International Journal of Theoretical and Applied Finance
Vol. 14, No. 5 (2011) 669–708

DOI: 10.1142/S0219024911006401

A COMPARISON OF PRICING KERNELS FOR GARCH OPTION PRICING WITH GENERALIZED HYPERBOLIC DISTRIBUTIONS

ALEXANDRU BADESCU*

Department of Mathematics and Statistics
University of Calgary, Calgary, Canada
abadescu@math.ucalgary.ca

ROBERT J. ELLIOTT

Department of Mathematical Sciences
University of Adelaide, Adelaide, Australia
and
Haskayne School of Business, University of Calgary
Calgary, Alberta, Canada
relliott@ucalgary.ca

REG KULPERGER

Department of Statistics and Actuarial Science
University of Western Ontario
London, Ontario, Canada
kulperger@stats.uwo.ca

JARKKO MIETTINEN

Department of Mathematics and Statistics
University of Helsinki, Helsinki, Finland
jjmietti@mappi.helsinki.fi

TAK KUEN SIU

Department of Applied Finance and Actuarial Studies
Faculty of Business and Economics
Macquarie University, Sydney, Australia
Ken.Siu@efs.mq.edu.au

Received 29 January 2010
Accepted 7 February 2011

Under discrete-time GARCH models markets are incomplete so there is more than one price kernel for valuing contingent claims. This motivates the quest for selecting an appropriate price kernel. Different methods have been proposed for the choice of a price kernel. Some of them can be justified by economic equilibrium arguments. This paper stud-

*Corresponding author.

ies risk-neutral dynamics of various classes of Generalized Hyperbolic GARCH models arising from different price kernels. We discuss the properties of these dynamics and show that for some special cases, some pricing kernels considered here lead to similar risk neutral GARCH dynamics. Real data examples for pricing European options on the S&P 500 index emphasize the importance of the choice of a price kernel.

Keywords: Option pricing; risk neutral valuation; Generalized Hyperbolic GARCH; extended Girsanov principle; Esscher transform; mean correcting martingale measure; Radon-Nikodym derivative.

1. Introduction

Empirical studies on asset price dynamics have shown evidence against the constant volatility assumption in the Black and Scholes [7] and Merton [42] option pricing model. This motivates a considerable amount of literature on developing asset price models which can accommodate the time variation of volatility. Notable examples are the autoregressive conditional heteroskedastic (ARCH) model of Engle [25], the generalized autoregressive conditional heteroskedastic (GARCH) model of Bollerslev [8] and Taylor [55] and the stochastic volatility (SV) model of Taylor [54], whose foundation was laid down by Clark [17] and Tauchen and Pitts [53].[1] Besides time variation of volatility, these models can incorporate other stylised features of returns, such as the heavy-tailedness of the return's distribution, volatility clustering, low autocorrelation of returns and positive autocorrelation of squared returns.

In the past two decades, option valuation under ARCH-type models has become an important topic in both option pricing theory and financial econometrics. Duan [19] developed an option pricing method under GARCH models with normal innovations. Heston and Nandi [35] derived a semi-closed-form formula for the price of a standard European option under a specific form of GARCH model. The general finding was quite promising since these models clearly outperform the homoskedastic models in explaining option price data. However, the approaches regarding the types of models and the pricing procedures used in these studies were quite different. For example, Duan considered a non-affine type of GARCH-in-mean model and used Monte-Carlo techniques for pricing European type options based on parameters estimated using only historical returns, while Heston and Nandi derive a semi-closed analytic formula and computed option prices using both information about return data and observed market quotes. Various empirical comparisons between affine and non-affine Gaussian GARCH models were done by Hsieh and Ritchken [37], and more recently, by Christoffersen *et al.* [15] for a more general volatility component model.

Another important issue addressed in the GARCH option pricing literature is the impact of different volatility specifications on option pricing. For example, Hardle and Hafner [33] proposed a model based on the Glosten *et al.* [29] asymmetric volatility process, namely, the GJR model, while Christoffersen and Jacobs [13]

[1]Taylor [56, 57] pointed out that a precursor of both the ARCH model and the SV model can be found in an unpublished manuscript by Rosenberg [45].

argued that a simple leverage effect in the conditional variance process outperforms most of the extensions considered in the literature relative to option prices. Multi component volatility models were also considered by Christoffersen *et al.* [15] by allowing a short-run and a long-run volatility dynamics to explain the variability exhibited in option prices data.

A second direction in the GARCH option pricing literature is to consider different distributions for GARCH innovations, since it was generally observed that the skewness and leptokurtosis of financial data cannot be captured by a GARCH model with normally distributed innovations (see, for example, Bollerslev [9], Baillie and Bollerslev [4], Hsieh [36]). Various parametric distributions were implemented in a GARCH framework for pricing European style options: Shifted Gamma (Siu *et al.* [50]), Inverse Gaussian (Christoffersen *et al.* [14]), Generalized Error (Duan [20], Christoffersen *et al.* [15]), α-stable (Menn and Rachev [41]), mixture of normals (Badescu *et al.* [3]), and Poisson-normal innovations (Duan *et al.* [21]). Extensive in and out-of sample empirical analyses generally show a significant improvement over Gaussian driven GARCH models relative to almost all maturities and moneyness considered. More recently, the normal variance-mean mixture distributions have been proposed as a general class of distributions for GARCH innovations. For example, the Normal Inverse Gaussian distribution (Jensen and Lunde [38]), Generalized Hyperbolic Skew Students t-Distribution (Aas and Haff, [1]), the z-distribution (Lanne and Saikkonen [39]) have been employed to incorporate some important empirical features of the returns' distribution. Much of the literature seems focussed on the use of the normal variance-mean mixture distributions for fitting financial returns data and financial risk management. A relatively small amount of work has been done studying the empirical performance of such models in option valuation. Some of these few examples include Stentoft [52] and Chorro *et al.* [12]. Stentoft [52] used the Normal Inverse Gaussian density for pricing American options, while Chorro *et al.* [12] investigated the pricing performance of GARCH models based on the Generalized Hyperbolic distribution. Since these studies used a similar model to the one in [12] we shall pay attention to the main differences between the two approaches.

