent's equilibrium price at time $t-1$ of the option with maturity at time T so that it is optimal for the agent not to buy or sell any units of the option at time $t-1$.

For each $t\in\mathcal{T}\backslash\{0\}$, the conditional expected utility function U_t on the economic agent's wealth at time t given $\mathcal{F}_t$ under $\mathcal{P}$ is given by:

$$\begin{aligned} & U_t(\eta_t) \\ = \; & E_{\mathcal{P}}\left\{u_t\left(m_tS_t + \eta_t\left[\tilde{V}(t,T|\mathcal{F}_{t+1}) - e^{r_t}\tilde{V}(t-1,T|\mathcal{F}_t)\right]\right)\Big|\mathcal{F}_t\right\} , \end{aligned} \tag{6-9}$$

where η_t is the choice variable.

The following proposition justifies the pricing result by the doubly stochastic Esscher transform.

Proposition 39 *For each* $t \in \mathcal{T}\backslash\{0\}$,

$$V(t-1,T|\mathcal{F}_t) = \tilde{V}(t-1,T|\mathcal{F}_t) \ , \tag{6-10}$$

and

$$\zeta_t = -\Theta_t \ . \tag{6-11}$$

Proof: We follow the argument in Gerber and Shiu (1994)(31). The optimal condition of the conditional expected utility function on the economic agent's wealth is equivalent to that $U_t(\eta_t)$ attains its maximum value when $\eta_t = 0$, for each $t \in \mathcal{T}\backslash\{0\}$. Mathematically, this can be translated to:

$$U_t{}'(\eta_t)|_{\eta_t=0} = 0 \ , \tag{6-12}$$

where $U_t{}'$ is the derivative of U_t with respect to η_t.

For simplicity, we write $\tilde{V}_{t-1}$ for $\tilde{V}(t-1,T|\mathcal{F}_t)$. Then, the optimal condition implies that

$$\tilde{V}_{t-1} = e^{-r_t}\frac{E_{\mathcal{P}}[\tilde{V}_t u_t{}'(m_tS_t)|\mathcal{F}_t]}{E[u_t{}'(m_tS_t)|\mathcal{F}_t]} \ , \tag{6-13}$$

where $u_t{}'$ is the derivative of u_t with respect to η_t.
We notice that

$$u_t{}'(x) = x^{-\zeta_t}I_{\{\zeta_t\neq 1\}} + (1/x)I_{\{\zeta_t=1\}} \ . \tag{6-14}$$

Then,

$$\begin{aligned} \tilde{V}_{t-1} &= e^{-r_t}\left\{\frac{E[\tilde{V}_t(m_tS_t)^{-\zeta_t}|\mathcal{F}_t]}{E[(m_tS_t)^{-\zeta_t}|\mathcal{F}_t]}I_{\{\zeta_t\neq 1\}} + \frac{E[\tilde{V}_t/(m_tS_t)|\mathcal{F}_t]}{E[1/(m_tS_t)|\mathcal{F}_t]}I_{\{\zeta_t=1\}}\right\} \\ &= e^{-r_t}\frac{E[\tilde{V}_t(m_tS_t)^{-\zeta_t}|\mathcal{F}_t]}{E[(m_tS_t)^{-\zeta_t}|\mathcal{F}_t]} \ . \end{aligned} \tag{6-15}$$

Since the above result applies for any option, it can be applied for the underlying asset S. Hence,

$$\begin{aligned} S_{t-1} &= e^{-r_t}\frac{E(S_t^{1-\zeta_t}|\mathcal{F}_t)}{E(S_t^{-\zeta_t}|\mathcal{F}_t)} \\ &= e^{-r_t}S_{t-1}\frac{E(e^{(1-\zeta_t)Y_t}|\mathcal{F}_t)}{E(e^{-\zeta_t Y_t}|\mathcal{F}_t)} \\ &= e^{-r_t}S_{t-1}\frac{M_Y(t,1-\zeta_t)}{M_Y(t,-\zeta_t)} \ . \end{aligned} \tag{6-16}$$

Then,

$$\frac{M_Y(t,1-\zeta_t)}{M_Y(t,-\zeta_t)} = e^{r_t} \ . \tag{6-17}$$

Hence, by uniqueness of Θ_t satisfying the martingale condition,

$$\zeta_t = -\Theta_t \ , \tag{6-18}$$

and

$$\tilde{V}_{t-1} = V_{t-1} \ . \tag{6-19}$$

□

6.2 Parametric Cases

In this section, we deal with some parametric cases, namely the Normal-Mixture (NM) GARCH model and the mixture GARCH model with innovations having a finite mixture of shifted gamma distributions. We consider K-component first-order mixture GARCH models in this section. The derivation of the pricing result for the case when the innovations have a finite mixture of shifted inverse Gaussian distributions is very similar to that of the finite mixture of shifted gamma distributions. We are able to preserve the parametric forms of the distributions for the GARCH innovations under the change of probability measures using the doubly stochastic Esscher tranform in Section 2.

6.2.1 Normal-Mixture (NM) GARCH models

First, we assume that under $\mathcal{P}$, the conditional distribution of ξ_t given $\mathcal{F}^S_{t-1}$ is a normal distribution $N(0, <\mathcal{H}_t, U_t>)$ with mean 0 and conditional variance $< \mathcal{H}_t, U_t >$, for each $t \in \mathcal{T} \backslash \{0\}$. Then, the conditional distribution of ξ_t given $\mathcal{F}^S_{t-1}$ under $\mathcal{P}$ is given by the following finite mixture of normal distributions:

$$F_\xi(x|\mathcal{F}^S_{t-1}) = \sum_{k=1}^{K} p_k N(x|0, h_{kt}) \ , \quad \sum_{k=1}^{K} p_k = 1 \ . \tag{6-20}$$