Non-parametric techniques were also investigated for pricing options within the GARCH setup. For example, Barone-Adesi *et al.* [6] proposed an option pricing algorithm by calibrating an asymmetric GARCH model to observed market quotes using the empirical density function of the filtered historical innovations, while Badescu and Kulperger [2] investigated a semi-parametric pricing method based on standardized residuals and a kernel density estimator of the innovations distribution.

Return dynamics and volatility specifications are not the only issues one should pay attention to when pricing contingent claims. The valuation problem in the incomplete markets included in these discrete time models has to be investigated carefully as well. From a theoretical perspective, the choice of an appropriate price kernel should be justified by some economic arguments. From a practical viewpoint, the choice of a price kernel may be dictated by analytical tractability, or mathematical convenience. The traditional method for derivative pricing in the

GARCH setup is based on the Risk Neutral Valuation Relationship (RNVR) introduced by Rubinstein [46] and Brennan [10] for discrete time models and normally distributed asset returns. The crucial assumption for constructing this price kernel is the bivariate normal distribution of the returns and the logarithm of the stochastic discount factor (SDF), while an economic justification is based on a Lucas-type equilibrium pioneered by Lucas [40]. Duan [19] analyzed the performance of a Gaussian GARCH option pricing model by introducing a local version of the RNVR (LRNVR), while Heston and Nandi [35] used a similar assumption for developing semi-closed-form pricing formulas for European options for a GARCH model with normally distributed innovations. Since the above method cannot be directly applied when relaxing the conditional normality assumption of the asset returns, researchers try to exploit other possible choices for price kernels which might be consistent with some economic equilibrium settings. For example, Duan [20] introduced a generalization of the local risk neutral valuation relationship, namely, the generalized LRNVR (GLRNVR) for a GARCH model with generalized error distribution, namely, the GED-GARCH model. His construction was further investigated by Stentoft [52], Christoffersen *et al.* [15]. Badescu *et al.* [3] used a similar assumption as Garcia *et al.* [28] to build an equilibrium measure for pricing under mixture normal GARCH models. Another martingale measure which can accommodate for non-normality in both discrete and continuous time settings is the Esscher transform introduced in the option pricing literature by Gerber and Shiu [30]. The conditional version of this transformation, see Bühlmann *et al.* [11], was first applied in a GARCH setting by Siu *et al.* [50] with a numerical illustration for shifted Gamma innovations. This principle was used for valuation in Badescu and Kulperger [2] and Christoffersen *et al.* [16] amongst others, although in the latter paper the measure change was also derived using an extended Girsanov-type argument.

Two main issues are addressed in this paper. Firstly, we consider a Generalized Hyperbolic GARCH option pricing model which nests many of the existing models in the literature. Secondly, we study the relationship among various choices of price kernels for different GARCH settings and investigate empirically the sensitivity of option prices with respect to the choice of the price kernels.

In the first part of the paper we briefly describe three price kernels within a general GARCH framework by specifying the parametric forms of the stochastic discount factors: a mean correcting martingale measure (MCMM), a conditional Esscher transform, and the GLRNVR. The MCMM approach is based on the extended Girsanov transformation that states the risk neutral dynamics of asset returns are obtained by shifting the mean of the density under the physical measure while the variance remains unchanged. We show that this approach is equivalent to a discrete-time version of the extended Girsanov transformation of Elliott and Madan [24] for a general, discrete-time, asset price model. The MCMM approach was later applied for option valuation in the context of a continuous-time, Geometric Lévy process, (see, for example, Schoutens [48]). An interesting feature of this transformation is that it preserves the parametric form of the return distribution after the measure change, and, in some special cases, it can be linked to equilibrium measures as done in Schroder [49].

In the second part of this paper we consider the GH-GARCH model. The family of Generalized Hyperbolic distributions, which is a subclass of the more general normal variance-mean mixture distributions, has become a standard class of returns' distributions in the literature due to its capability to provide a good fit to financial returns data.

Based on the price kernels mentioned above we derive risk neutralized dynamics for the GH-GARCH model which will be used for evaluating option prices. Although the MCMM and the conditional Esscher transform preserve the same parametric form of the probability law for returns after the measure change in our setting, they lead to different return dynamics, and therefore, to different option prices. It is interesting to note that in the Gaussian limit case all of the three price kernels considered lead to the same risk neutralized return dynamics. Following the approach of Brennan [10] we are able to show that the conditional Esscher transform used in the option valuation under the GH-GARCH model is consistent with that arising from an equilibrium pricing measure by making use of a similar assumption regarding the bivariate distribution of the asset returns and the logarithm of a price kernel. To provide empirical support and to illustrate the implementation of the results developed here, we conduct a real data experiment by computing European call option prices written on S&P 500 Index for various subclasses and limits of Generalized Hyperbolic distributions under all of the three pricing kernels.

We use Monte Carlo simulations for approximating the option prices. Depending on the tractability, asset return paths can be simulated either under a risk neutral measure, or under the physical measure, where in the latter, the option prices are computed based on simulated paths of the Radon-Nikodym derivatives.

The results of our empirical analysis reveal that all models with conditional skewness outperform considerably the GARCH one with Gaussian innovations under all of the three price kernels. This is a natural consequence of the enhanced overall fit offered by these non-Gaussian models. Another important conclusion is that, except for the Gaussian GARCH case, option prices are sensitive to the choices of various price kernels. This provides empirical support to our conjectures and theoretical results. Based on three standard performance indicators, we conclude that the best "overall" model is the Normal Inverse Gaussian GARCH in terms of fitting market returns data, while the conditional Esscher transform gives rise to option prices closest to observed market option prices data with corresponding strikes and maturities.