The conditional distribution of Y_t given $\mathcal{F}^S_{t-1}$ under $\mathcal{P}$ is a normal distribution $N(r_t + \lambda_t\sqrt{< \mathcal{H}_t, U_t >} - \frac{1}{2} < \mathcal{H}_t, U_t >, < \mathcal{H}_t, U_t >)$ with mean $r_t + \lambda_t\sqrt{< \mathcal{H}_t, U_t >} - \frac{1}{2} < \mathcal{H}_t, U_t >$ and variance $< \mathcal{H}_t, U_t >$, where the distribution of U_t is given by:

$$P(U_t = e_k) = p_k \ ,$$

and the dynamics of h_k is governed by:

$$h_{kt} = \alpha_{k0} + \alpha_{k1}\xi^2_{t-1} + \sum_{k=1}^{K} \beta_{ki1} h_{i,t-1} \ ,$$

for each $k = 1, 2, \ldots, K$. Hence we have

$$\begin{aligned} Y_t &= r_t + \lambda_t\sqrt{< \mathcal{H}_t, U_t >} - \frac{1}{2} < \mathcal{H}_t, U_t > + \xi_t \ , \\ \xi_t|\mathcal{F}^S_{t-1} &\sim NM(p_1, \ldots, p_K; h_{1t}, \ldots, h_{Kt}) \ , \\ h_{kt} &= \alpha_{k0} + \alpha_{k1}\xi^2_{t-1} + \sum_{i=1}^{K} \beta_{ki1} h_{i,t-1} \ . \end{aligned}$$

Then,

$$M_{Y_t|\mathcal{F}^S_{t-1}}(t, \Theta_t) = \exp\left[\Theta_t\left(r_t + \lambda_t\sqrt{< \mathcal{H}_t, U_t >} - \frac{1}{2} < \mathcal{H}_t, U_t >\right) + \frac{1}{2}\Theta_t^2 < \mathcal{H}_t, U_t > \right] \tag{6-21}$$

From the martingale condition proposition, the risk-neutralized stochastic Esscher parameter Θ_t is given by:

$$\Theta_t = -\frac{\lambda_t}{\sqrt{< \mathcal{H}_t, U_t >}} = -\lambda_t \sum_{k=1}^{K} \frac{1}{\sqrt{h_{kt}}} < U_t, e_k > \ , \quad t \in \mathcal{T} \backslash \{0\} \ . \tag{6-22}$$

The moment generating function $M_{Y_t|\mathcal{F}^S_{t-1}}(z, t; \Theta)$ is given by:

$$M^\theta_{Y_t|\mathcal{F}^S_{t-1}}(z, t; \Theta) \ = \ \exp[z(r - \frac{1}{2} < \mathcal{H}_t, U_t >) + \frac{1}{2}z^2 < \mathcal{H}_t, U_t >] \ . \tag{6-23}$$

Hence, under $\mathcal{P}_\Theta$, the conditional distribution of Y_t given $\mathcal{F}^S_{t-1}$ is a normal distribution with conditional mean $r_t - \frac{1}{2} < \mathcal{H}_t, U_t >$ and conditional variance $< \mathcal{H}_t, U_t >$, where

$$P(U_t = e_k) = p_k \ , \tag{6-24}$$

and

$$h_{kt} \;=\; \alpha_{k0} + \alpha_{k1}\xi_{t-1}^2 + \sum_{i=1}^{K} \beta_{ki1} h_{i,t-1} \ . \tag{6-25}$$

for each $k = 1, 2, \ldots, K$.

Note that $\xi_t | \mathcal{F}_{t-1}^S \sim N(-\lambda_t\sqrt{< \mathcal{H}_t, U_t >}, < \mathcal{H}_t, U_t >)$ under $\mathcal{P}_\Theta$. Let $\epsilon_{t-1} := \xi_{t-1} + \lambda_{t-1}\sqrt{< \mathcal{H}_{t-1}, U_{t-1} >}$. Then, under $\mathcal{P}_\Theta$, $\epsilon_t | \mathcal{F}_{t-1}^S \sim N(0, < \mathcal{H}_t, U_t >)$. We can write the dynamics of h_{kt} as follows:

$$h_{kt} \;=\; \alpha_{k0} + \alpha_{k1}(\epsilon_{t-1} - \lambda_{t-1}\sqrt{< \mathcal{H}_{t-1}, U_{t-1} >})^2 + \sum_{i=1}^{K} \beta_{ki1} h_{i,t-1} \ . \tag{6-26}$$

We have just proved the following Theorem:

Theorem 40 *Let*

$$\begin{aligned} Y_t &= r_t + \lambda_t\sqrt{< \mathcal{H}_t, U_t >} - \frac{1}{2} < \mathcal{H}_t, U_t > + \xi_t \ , \\ \xi_t | \mathcal{F}_{t-1}^S &\sim NM(p_1, \ldots, p_K; h_{1t}, \ldots, h_{Kt}) \ , \\ h_{kt} &= \alpha_{k0} + \alpha_{k1}\xi_{t-1}^2 + \sum_{i=1}^{K} \beta_{ki1} h_{i,t-1} \ . \end{aligned}$$

be the model in the real world probability P. Then, the risk neutral version of the model is:

$$\begin{aligned} Y_t &= r_t - \frac{1}{2}\langle H_t, U_t\rangle + \epsilon_t \\ \epsilon_t | \mathcal{F}_{t-1}^S &\sim N\left(0, \langle H_t, U_t\rangle\right) \\ h_{kt} &= \alpha_{k0} + \alpha_{k1}\left(\epsilon_{t-1} - \lambda_{t-1}\sqrt{\langle H_{t-1}, U_{t-1}\rangle}\right)^2 + \sum_{i=1}^{K} \beta_{ki1} h_{i,t-1} \end{aligned}$$

6.2.2
Shifted-Gamma-Mixture (SGM) GARCH models

We assume that the innovations for the general mixture GARCH model in Section 2 follow a finite mixture of shifted gamma distributions. First, for each $k = 1, 2, \ldots, K$, we consider a sequence of i.i.d. random variables $\{X_t^{(k)}\}_{t \in \mathcal{T}\setminus\{0\}}$

with common distribution being a gamma distribution $Ga(a_k, b)$ with shape parameter a_k and scale parameter b. For each $t \in \mathcal{T}\backslash\{0\}$, we define the random variable $\nu_t^{(k)}$ by standardizing the gamma random variable X_t as follows:

$$\nu_t^{(k)} := \frac{X_t^{(k)} - a_k/b}{\sqrt{a_k/b^2}} \ . \tag{6-27}$$

Note that $\nu_t^{(k)}$ follows a standard shifted gamma distribution $SGa(\cdot|0,1)$ with zero mean and unit variance, for each $k = 1, 2, \ldots, K$. For each $t \in \mathcal{T}\backslash\{0\}$, let $V_t := (\nu_1, \nu_2, \ldots, \nu_K)$ and ν_t be defined as follows:

$$\nu_t :=< V_t, U_t >= \sum_{k=1}^{K} < V_t, e_k >< U_t, e_k >= \sum_{k=1}^{K} \nu_k < U_t, e_k > \ . \tag{6-28}$$

Then, ν_t follows a standard shifted gamma distribution $SGa(\cdot|0,1)$ with zero mean and unit variance.

We assume that $\xi_t := -\sqrt{< \mathcal{H}_t, U_t >}\nu_t$; then, $\xi_t = -\sum_{k=1}^{K} \sqrt{h_{kt}}\nu_t^{(k)} < U_t, e_k >$. Hence, the conditional distribution of ξ_t given $\mathcal{F}_{t-1}^S$ under $\mathcal{P}$ is minus a shifted gamma distribution $SGa(\cdot|0, < \mathcal{H}_t, U_t >)$ with zero mean and variance $< \mathcal{H}_t, U_t >= \sum_{k=1}^{K} h_{kt} < U_t, e_k >$. Then, the conditional distribution of ξ_t given $\mathcal{F}_{t-1}^S$ is given by the following finite mixture of shifted gamma distributions:

$$F_\xi(x|\mathcal{F}_{t-1}^S) = \sum_{k=1}^{K} p_k SGa(x|0, h_{kt}) \ , \quad \sum_{k=1}^{K} p_k = 1 \ , \tag{6-29}$$

where $SGa(\cdot|0, h_{kt})$ is the probability distribution of a shifted gamma distribution with mean 0 and variance h_{kt}.

We write $\xi_t|\mathcal{F}_{t-1}^S \sim -MSGa(p_1, \ldots, p_K; h_{1t}, \ldots, h_{Kt})$. Hence, under $\mathcal{P}$, the conditional distribution of Y_t given $\mathcal{F}_{t-1}^S$ is minus a shifted gamma distribution $SGa(\cdot|r_t + \lambda_t\sqrt{< \mathcal{H}_t, U_t >} - \frac{1}{2} < \mathcal{H}_t, U_t >, < \mathcal{H}_t, U_t >)$ with mean $r_t + \lambda_t\sqrt{< \mathcal{H}_t, U_t >} - \frac{1}{2} < \mathcal{H}_t, U_t >$ and variance $< \mathcal{H}_t, U_t >$. Then, under $\mathcal{P}$, the dynamics of Y_t is governed by the following K-component first-order mixture GARCH model with innovations having a finite mixture of shifted gamma distributions:

$$\begin{aligned} Y_t &= r_t + \lambda_t\sqrt{< \mathcal{H}_t, U_t >} - \frac{1}{2} < \mathcal{H}_t, U_t > + \xi_t \ , \\ \xi_t|\mathcal{F}_{t-1}^S &\sim -MSGa(p_1, \ldots, p_K; h_{1t}, \ldots, h_{Kt}) \ , \\ h_{kt} &= \alpha_{k0} + \alpha_{k1}\xi_{t-1}^2 + \sum_{i=1}^{K} \beta_{ki1} h_{i,t-1} \ . \end{aligned} \tag{6-30}$$

Let $A := (a_1, a_2, \dots, a_K)$ and A_t be defined as follows:

$$A_t :=< A, U_t >= \sum_{k=1}^{K} a_k < U_t, e_k > \quad . \tag{6-31}$$

We can also write the dynamics of $\{Y_t\}_{t \in \mathcal{T}\setminus\{0\}}$ under $\mathcal{P}$ as follows:

$$\begin{aligned} Y_t &= \sum_{k=1}^{K} \Bigg(r_t + \lambda_t \sqrt{< \mathcal{H}_t, e_k >} - \frac{1}{2} < \mathcal{H}_t, e_k > + \sqrt{< A, e_k >< \mathcal{H}_t, e_k >} \\ &\quad - b \sqrt{\frac{< \mathcal{H}_t, e_k >}{< A, e_k >}} X_t^{(k)} \Bigg) < U_t, e_k > \\ &= \sum_{k=1}^{K} \left(r_t + \lambda_t \sqrt{h_{kt}} - \frac{1}{2} h_{kt} + \sqrt{a_k h_{kt}} - b \sqrt{\frac{h_{kt}}{a_k}} X_t^{(k)} \right) < U_t, e_k > \end{aligned} \tag{6-32}$$

Define B_t as $\sqrt{\frac{A_t}{H_t}}$; that is,

$$B_t = \sum_{k=1}^{K} \sqrt{\frac{a_k}{h_{kt}}} < U_t, e_k > \quad . \tag{6-33}$$

Then, the conditional distribution of Y_t given $\mathcal{F}_{t-1}^S$ is minus a shifted gamma distribution with shape parameter A_t, scale parameter B_t and shifted parameter $-r_t - \lambda_t \sqrt{< \mathcal{H}_t, U_t >} + \frac{1}{2} < \mathcal{H}_t, U_t > - \sqrt{< A, U_t >< \mathcal{H}_t, U_t >}$. Hence, the moment generating function of Y_t given $\mathcal{F}_{t-1}^S$ under $\mathcal{P}$ is given by:

$$M_{Y|\mathcal{F}_{t-1}^S}(t, \Theta_t) = \left(\frac{B_t}{B_t + \Theta_t} \right)^{A_t} e^{\left(r_t + \lambda_t \sqrt{H_t} - \frac{1}{2} H_t + \sqrt{A_t H_t} \right) \Theta_t} \quad , \tag{6-34}$$

where $0 < B_t + \Theta_t$.

Hence, under $\mathcal{P}_\Theta$, the moment generating function of Y_t given $\mathcal{F}_{t-1}^S$ is given by:

$$M_{Y|\mathcal{F}_{t-1}^S}^{\theta}(t, z; \Theta_t) = \left(\frac{B_t + \Theta_t}{B_t + \Theta_t - z} \right)^{A_t} e^{\left(r_t + \lambda_t \sqrt{H_t} - \frac{1}{2} H_t + \sqrt{A_t H_t} \right) z} \quad . \tag{6-35}$$

From the martingale condition, the risk-neutralized stochastic Esscher parameter Θ_t is given by:

$$\Theta_t = \left[e^{\frac{\lambda_t \sqrt{H_t} - \frac{1}{2} H_t + \sqrt{A_t H_t}}{A_t}} - 1 \right]^{-1} - B_t \tag{6-36}$$