An investigation for moneyness and maturity indicates that the conditional Esscher transform is, in general, the most suitable one for valuing long-term, out-of-money options. The GLRNVR (except for the NIG case) and the MCMM are appropriate for valuing short-term options. This has important implications for both actuarial and financial pricing since the Esscher transform has long been used in valuing insurance contracts in the actuarial science literature and the terms of insurance contracts and products are relatively longer than those of financial products. Therefore, some valuation tools which are suitable for valuing financial products might not be appropriate for valuing long-term financial and insurance products. Indeed, many modern insurance products with embedded options, such as unit-linked life policies, usually have long maturities and guarantees underlying some of these policies would be likely to be of value. The empirical prominency of the conditional Esscher transform for valuing long-term, out-of-money options makes it a tempting choice of valuing the embedded options underlying these modern insurance products.

As previously mentioned, Chorro *et al.* [12] considered GARCH models with innovations having a GH distribution for option pricing. However, there are some fundamental differences between our proposed model and theirs. Firstly, we derive risk neutral dynamics for the GH-GARCH model using pricing kernels arising from three different approaches, whereas, Chorro *et al.* [12] only focused on the pricing method based on the conditional Esscher transform. Note that a number of GARCH option pricing models may be regarded as particular cases of the GH-GARCH models. We provide details on how to obtain these particular models as special cases in the appendix. Furthermore, we present a detailed comparison between the pricing kernels and the consumption based models. Secondly, the numerical implementation in our present paper and that in Chorro *et al.* [12] are different. They used a two-stage estimation procedure consisting of the Quasi Maximum Likelihood Estimation (QMLE) at the first stage and the MLE to estimate the unknown parameters in the GH distribution using standardized residuals at the second stage. This method is similar to that in Siu *et al.* [50]. Here we directly use the classical MLE algorithm to estimate the unknown model parameters based on historical returns. Lastly, Chorro *et al.* [12] computed the Monte Carlo option prices by simulating asset returns under a risk-neutral measure selected by the Esscher transform. We use simulated return paths under the historical measure P and make use of the Radon-Nikodym derivative determined by the Esscher transform. There are at least two advantages of this approach. Firstly, it does not require the risk-neutral return dynamics, which are not easy to find in some cases. Secondly, it is a kind of variance reduction method. In a recent paper by Siu and Yang [51], this simulation approach has been applied for option valuation under regime-switching models.

The rest of the paper is organized as follows: Sec. 2 briefly describes the three price kernels in a general GARCH setting. In Sec. 3 we introduce the GH-GARCH model and investigate the consistency of the proposed risk neutral measures and equilibrium measures. Risk neutralized return dynamics for various subclasses and limits of GH-GARCH models are derived in the appendix. Model estimation and the option pricing procedures are described in Sec. 4. The conclusions and further research are illustrated in Sec. 5.

2. Pricing Kernels for a General GARCH Setup

Consider a discrete time economy with two primitive securities, namely, a risk-free asset, say a bond, and a risky asset, say a stock. Suppose $\mathcal{T} = \{t|t = 0, 1, \ldots, T\}$,

($T < \infty$), is the set of trading dates. To model uncertainty, we fix a complete, filtered, probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$, where P is a historical probability measure[2] and $\{\mathcal{F}_t\}$ is a filtration, or a family of increasing information sets, representing the resolution of uncertainty based on information generated by market prices over time; we assume $\mathcal{F}_0 = \sigma\{\emptyset, \Omega\}$, the trivial information set, and $\mathcal{F}_T = \mathcal{F}$. We start by considering the following GARCH-in-mean model for the return $y_t := \ln(S_t/S_{t-1})$, where S_t is the stock price at time t.

$$y_t = m_t + \sigma_t \varepsilon_t \tag{2.1}$$

$$\sigma_t^2 = a_0 + a_1 \sigma_{t-1}^2 \varphi(\varepsilon_{t-1}) + b_1 \sigma_{t-1}^2. \tag{2.2}$$

Here

(1) $\{\varepsilon_t\}_{0 \le t \le T}$ is a sequence of independent and identically distributed (i.i.d.) random variables with common distribution $\mathbf{D}(0, 1)$, a distribution with zero mean and unit variance;
(2) the conditional mean return m_t it is assumed to be an $\{\mathcal{F}_t\}$-predictable process. In many studies, m_t is assumed to be a function of the conditional variance σ_t^2 of the return and a risk premium quantifier at time t;
(3) the function $\varphi(\cdot)$ describes the impact of random shock of return ϵ_{t-1} on the conditional variance σ_t^2. This is called the news impact curve, which was introduced by Pagan and Schwert [43] and Engle and Ng [26];
(4) a_0, a_1 and b_1 are the coefficients of the GARCH model, where $a_0 > 0$ and $a_1, b_1 \ge 0$, and these parameters and the function $\varphi(\cdot)$ are such that the conditional variance dynamics are covariance stationary.

Throughout this paper we assume that the conditional cumulant function of ε_t given $\mathcal{F}_{t-1}$ under P exists in a neighborhood of zero; that is,

$$\kappa_{\varepsilon_t}^P(u) = \ln E^P[e^{u\varepsilon_t} | \mathcal{F}_{t-1}] < \infty, \quad u \in (-L, L), \ \ L > 0. \tag{2.3}$$

Stochastic discount factors (SDF) pioneered by Hansen and Richard [32] and Harrison and Kreps [34] are important tools for derivative valuation. Here we consider different parametric forms of stochastic discount factors, or price kernels, for valuing European-style options. The main idea of using stochastic discount factors for derivative valuation is described below.

Let $\{M_t\}$ be a positive-valued, $\{\mathcal{F}_t\}$-adapted, stochastic process defined on $(\Omega, \mathcal{F}, P)$ such that the following no-arbitrage conditions hold:

$$E^P[M_t | \mathcal{F}_{t-1}] = e^{-r_t} \tag{2.4}$$

$$E^P[M_t e^{y_t} | \mathcal{F}_{t-1}] = 1. \tag{2.5}$$

[2]There are different names for the historical probability measure. It is also called a statistical probability measure, a data-generating probability measure, a physical probability measure and a real-world probability measure. In some literature P is also referred to as a subjective probability measure. Here we use the term "historical probability" for the measure P.

Here $\{r_t\}$ is an $\{\mathcal{F}_t\}$-predictable process, where r_t is the interest rate in the period from time $t-1$ to time t. Condition (2.4) ensures that the probability measure induced by M_t is well-defined, while (2.5) ensures that discounted asset prices are martingales under this new measure.

A price at time t of a European option with payoff $h(S_T)$ and expiration time T associated with a SDF $\{M_t\}$ is given by:

$$\Pi_t^M(h(S_T)) = E^P[M_{t+1}, \ldots, M_T h(S_T)|\mathcal{F}_t]. \tag{2.6}$$

Here E^P is expectation under P.

There are various admissible choices of SDF's that satisfy the no-arbitrage conditions and are consistent with the above valuation formula. Among them, we describe in the following subsections three potential candidates of SDF's for option valuation in the GARCH framework given by (2.1)–(2.2).

2.1. *Mean correcting martingale measure*

The mean correcting martingale measure (MCMM) is a popular tool for derivative valuation under markets driven by continuous-time, exponential Levy processes (see, for example, Schoutens, [48]). The construction of the MCMM is rather simple and it is based on a Girsanov-type transformation that preserves the returns distribution after the measure change by shifting the mean while keeping the variance unchanged. Along this line, we construct a discrete-time version of the MCMM in our proposed GARCH framework via the SDF approach. We assume that the state price process $\{M_t\}$ that obeys the no-arbitrage conditions (2.4)–(2.5) has the following form:

$$M_t = e^{-r_t}\frac{f_t^P(\varepsilon_t + \varrho_t)}{f_t^P(\varepsilon_t)} \tag{2.7}$$

where

(1) $f_t^P(\cdot)$ is the conditional probability density of the innovation ε_t at time t given $\mathcal{F}_{t-1}$ under the historical measure P;
(2) the market price of risk process, denoted as $\{\varrho_t\}$, is an $\{\mathcal{F}_t\}$-predictable process and is uniquely determined from condition (2.5) for any $t \in \mathcal{T}\backslash\{0\}$;
(3) it is clear by definition that the parametric form of the SDF from (2.7) satisfies (2.4).

In the general GARCH framework considered above, finding ϱ_t is equivalent to identifying quantity m_t^{shift}, which represents a shift to the conditional mean return m_t such that the discounted stock price is an $\{\mathcal{F}_t\}$-martingale under a newly defined probability measure $Q^{(m)}$ ($Q^{(m)}$ is the probability measure induced by the SDF from (2.7)). Consequently, under $Q^{(m)}$,

$$y_t = m_t + m_t^{\text{shift}} + \sigma_t\varepsilon_t', \quad \varepsilon_t' \sim \mathbf{D}(0,1) \text{ iid}. \tag{2.8}$$

Imposing now the martingale condition of the discounted asset price under $Q^{(m)}$ leads to the following form for the mean shift:

$$m_t^{\text{shift}} = r_t - m_t - \kappa_{\varepsilon_t}^P(\sigma_t), \tag{2.9}$$

or equivalently, the market price of risk has the following form:

$$\varrho_t = \frac{m_t + \kappa_{\varepsilon_t}^P(\sigma_t) - r_t}{\sigma_t}. \tag{2.10}$$

Therefore, we can conclude that under a price kernel arising from the mean correcting martingale measure, $Q^{(m)}$ the dynamics of the conditional return and variance are given by:

$$y_t = r_t - \kappa_{\varepsilon_t}^P(\sigma_t) + \sigma_t \varepsilon'_t \tag{2.11}$$

$$\sigma_t^2 = a_0 + a_1 \sigma_{t-1}^2 \varphi(\varepsilon'_{t-1} - \varrho_{t-1}) + b_1 \sigma_{t-1}^2. \tag{2.12}$$

An obvious advantage of using $Q^{(m)}$ for specifying a price kernel is its tractability. Indeed, the risk-neutral dynamics arising from $Q^{(m)}$ can easily be obtained for various non-normal distributions of GARCH innovations. However, economic justifications for the choice of a price kernel based on the MCMM seems to be lacking. Here we show that in our proposed GARCH setting, the MCMM construction of a price kernel is equivalent to the construction based on the extended Girsanov's transformation pioneered by Elliott and Madan [24] for a general, discrete-time, model. The construction in [24], based on an argument similar to a continuous-time Girsanov transformation, involves a multiplicative Doob's decomposition of the discounted stock price as the product of a predictable process and a martingale under the historical measure P. The main idea of the Elliott-Madan approach is to identify a new probability measure such that the probability law of the discounted stock price under the newly defined measure is identical to the probability law of its martingale component under the historical measure. To implement the Elliott-Madan method in our proposed GARCH model, we assume the following parametric form of the SDF:

$$M_t = e^{-r_t} \frac{g_t^P(y_t - r_t + \kappa_{y_t}^P(1))}{g_t^P(y_t)}, \tag{2.13}$$

where

(1) $g_t^P(\cdot)$ is the conditional probability density function of the return y_t at time t given $\mathcal{F}_{t-1}$ under P;
(2) $\kappa_{y_t}^P(\cdot)$ is the conditional cumulant generating function of y_t given $\mathcal{F}_{t-1}$ under P.

By a change of variables, it is not difficult to show that the above parametric form of the SDF coincides with the SDF specified by the MCMM in (2.7).

As in [24], the extended Girsanov's transformation is defined with respect to the natural filtration generated by asset price dynamics which, in our GARCH setting, corresponds to $\mathcal{F}_t = \sigma(\varepsilon_s : s \leq t)$. This specification of a price kernel is consistent with a weak form of market efficiency.