Then, we define the parameter $\tilde{B}_t$ as follows:

$$
\begin{aligned}
\tilde{B}_t &:= B_t + \Theta_t \\
&= \left[e^{\frac{\lambda_t \sqrt{H_t} - \frac{1}{2} H_t + \sqrt{A_t H_t}}{A_t}} - 1 \right]^{-1} \\
&= \sum_{k=1}^{K} \left[e^{\left(\frac{\lambda_t \sqrt{h_{kt}} - \frac{1}{2} h_{kt} + \sqrt{a_k h_{kt}}}{a_k} \right)} - 1 \right]^{-1} < U_t, e_k > \ , \ t \in \mathcal{T} \backslash \{0\} \ .
\end{aligned}
$$

Hence, under $\mathcal{P}_\Theta$, $Y_t | \mathcal{F}^S_{t-1}$ follows a minus shifted gamma distribution with shape parameter A_t, scale parameter $\tilde{B}_t$ and shifted parameter $-r_t - \lambda_t \sqrt{< \mathcal{H}_t, U_t >} + \frac{1}{2} < \mathcal{H}_t, U_t > - \sqrt{< A, U_t >< \mathcal{H}_t, U_t >}$, where

$$
P(U_t = e_k) = p_k \ . \tag{6-37}
$$

Let X^q_t denote a random variable such that $X^q_t | \mathcal{F}^S_{t-1} \sim -\frac{1}{\tilde{B}_t} Ga(A_t, 1)$. Then,

$$
h_{kt} = \alpha_{k0} + \alpha_{k1} \left(X^q_{t-1} + \sqrt{A_{t-1} H_{t-1}} \right)^2 + \sum_{i=1}^{K} \beta_{ki1} h_{i,t-1} \ ,
$$

where $k = 1, 2, \ldots, K$.

From the discussion above we have:

Theorem 41 *Let the model under P be:*

$$
\begin{aligned}
Y_t &= r_t + \lambda_t \sqrt{< \mathcal{H}_t, U_t >} - \frac{1}{2} < \mathcal{H}_t, U_t > + \xi_t \ , \\
\xi_t | \mathcal{F}^S_{t-1} &\sim -MSGa(p_1, \ldots, p_K; h_{1t}, \ldots, h_{Kt}) \\
h_{kt} &= \alpha_{k0} + \alpha_{k1} \xi^2_{t-1} + \sum_{i=1}^{K} \beta_{ki1} h_{i,t-1} \ .
\end{aligned}
$$

Then, under the risk neutral measure the model becomes:

$$
\begin{aligned}
Y_t &= r_t + \lambda_t \sqrt{< \mathcal{H}_t, U_t >} - \frac{1}{2} < \mathcal{H}_t, U_t > + \sqrt{< A, U_t >< \mathcal{H}_t, U_t >} + X^q_t \ , \\
X^q_t | \mathcal{F}^S_{t-1} &\sim -\frac{1}{\tilde{B}_t} Ga(A_t, 1) \\
h_{kt} &= \alpha_{k0} + \alpha_{k1} \left(X^q_{t-1} + \sqrt{A_{t-1} H_{t-1}} \right)^2 + \sum_{i=1}^{K} \beta_{ki1} h_{i,t-1} \ ,
\end{aligned}
$$

Finally, we note that the scale parameter b does not appear in both the real-world and the risk-neutral dynamics due to the structure of the mixture shifted-gamma model for the innovations.

6.3 Empirical Results and Discussion

In this section we conduct a simulation exercise in order to compare the option prices induced by assuming different processes for the log-returns of the underlying asset.

We consider the two-component mixture GARCH(1, 1) models with a finite mixture of normal innovations and a finite mixture of shifted gamma innovations for illustration. We employ the daily close values of S&P500 index from November 7, 1980, to November 4, 2005, 25 years of data with 6,130 observations. The data were obtained from the database in Yahoo Finance. We estimate the parameters underlying the option pricing models for empirical studies using the 25-year S&P500 index data. We assume that the risk-free interest rate r is zero and that the unit risk premium is a constant λ throughout this section as in Duan (1995)(17) and Siu et al. (2004)(43).

We shall discuss the estimation procedures and the estimation results for the two-mixture GARCH models in the sequel. Tong (1990)(45) and Fan and Yao (2003)(27) mentioned that the conditional maximum likelihood estimation enjoys some desirable sampling and asymptotic properties. For estimating the two-component mixture GARCH model with a finite mixture of normal innovations, we employ the conditional maximum likelihood estimation. For estimating the two-component mixture GARCH models with a finite mixture of shifted gamma distributions, we use a two-stage estimation procedure adopted in Siu et al. (2004) (43). In the first-stage, the quasi-maximum likelihood estimation (QMLE) is used to estimate the mixture GARCH parameters, namely p_k, α_{k0}, α_{k1}, β_{k11} and β_{k21}, where $k = 1, 2$. The QMLE is an approximation to the exact MLE by assuming that the innovations of a time series model follow a normal distribution even though the "true" innovations may not be normally distributed. In the QMLE, the exact likelihood is replaced by the normal likelihood. Franses and van Dijk (2000) (30) mentioned that if it is not sure whether the specified parametric assumption for the GARCH innovations is correct, the QMLE can be employed for estimation. Fan and Yao (2003) (27) pointed out that one does not know the "true" distribution for the innovations in practice and, hence, the QMLE can provide a practical way to estimate the parameters in a GARCH model. In fact, the QMLE method can provide market practitioners with a convenient way to estimate non-normal GARCH models since many standard computing and statistical packages have included the conditional MLE for GARCH models with normal innovations. One undesirable feature of the QMLE is that the standard errors of the estimates are larger than those obtained by the exact MLE in the finite sample case. However, the QMLE also enjoys some desirable asymptotic properties and the standard errors of the estimates from the QMLE can be reduced by estimating the model using a large data set. In the second-stage of our estimation, we adopt the method of moments approach to estimate the unknown parameters in the finite mixture of shifted gamma distributions. Tong (1990) (45) mentioned that the method of moments is one of the common approaches to estimate parametric non-linear time series models. Gerber and Shiu (1994)(31) employed the method of moments approach to estimate the unknown parameter in the shifted gamma distribution underlying their option pricing model. Taylor (1986)(44) pioneered the use of the method of moments approach for estimating stochastic volatility models. Here, given the observed market data of the logarithmic returns $\{Y_1, Y_2, \ldots, Y_N\}$ and the values of the estimated parameters $\hat{p}_k$, $\hat{\alpha}_{k0}$, $\hat{\alpha}_{k1}$, $\hat{\beta}_{k11}$ and $\hat{\beta}_{k21}$ $(k = 1, 2)$, we can employ the three realized time series $\{\xi_1, \xi_2, \ldots, \xi_N\}$, $\{h_{11}, h_{12}, \ldots, h_{1N}\}$ and $\{h_{21}, h_{22}, \ldots, h_{2N}\}$ to evaluate the method of moments estimators for the shifted gamma parameters, namely $\hat{a}_1$ and $\hat{a}_2$. By using the method of moments approach for the mixture GARCH model and matching the theoretical and the empirical third moments of the shifed-gamma innovations, we provide the following formula for the method of moments estimators $\hat{a}_1$ and $\hat{a}_2$:

$$\hat{a}_1 = \frac{1}{\left(\mathbb{B}+\sqrt{\frac{1-p_1}{p_1}\left(\mathbb{A}-\mathbb{B}^2\right)}\right)^2} \quad \textit{and } \hat{a}_2 = \frac{1}{\left(\mathbb{B}+\sqrt{\frac{p_1}{1-p_1}\left(\mathbb{A}-\mathbb{B}^2\right)}\right)^2}$$

where

$$\mathbb{B} = \frac{1}{2}\frac{\sum_{t=1}^N \xi_t^3}{\sum_{t=1}^N h_t^{3/2}} \quad \textit{and } \mathbb{A} = \frac{1}{6}\frac{\sum_{t=1}^N \xi_t^4}{\sum_{t=1}^N h_t^2} - \frac{1}{2}$$

and

$$\begin{aligned}
\xi_t^3 &= p_1\left[Y_t - \left(\lambda\sqrt{h_{1,t}} - \frac{1}{2}h_{1,t}\right)\right]^3 + (1-p_1)\left[Y_t - \left(\lambda\sqrt{h_{2,t}} - \frac{1}{2}h_{2,t}\right)\right]^3 \\
h_t^{3/2} &= p_1 h_{1,t}^{3/2} + (1-p_1)\, h_{2,t}^{3/2} \\
\xi_t^4 &= p_1\left[Y_t - \left(\lambda\sqrt{h_{1,t}} - \frac{1}{2}h_{1,t}\right)\right]^4 + (1-p_1)\left[Y_t - \left(\lambda\sqrt{h_{2,t}} - \frac{1}{2}h_{2,t}\right)\right]^4 \\
h_t^2 &= p_1 h_{1,t}^2 + (1-p_1)\, h_{2,t}^2
\end{aligned}$$

The first table below, displays the estimation results for the mixture GARCH parameters p_k, α_{k0}, α_{k1}, β_{k11} and β_{k21} $(k = 1, 2)$ using the QMLE in the first-stage of the estimation procedure. Both the estimation results based on the 25-year S&P 500 index data are presented. The estimation results were obtained using R package "rgenoud" to search for an optimum of the log-likelihood.

We consider five pricing schemes for options with 90 days to maturity: the classical Black and Scholes formulae assuming a GBM process and the discrete-time doubly stochastic Esscher transform method for GARCH and Mixture of GARCHs (henceforth MGARCH) processes, each one with Normal and shifted-Gamma innovations(henceforth NMGARCH and SGMGARCH respectively). In the Normal and Gamma codes we simulated 10000 times but as in the Normal case we used the antithetic method we actually have 20000 prices. In both cases the control variate technique was used.

The five pricing schemes are applied to two artificial series produced by a Mixture of GARCH model with Normal and Shifted-Gamma innovations with 3200 data points, obtained after a warm-up period of 1000 observations. The MGARCH parameters are given in the table 6.1.

Mixture of GARCHs Parameters	
α	$[1.275531 \times 10^{-4}, 0.433127; 1.965270 \times 10^{-9}, 4.315862 \times 10^{-2}]$
β	$[0.383459, 0.182873; 3.035471 \times 10^{-3}, 0.935921]$
λ	7.715479×10^{-2}
p	$[0.068147, 0.9319]$
a (Gamma case)	$[0.0964, 6.0256]$

Table 6.1: Parameters for the Mixture of GARCHs. Each line in the matrices α, β contains the parameters of a regime.

In order to find the Black and Scholes price we only need the volatility estimated by the sample variance, 2.0186×10^{-4}, for the Normal data and 6.0211×10^{-5} for the Gamma data. because the drift is not required by the Black Scholes formulae.

For the estimation of the GARCH parameters we use an iterated two-stage method. Initially, we suppose h_t a constant equal to the sample variance. Then, we estimate the risk premium by weighted least squares(WLS). Next, we fit a GARCH(1,1) model to the residuals of the WLS by performing a Quasi Maximum Likelihhod. We iterate these two steps until convergence is attained. The estimated parameters are shown in table 6.2.

GARCH Parameters in Normal Case	
α	5.1325×10^{-6}
β	0.9114
λ	0.0673
Risk Premium	0.1343

Table 6.2: Estimated GARCH Parameters in the Normal Case

We estimated the parameters using the two stage procedure as described before and then for finding a we used the method of moments as in Siu et al(2004)(43) to obtain the expression:

$$\hat{a} = \left[\frac{2\sum_{t=1}^{T} h_t^{3/2}}{\sum_{t=1}^{T} \xi_t^3} \right]^2 \tag{6-38}$$

which led us to the parameters shown in table 6.3.

GARCH Parameters in Shifted-Gamma Case	
α	3.7707×10^{-6}
β	0.8631
λ	0.0702
Risk Premium	0.0798
a	22.5253

Table 6.3: Estimated GARCH Parameters in the Gamma Case

Some authors (Gerber and Shiu(1994) (31) and Siu et al. (2004) (43)) have been using shifted-gamma innovations to model log-returns in order to handle the skewness that real financial series usually exhibits as can be seen in Medeiros and Veiga (2009)(36). However, the skewness of the Gamma distribution is strictly positive whilst financial time series can present both signs. In practice, before adopting the shifted-gamma model, one may check if there is any skewness to be modeled. The sign of the asymmetry in the data, has to be taken into consideration too. Then, one should check for the sign of the skewness so as to select an appropriate formulation of the shifted-gamma innovations. In Section 4, we develop a model to incorporate negative skewness. The positive case is similar, and for the GARCH case it has already been documented in Siu et al.(2004) (43) but accounts for positive skewness only. Here, in the Shifted-Gamma case, we perform an experiment consisting of using a mixture of positive and negative noises. The fisrt regime, will provide positive innovations whilst the second negative ones.