Elliott and Madan [24] showed that the specification of a price kernel by the extended Girsanov's transformation is justified by quadratic hedging strategies that minimize the conditional variance of the discounted, risk-adjusted, hedging cost. Here we justify the use of the extended Girsanov's transformation from a different perspective and consider the use of equilibrium models to justify the parametric specification of a price kernel by the extended Girsanov's transformation. We shall discuss this in detail in Sec. 3 when analyzing some special cases of our proposed GARCH model.

2.2. *Conditional Esscher transform*

The Esscher transform is a well-known tool in actuarial science and its use for option valuation were first proposed by the seminal work of Gerber and Shiu [30]. Using this approach, Gerber and Shiu derived European option prices for some important models for the logarithm of the stock price including a Brownian motion with constant drift, a Binomial distribution, a shifted Gamma process, a shifted Poisson process and a shifted inverse Gaussian process. In the case when the logarithm of the stock price follows a Brownian motion with drift, the Gerber-Shiu option valuation formula coincides with the celebrated Black-Scholes-Merton option pricing formula.

Since the GARCH setup does not satisfy the assumption of independent increments imposed in the Gerber-Shiu option pricing model, one may need to modify the Gerber-Shiu option pricing model when valuing options in the GARCH setup. In particular, a conditional version of the Esscher transform, introduced by Bühlmann *et al.* [11], provides a promising way to accommodate the dependent increments arising from the GARCH structure. The key idea of the conditional Esscher transform is to consider the following exponential-affine function for a price kernel:

$$M_t = e^{\theta_t^* y_t - \kappa_{y_t}^P(\theta_t^*) - r_t}, \tag{2.14}$$

where $\{\theta_t^*\}$ is the unique predictable process that satisfies the no-arbitrage condition (2.5).

More specifically, for each $t \in \mathcal{T}$, we need to solve for θ_t^* from:

$$\kappa_{y_t}^P(1 + \theta_t^*) - \kappa_{y_t}^P(\theta_t^*) = r_t. \tag{2.15}$$

We call (2.15) a martingale equation. The existence of a solution of (2.15) is guaranteed by the existence of the cumulant generating function imposed at the beginning of this section. The Esscher pricing kernel in (2.14) resembles the one adopted in Gourieroux and Monfort [31]. In general, one may need to impose some restrictions on the model parameters to ensure the existence of the cumulant generating function. When we consider some parametric cases of the Generalized Hyperbolic innovations, some constraints on the parameters of the distribution of the innovations are required. A discussion for the uniqueness of a solution of a similar martingale solution can be found in [30] for the independent increments setup and [16] for the GARCH setup. Note that the uniqueness of a conditional Esscher transformation for the process y_t such that the martingale equation (2.15) is satisfied does not mean the uniqueness of a measure change for risk-neutralization.

We denote by $Q^{(e)}$ the risk neutral probability measure induced by the SDF from (2.14). Following Siu *et al.* [50], the cumulant generating function of the returns under $Q^{(e)}$ is given by:

$$\kappa_{y_t}^{Q^{(e)}}(u) = \kappa_{y_t}^{P}(u + \theta_t^*) - \kappa_{y_t}^{P}(\theta_t^*), \quad u \in (-L, L), \ \ L > 0. \tag{2.16}$$

The conditional Esscher transform is a convenient tool for derivative valuation when the distributions of asset returns are non-normal. A justification for the choice of an exponential-affine price kernel is provided by Gerber and Shiu [30]. They showed that the choice is the same as the one arising from the solution of a utility maximization problem of an economic agent with a power utility.

Since the consistency between the choice of SDF's and consumption based models is essential in both the empirical asset pricing and financial econometric literatures, we show in Sec. 3 how a price kernel specified by the conditional Esscher transform can be derived from an equilibrium condition under the GH-GARCH model.

2.3. *Generalized local risk neutral valuation relationship*

Duan [20] extended the LRNVR to the GLRNVR to deal with option valuation under GARCH models with non-normal innovations. The key idea of the GLRNVR is the introduction of a normal transformation which transforms a non-normal innovation term into a normal one. This is similar to the construction of the Wang transform introduced by Wang ([58, 59]) to the actuarial science literature. Here a new equilibrium price kernel is constructed via a conditional normality transformation based on the idea of the GLRNVR. The key assumption here is also a conditional bivariate normal distribution of the logarithm of the stochastic discount factor and the transformed normal random variable $\Phi^{-1}(F(\varepsilon_t))$, where $\Phi(\cdot)$ is the standard normal distribution function and F is the innovation's distribution function under the physical measure P.

The return dynamics under the GLRNVR measure, denoted as $Q^{(g)}$, are governed by:

$$y_t = m_t + \sigma_t \eta_t, \tag{2.17}$$

$$\eta_t = F^{-1}(\Phi(\xi_t - \upsilon_t)), \tag{2.18}$$

$$\sigma_t^2 = a_0 + a_1 \sigma_{t-1}^2 \varphi(\eta_{t-1}) + b_1 \sigma_{t-1}^2 \tag{2.19}$$

where

(1) $\{\xi_t\}$ are i.i.d. standard normal random variables under $Q^{(g)}$;
(2) $F^{-1}(\cdot)$ is the inverse of the innovation's distribution function under the physical measure P;
(3) υ_t is the risk premium parameter at time t.

We denote by $\kappa_{\eta_t}^{Q^{(g)}}$ the conditional cumulant generating function of η_t given $\mathcal{F}_{t-1}$ under $Q^{(g)}$. As noted in Stentoft [52], the main difficulty in implementing this transformation is that solving for υ_t from the martingale equation:

$$m_t = r_t - \kappa_{\eta_t}^{Q^{(g)}}(\sigma_t), \tag{2.20}$$

can be computationally demanding. Christoffersen *et al.* [15] used a linear approximation to the Generalized Error Distribution to solve for υ_t. To simplify the implementation of this method one may assume υ_t to be constant for any $t \in \mathcal{T}$. For example, Stentoft [52] estimated this parameter using historical returns by substituting the right hand side of equation (2.20) into the return dynamics under P from (2.1). However, this approach requires calculating an approximation of $\kappa_{\eta_t}^{Q^{(g)}}(\sigma_t)$ which may slow down the computations. In our numerical study we consider a specific form of m_t which incorporates a constant market price of risk and we use its estimated value from historical returns to approximate υ_t. One disadvantage of the GLRNVR is that, except for the Gaussian innovation case, one has to solve for the unknown risk premium parameter υ_t from (2.20). Consequently, it is difficult to assess the accuracy of this method relative to the other two proposed risk neutral measures when υ_t is constant.

3. Generalized Hyperbolic GARCH Model (GH-GARCH)

In this section we consider a GARCH model based on the Generalized Hyperbolic distribution. By invoking the use of the martingale measures discussed in the previous section, we derive risk-neutral dynamics for the asset returns. Furthermore, we show that the conditional Esscher transform and Duan's LRNVR are consistent with SDF's derived based on consumption based CAPM models under some assumptions regarding the bivariate distribution of the asset returns and the marginal utility of consumption.

3.1. *Risk neutral specifications under GH-GARCH*

The Generalized Hyperbolic (GH) distribution is a semi-heavy-tailed distribution that captures many important features exhibited by the empirical distribution of financial returns data (see Barndorff-Nielsen and Shepard [5], Eberlein [22], amongst others). The GH distribution has a mixture representation. In particular, it can be constructed by assuming that the mixing random variable Y has a Generalized

Inverse Gaussian (GIG) distribution with the following probability density function:

$$f_{\mathrm{GIG}}(y,\lambda,\delta,\gamma) = \left(\frac{\gamma}{\delta}\right)^{\lambda} \frac{1}{2K_\lambda(\delta\gamma)} y^{\lambda-1} \exp\left[-\frac{1}{2}(\delta^2 y^{-1} + \gamma^2 y)\right], \tag{3.1}$$

where $y > 0$; K_λ is the modified Bessel function of the third kind[3] associated with the parameter λ; the parameters should satisfy the following conditions:

- $\delta \geq 0, |\gamma| > 0$ if $\lambda > 0$,
- $\delta > 0, |\gamma| > 0$ if $\lambda = 0$,
- $\delta > 0, |\gamma| \geq 0$ if $\lambda < 0$.

Suppose a random variable X follows the GH distribution with parameters $\lambda, \alpha, \beta, \delta, \mu$. We write

$$X \sim \mathbf{GH}(\lambda, \alpha, \beta, \delta, \mu). \tag{3.2}$$

Then the probability density function of X is given by:

$$\begin{aligned} f_{\mathrm{GH}}(x,\lambda,\alpha,\beta,\delta,\mu) = {} & \frac{\gamma^\lambda}{\sqrt{2\pi}\delta^\lambda K_\lambda(\delta\gamma)} \cdot \frac{K_{\lambda-\frac{1}{2}}(\alpha\sqrt{\delta^2+(x-\mu)^2})}{(\sqrt{\delta^2+(x-\mu)^2}/\alpha)^{\frac{1}{2}-\lambda}} \\ & \cdot \exp[\beta(x-\mu)], \quad x \in \mathbf{R}, \end{aligned} \tag{3.3}$$

where α is the kurtosis; β is the skewness; δ the scale parameter; μ is the location parameter; $\gamma^2 = \alpha^2 - \beta^2$.

Clearly, this distribution is well defined if $|\beta| \leq \alpha$. It is easy to check that if $X \sim \mathbf{GH}(\lambda, \alpha, \beta, \delta, \mu)$, then $m + \sigma X \sim \mathbf{GH}(\lambda, \alpha/|\sigma|, \beta/|\sigma|, \delta|\sigma|, \sigma\mu + m)$. However, there are various scale and location invariant parametrizations[4] of the generalized hyperbolic distribution proposed in the literature. For our numerical purposes we use two such parametrizations:

- Parametrization (1) $\tilde{\alpha} = \alpha\delta, \tilde{\beta} = \beta\delta$,
- Parametrization (2) $\xi = \delta\gamma, \rho = \beta/\alpha$.

The cumulant generating function of X takes the following form, (see Eberlein and Hammerstein [23]):

$$\begin{aligned} \kappa_{\mathrm{GH}}(u) = {} & \mu u + \frac{\lambda}{2} \ln\left(\frac{\alpha^2-\beta^2}{\alpha^2-(\beta+u)^2}\right) \\ & + \ln \frac{K_\lambda(\delta\sqrt{\alpha^2-(\beta+u)^2})}{K_\lambda(\delta\sqrt{\alpha^2-\beta^2})}, \quad |\beta+u| < \alpha. \end{aligned} \tag{3.4}$$

[3]For more information about the modified Bessel function interested readers may refer to Appendix B in Prause (1999).

[4]In general, a random variable Y belongs to a location-scale family of distributions if the cumulative distribution function of Y can be expressed as $F(y;\mu,\sigma) = \Phi(\frac{y-\mu}{\sigma})$, where $-\infty < \mu < \infty$ is a location parameter, $\sigma > 0$ is a scale parameter and Φ is the cumulative distribution function of the standardized random variable $\frac{Y-\mu}{\sigma}$.