The resulting prices are presented in tables that follows. We show both Call and Put option prices. After each option price table there is another table with the ratio between the price and the the Black Scholes price for each model. A graph of these comparative tables is also shown.

Artificial MGARCH Call Prices (Normal data)

K/S_0	BS	NMGARCH	GARCH-Normal	SGMGARCH	GARCH-Gamma
0.80	20.2451	20.0396	20.3704	20.6131	19.9422
0.85	15.6879	15.1600	15.8647	16.3017	15.0241
0.90	11.5852	10.5621	11.7831	12.4678	10.3457
0.95	8.1130	6.6096	8.3018	9.2143	6.3225
1.00	5.3731	3.6432	5.5584	6.6011	3.3357
1.05	3.3638	1.7641	3.5860	4.6039	1.5130
1.10	1.9930	0.7591	2.2336	3.1507	0.5988
1.15	1.1203	0.3048	1.3582	2.1351	0.2060
1.20	0.5994	0.1208	0.8043	1.4586	0.0720

Table 6.4: Artificial MGARCH Call Prices, and T=90. The parameters used are in table 1

Table 6.5: Artificial MGARCH Call Price ratios (Normal data)

K/S_0	BS	NMGARCH	GARCH-Normal	SGMGARCH	GARCH-Gamma
0.8	1.0000	0.9898	1.0062	1.0182	0.9850
0.85	1.0000	0.9663	1.0113	1.0391	0.9577
0.9	1.0000	0.9117	1.0171	1.0762	0.8930
0.95	1.0000	0.8147	1.0233	1.1357	0.7793
1	1.0000	0.6780	1.0345	1.2285	0.6208
1.05	1.0000	0.5244	1.0661	1.3687	0.4498
1.1	1.0000	0.3809	1.1207	1.5809	0.3005
1.15	1.0000	0.2721	1.2124	1.9058	0.1839
1.2	1.0000	0.2015	1.3418	2.4334	0.1201

Call Option ratio (Normal data)

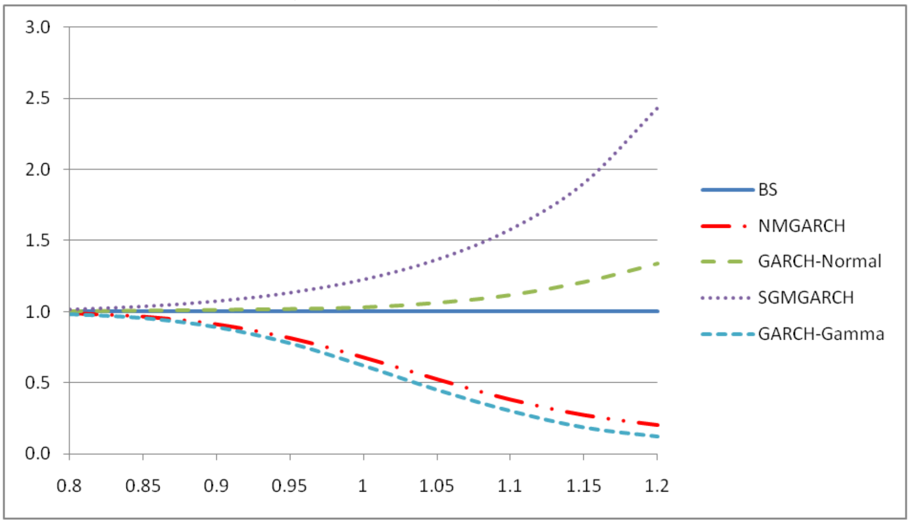

Figure 6.1: Graph of the ratio of Call options

Artificial MGARCH Call Prices (Gamma data)

K/S_0	BS	SGMGARCH	GARCH-Gamma	NMGARCH	GARCH-Normal
0.80	20.0023	20.6131	20.0617	20.0396	19.9935
0.85	15.0324	16.3017	15.0553	15.1600	15.0185
0.90	10.2388	12.4678	10.0617	10.5621	10.1955
0.95	6.0304	9.2143	5.3346	6.6096	5.9153
1.00	2.9361	6.6011	1.8991	3.6432	2.7815
1.05	1.1469	4.6039	0.4111	1.7641	1.0286
1.10	0.3553	3.1507	0.0464	0.7591	0.3053
1.15	0.0875	2.1351	0.0042	0.3048	0.0724
1.20	0.0174	1.4586	0.0010	0.1208	0.0157

Table 6.6: Artificial MGARCH Call Prices, and T=90. The parameters used are in table 1

Table 6.7: Artificial MGARCH Call Price ratios (Gamma data)

K/S_0	BS	SGMGARCH	GARCH-Gamma	NMGARCH	GARCH-Normal
0.8	1.0000	1.0305	1.0030	1.0019	0.9996
0.85	1.0000	1.0844	1.0015	1.0085	0.9991
0.9	1.0000	1.2177	0.9827	1.0316	0.9958
0.95	1.0000	1.5280	0.8846	1.0960	0.9809
1	1.0000	2.2483	0.6468	1.2408	0.9473
1.05	1.0000	4.0142	0.3584	1.5381	0.8969
1.1	1.0000	8.8677	0.1306	2.1365	0.8593
1.15	1.0000	24.4011	0.0480	3.4834	0.8274
1.2	1.0000	83.8276	0.0575	6.9425	0.9023

Call Option ratio (Gamma data)

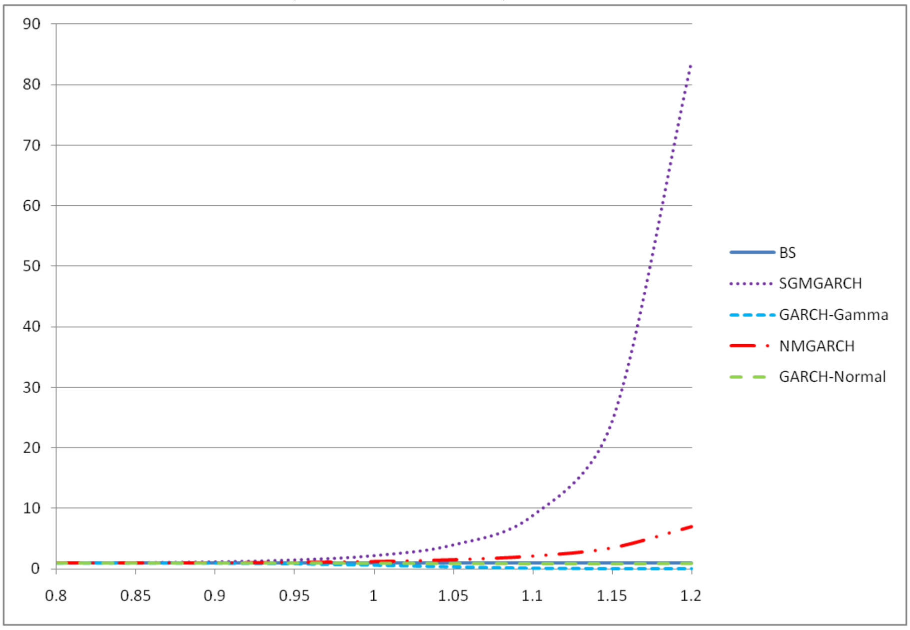

Figure 6.2: Graph of the ratio of Call options

Artificial MGARCH Put Prices (Normal data)