The mean and the variance of X are, respectively, given by:

$$E[X] = \mu + \frac{\delta\beta}{\gamma} R_\lambda(\delta\gamma), \tag{3.5}$$

$$\mathrm{Var}[X] = \frac{\delta}{\gamma} R_\lambda(\delta\gamma) + \frac{\beta^2\delta^2}{\gamma^2} S_\lambda(\delta\gamma). \tag{3.6}$$

where the functions R_λ and S_λ are defined for all $u \in \mathbf{R}^+$ by:

$$R_\lambda(u) = \frac{K_{\lambda+1}(u)}{K_\lambda(u)}, \tag{3.7}$$

$$S_\lambda(u) = \frac{K_{\lambda+2}(u)K_\lambda(u) - K^2_{\lambda+1}(u)}{K^2_\lambda(u)}. \tag{3.8}$$

In the following, we implement the Generalized Hyperbolic distribution in the GARCH setup described in the previous section. We consider that under the physical measure P the return and conditional variance dynamics are, respectively, governed by:

$$y_t = m_t + \sigma_t \varepsilon_t, \tag{3.9}$$

$$\varepsilon_t \sim \mathbf{GH}(\lambda, \alpha, \beta, \delta, \mu), \tag{3.10}$$

$$\sigma_t^2 = a_0 + a_1 \sigma_{t-1}^2 \varphi(\varepsilon_{t-1}) + b_1 \sigma_{t-1}^2. \tag{3.11}$$

We assume that $\{\varepsilon_t\}$ are i.i.d. with common probability density function (3.3). For the purpose of simulation, we assume the parameter λ can either take a predetermined value or need to be estimated from the data, depending on the special case of GH distribution considered. To standardize the innovation process, we impose the model parameters $(\lambda, \alpha, \beta, \delta, \mu)$ specifying the Generalized Hyperbolic distribution to satisfy the following conditions:

$$\mu + \frac{\delta\beta}{\gamma} R_\lambda(\delta\gamma) = 0, \tag{3.12}$$

$$\frac{\delta}{\gamma} R_\lambda(\delta\gamma) + \frac{\beta^2\delta^2}{\gamma^2} S_\lambda(\delta\gamma) = 1. \tag{3.13}$$

Using (3.3), the conditional probability density function of the return y_t given $\mathcal{F}_{t-1}$ under P has the following form:

$$f^P_{y_t}(y) = \frac{(\alpha^2 - \beta^2)^{\frac{\lambda}{2}}}{\sqrt{2\pi}\sigma_t \delta^\lambda K_\lambda(\delta\sqrt{\alpha^2 - \beta^2})} \cdot \frac{K_{\lambda - \frac{1}{2}}\left(\alpha\sqrt{\delta^2 + (\frac{y - m_t}{\sigma_t} - \mu)^2}\right)}{\left(\sqrt{\delta^2 + (\frac{y - m_t}{\sigma_t} - \mu)^2}/\alpha\right)^{\frac{1}{2} - \lambda}}$$
$$\cdot \exp\left[\beta\left(\frac{y - m_t}{\sigma_t} - \mu\right)\right], \quad y \in \mathbf{R} \tag{3.14}$$

The following proposition gives the risk-neutralized dynamics of the return arising from the conditional Esscher transform. This result is similar to the one derived by Chorro *et al.* [12] in Proposition 5.

Theorem 3.1. *Suppose the asset return process* $y := \{y_t\}_{t\in\mathcal{T}}$ *satisfies* (3.9)–(3.11). *Under the risk-neutral conditional Esscher transform* $Q^{(e)}$ *the dynamics of the return have the following representation*:

$$y_t = m_t + \sigma_t(\mu + \vartheta_{1t}) + \sigma_t\vartheta_{2t}\xi_t, \tag{3.15}$$

$$\xi_t|_{\mathcal{F}_{t-1}} \sim \mathbf{GH}\left(\lambda, \alpha\vartheta_{2t}, \beta_{1t}\vartheta_{2t}, \frac{\delta}{\vartheta_{2t}}, -\frac{\vartheta_{1t}}{\vartheta_{2t}}\right), \tag{3.16}$$

where $\{\beta_{1t}\}$, $\{\vartheta_{1t}\}$ *and* $\{\vartheta_{2t}\}$ *are some* $\{\mathcal{F}_t\}$*-predictable processes given by*:

$$\beta_{1t} = \beta + \theta_t^*\sigma_t, \tag{3.17}$$

$$\vartheta_{1t} = \frac{\delta\beta_{1t}}{\sqrt{\alpha^2-\beta_{1t}^2}}R_\lambda\left(\delta\sqrt{\alpha^2-\beta_{1t}^2}\right), \tag{3.18}$$

$$\vartheta_{2t} = \left(\frac{\delta}{\sqrt{\alpha^2-\beta_{1t}^2}}R_\lambda(\delta\sqrt{\alpha^2-\beta_{1t}^2}) + \frac{\delta^2\beta_{1t}^2}{\alpha^2-\beta_{1t}^2}S_\lambda(\delta\sqrt{\alpha^2-\beta_{1t}^2})\right)^{\frac{1}{2}}, \tag{3.19}$$

such that for each $t \in \mathcal{T}\backslash\{0\}$, ξ_t *has zero conditional mean and unit conditional variance given* $\mathcal{F}_{t-1}$ *and* θ_t^* *is the unique predictable solution of the following martingale equation*:

$$\frac{\lambda}{2}\ln\left(\frac{\alpha^2-(\beta+(1+\theta_t^*)\sigma_t)^2}{\alpha^2-(\beta+\theta_t^*\sigma_t)^2}\right) + \ln\frac{K_\lambda(\delta\sqrt{\alpha^2-(\beta+\theta_t^*\sigma_t)^2})}{K_\lambda(\delta\sqrt{\alpha^2-(\beta+(1+\theta_t^*)\sigma_t)^2})}$$
$$= \mu\sigma_t + m_t - r_t. \tag{3.20}$$

Proof. First we notice from (3.9) that under P the returns are conditionally generalized hyperbolic distributed, $y_t|\mathcal{F}_{t-1} \sim \mathbf{GH}(\lambda, \alpha/\sigma_t, \beta/\sigma_t, \delta\sigma_t, m_t + \sigma_t\mu)$, and using (3.4) we can express the conditional cumulant function of the returns under the physical measure P as:

$$\kappa_{y_t}^P(u) = (m_t+\mu\sigma_t)u + \frac{\lambda}{2}\ln\left(\frac{\alpha^2-\beta^2}{\alpha^2-(\beta+u\sigma_t)^2}\right) + \ln\frac{K_\lambda(\delta\sqrt{\alpha^2-(\beta+u\sigma_t)^2})}{K_\lambda(\delta\sqrt{\alpha^2-\beta^2})},$$
$$|\beta+u\sigma_t| < \alpha.$$