K/S_0	BS	NMGARCH	GARCH-Normal	SGMGARCH	GARCH-Gamma
0.80	0.2451	0.0487	0.3486	0.4922	0.0259
0.85	0.6879	0.1752	0.8335	1.0361	0.0844
0.90	1.5852	0.5728	1.7654	2.0291	0.3973
0.95	3.1130	1.5923	3.3021	3.6551	1.3615
1.00	5.3731	3.6200	5.5684	5.9864	3.3947
1.05	8.3638	6.7476	8.5754	8.9910	6.6014
1.10	11.9930	10.7743	12.2063	12.5913	10.7275
1.15	16.1203	15.3253	16.3190	16.6484	15.3697
1.20	20.5994	20.1390	20.7742	21.0324	20.2601

Table 6.8: Artificial MGARCH Put Prices, and T=90. The parameters used are in table 1

Table 6.9: Artificial MGARCH Put Price ratios (Normal data)

K/S_0	BS	NMGARCH	GARCH-Normal	SGMGARCH	GARCH-Gamma
0.8	1.0000	0.1987	1.4223	2.0082	0.1057
0.85	1.0000	0.2547	1.2117	1.5062	0.1227
0.9	1.0000	0.3613	1.1137	1.2800	0.2506
0.95	1.0000	0.5115	1.0607	1.1741	0.4374
1	1.0000	0.6737	1.0363	1.1141	0.6318
1.05	1.0000	0.8068	1.0253	1.0750	0.7893
1.1	1.0000	0.8984	1.0178	1.0499	0.8945
1.15	1.0000	0.9507	1.0123	1.0328	0.9534
1.2	1.0000	0.9776	1.0085	1.0210	0.9835

Put Option ratio (Normal data)

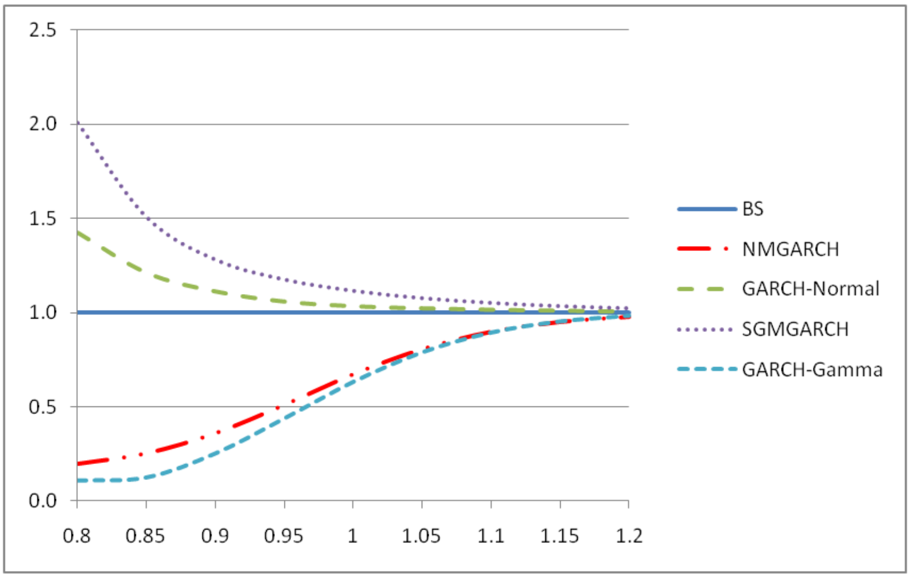

Figure 6.3: Graph of the ratio of Put Options

Artificial MGARCH Put Prices (Gamma data)

K/S_0	BS	SGMGARCH	GARCH-Gamma	NMGARCH	GARCH-Normal
0.80	0.0023	0.4922	0.0097	0.0487	0.0030
0.85	0.0324	1.0361	0.0122	0.1752	0.0305
0.90	0.2388	2.0291	0.0489	0.5728	0.2198
0.95	1.0304	3.6551	0.4256	1.5923	0.9632
1.00	2.9361	5.9864	2.0676	3.6200	2.8462
1.05	6.1469	8.9910	5.5272	6.7476	6.0732
1.10	10.3553	12.5913	10.1236	10.7743	10.3267
1.15	15.0875	16.6484	15.0670	15.3253	15.0818
1.20	20.0174	21.0324	20.0585	20.1390	20.0158

Table 6.10: Artificial MGARCH Put Prices, and T=90. The parameters used are in table 1

Table 6.11: Artificial MGARCH Put Price ratios (Gamma data)

K/S_0	BS	SGMGARCH	GARCH-Gamma	NMGARCH	GARCH-Normal
0.8	1.0000	214.0000	4.2174	21.1739	1.3043
0.85	1.0000	31.9784	0.3765	5.4074	0.9414
0.9	1.0000	8.4971	0.2048	2.3987	0.9204
0.95	1.0000	3.5473	0.4130	1.5453	0.9348
1	1.0000	2.0389	0.7042	1.2329	0.9694
1.05	1.0000	1.4627	0.8992	1.0977	0.9880
1.1	1.0000	1.2159	0.9776	1.0405	0.9972
1.15	1.0000	1.1035	0.9986	1.0158	0.9996
1.2	1.0000	1.0507	1.0021	1.0061	0.9999

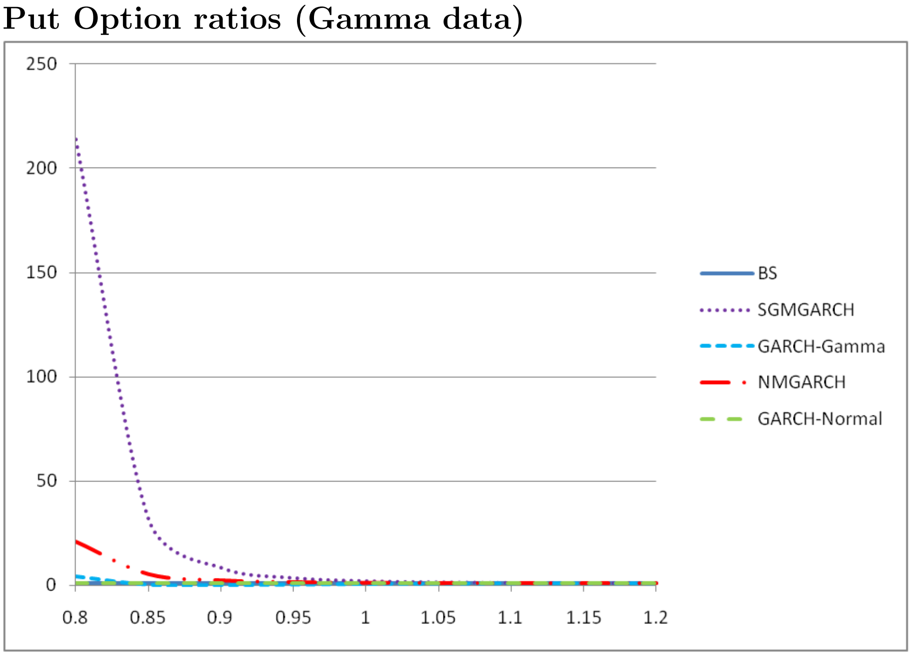

Figure 6.4: Graph of the ratio of Put options

We can see that Calls and Puts have a different behavior. In tables 6.9 and 6.11, the put option price ratios have their largest values deep in the money. The more pronounced effect is in the SGMGARCH scheme. The values are larger than those in other schemes. Tables 6.5 and 6.7, on the other hand, show their largest values deep out the money. The more pronounced effect is again in the SGMGARCH scheme. The values of the SGMGARCH scheme are much more drastic when the data is generated by Shifted-Gamma noises than with Normal noises as we can see in the figures. The second largest values alternates between the NMGARCH scheme for Shift-Gamma data and the GARCH-Normal for Normal data. The smallest values appears in the GARCH-Gamma scheme.

To illustrate the changes in the option prices when we change the measures, we simulated prices under both the physical and risk neutral measures in all schemes. Then we checked for the proportions of scenarios where the options were exercised, which then give an estimate of the real-world probability of exercising an option. We chose $S_0 = 100$ and $K = 100$ to perform this exercise. We notice that in all schemes presented in table 6.12, the prices under the risk neutral measure are less likely to exceed the strike price than the prices under the physical measure. Note also that the difference is more pronounced than in the FC-GARCH schemes in chapter 5.

Table 6.12: Average rate of exercising

$Model/rate$	Risk Neutral Measure	Physical Measure
NMGARCH	0.4824	0.7720
SGMGARCH	0.4665	0.7620
Gamma GARCH	0.4832	0.7652
Normal Garch	0.4692	0.8872
GARCH-Gamma with Normal data	0.4827	0.8813
GARCH-Normal with Gamma data	0.4856	0.7639

6.4 Sensitivity Analysis

Now we are going to check how the option prices change when some of the parameters are disturbed. We performed simulations imposing a small variation around the values of the parameters.

Proceeding with this exercise we capture the importance of each parameter in the option prices. We perform this analysis with the MGARCH models having Normal and Shifted-Gamma innovations.

6.4.1 Normal innovations

If we increase the value of α_{12}, by steps of 0.05 the option price does not change significantly. We made 5 variations in steps of 0.01 for α_{22} and as we increase its value, the option value also increases.

We made β_{11} with increments of 0.05, and for β_{21} in steps of 0.001 and the option price slightly increased. For β_{12} the effect is not clear. On the other hand for the β_{22} with 0.05 increments, we can clearly see the increase effect on option prices.

The p parameter increases the value of the option as is shown in figure 11. We performed 5 varations from 0.0081 to 0.1281 in steps of 0.03.

The risk premium didn't show any clear effect. The graphics shows five variations of the risk premium with steps of 0.05 and 0.1. It seems having an increasing trend, but this increase is very slightly even performing steps of 0.1.

6.4.2 Shifted Gamma innovations

If we increase the value of α_{12} and α_{22} , by steps of 0.05 and 0.01 respectively the option price does not have a monotonic pattern.

We made for β_{11} increments of 0.05, and the option price does not have a monotonic pattern. For β_{12}, β_{21} and β_{22} the effect is not clear too.

We performed 5 varations of the p parameter from 0.0081 to 0.1281 in steps of 0.03. No clear effect was noticed.

The a_1 parameter doesn't have a clear effect on the option prices. a_1 was varied swith increments of 0.01 and 0.03. In any of the cases there was a monotonic pattern. a_2 started from its value having increments of 0.03 and also varied but without a monotonic clear effect.

The risk premium has no clear monotonic pattern influence on the option prices. We simulated with increments of 0.03 and also steps of 0.1, In summary, in the Gamma case, maybe due to the mixture of signs in the noises, even when the prices varied significantly, no monotonic pattern could be detected.

6.5 Summary and discussion

In this paper we adopted the discrete-time doubly stochastic Esscher transform to find a pricing kernel for the MGARCH models with two different parametric distributions for innovations, the Normal and the shifted-Gamma cases. We also performed simulations and showed tables comparing the Black Scholes prices and the GARCH prices to our simulation results of the MGARCH models as well as we performed a sensitivity analysis to understand how changes in some parameters affect the option valuation results.

In the tables and graphs of section 5, we noticed large values for the SGMGARCH scheme and small values for the GARCH schemes. In the sensitivity analysis, we noticed that the GARCH parameters β_{22} and p were the most sensitive to perturbations of these model parameters in the Normal case, while the other parameters and the risk premium have little or no impact on option prices. In the Gamma case, maybe due to the mixture of signs in the noises, we could not measure the importance of any of the parameters. Even when the prices varied significantly, no monotonic pattern could be detected.

The MGARCH models can capture features that some other models cannot like the high kurtosis, so the option prices are more precise if calculated in the way we did in this paper. Here we performed simulations with a mixture of 2 regimes but the model can mimic an economy with many regimes.