Using this expression, it is easy to verify that the martingale equation (2.15) leads to (3.20), provided that $-\alpha < \beta+\theta_t^*\sigma_t < \alpha-\sigma_t$, for all $t \in \mathcal{T}$. To identify the returns dynamics under $Q^{(e)}$ we start by evaluating the conditional cumulant generating function of y_t given $\mathcal{F}_{t-1}$ under $Q^{(e)}$.

$$
\begin{aligned}
\kappa_{y_t}^{Q^{(e)}}(u) &= \kappa_{y_t}^{P}(u+\theta_t^*) - \kappa_{y_t}^{P}(\theta_t^*) \\
&= (m_t + \mu\sigma_t)u + \frac{\lambda}{2}\ln\frac{\alpha^2 - (\beta+\theta_t^*\sigma_t)^2}{\alpha^2 - (\beta+\theta_t^*\sigma_t + u\sigma_t)^2} \\
&\quad + \ln\frac{K_\lambda(\delta\sqrt{\alpha^2 - (\beta+\theta_t^*\sigma_t + u\sigma_t)^2})}{K_\lambda(\delta\sqrt{\alpha^2 - (\beta+\theta_t^*\sigma_t)^2})}. \qquad (3.21)
\end{aligned}
$$

This expression is well defined provided that $|\beta + \theta_t^*\sigma_t + u\sigma_t| < \alpha$. From (3.21) we see that under $Q^{(e)}$, the conditional distribution of the return y_t given $\mathcal{F}_{t-1}$ is again a GH distribution as follows:

$$
y_t|\mathcal{F}_{t-1} \sim \mathbf{GH}(\lambda, \alpha/\sigma_t, \beta_{1t}/\sigma_t, \delta\sigma_t, m_t + \sigma_t\mu), \qquad (3.22)
$$

where $\beta_{1t} = \beta + \theta_t^*\sigma_t$. We notice that the conditional return distribution after the measure change arising from $Q^{(e)}$ is obtained by shifting only the skewness of the original distribution with θ_t^*. Therefore, the return dynamics under $Q^{(e)}$ are:

$$
y_t = m_t + \sigma_t\eta_t, \qquad (3.23)
$$

$$
\eta_t|_{\mathcal{F}_{t-1}} \sim \mathbf{GH}(\lambda, \alpha, \beta + \theta_t^*\sigma_t, \delta, \mu). \qquad (3.24)
$$

In order to represent y_t in the form given by (3.15)–(3.16) we denote by $\vartheta_{1t} = E^P[\eta_t|\mathcal{F}_{t-1}] - \mu$ and ϑ_{2t} the conditional standard deviation of η_t and we let $\xi_t = \eta_t/\vartheta_{2t} - (\vartheta_{1t} + \mu)/\vartheta_{2t}$. □

We conclude that, under the risk-neutral conditional Esscher transform $Q^{(e)}$, the conditional mean and standard deviation of the returns, $m^{Q^{(e)}}$ and $\sigma^{Q^{(e)}}$, are given by:

$$
m_t^{Q^{(e)}} = m_t + \sigma_t\mu + \frac{\sigma_t\delta(\beta+\theta_t^*\sigma_t)}{\sqrt{\alpha^2 - (\beta+\theta_t^*\sigma_t)^2}} R_\lambda(\delta\sqrt{\alpha^2 - (\beta+\theta_t^*\sigma_t)^2}), \qquad (3.25)
$$

$$
\begin{aligned}
\sigma_t^{Q^{(e)}} &= \sigma_t\left(\frac{\delta}{\sqrt{\alpha^2 - (\beta+\theta_t^*\sigma_t)^2}} R_\lambda(\delta\sqrt{\alpha^2 - (\beta+\theta_t^*\sigma_t)^2})\right. \\
&\quad \left. + \frac{\delta^2(\beta+\theta_t^*\sigma_t)^2}{\alpha^2 - (\beta+\theta_t^*\sigma_t)^2} S_\lambda(\delta\sqrt{\alpha^2 - (\beta+\theta_t^*\sigma_t)^2})\right)^{\frac{1}{2}}. \qquad (3.26)
\end{aligned}
$$

Thus, we identify a pricing kernel which preserves the parametric form of the return's distribution after the measure change. In particular, the conditional risk-neutral distribution of return is obtained by shifting the skewness parameter while keeping the other parameter constant. We remark that the risk-neutral return dynamics have no longer a linear GARCH form since the innovations $\{\xi_t\}$ are not independent and identically distributed and the conditional volatility of return cannot be updated under the conditional Esscher transform directly.

One can still simulate the process $\{y_t\}$ under this new probability measure using (3.15)–(3.16), where the conditional variance process is filtered according to (3.11).

In the following proposition, we give the risk-neutral dynamics of return arising from the mean correcting martingale measure for the GH-GARCH model from (3.9)–(3.11). This represents another measure change which preserves the parametric form of the return's distribution.

Theorem 3.2. *Suppose the asset return process $y := \{y_t\}_{t\in\mathcal{T}}$ satisfies equations* (3.9)–(3.11). *Under the mean correcting martingale measure $Q^{(m)}$, the dynamics of the return have the following form*:

$$y_t = r_t - \mu\sigma_t + \frac{\lambda}{2}\ln\frac{\alpha^2-(\beta+\sigma_t)^2}{\alpha^2-\beta^2} + \ln\frac{K_\lambda(\delta\sqrt{\alpha^2-\beta^2})}{K_\lambda(\delta\sqrt{\alpha^2-(\beta+\sigma_t)^2})} + \sigma_t\varepsilon'_t, \tag{3.27}$$

$$\varepsilon'_t \sim \mathbf{GH}(\lambda,\alpha,\beta,\delta,\mu), \tag{3.28}$$

$$\sigma_t^2 = a_0 + a_1\sigma_{t-1}^2\varphi\left(\varepsilon'_{t-1} + \frac{m_{t-1}^{\text{shift}}}{\sigma_{t-1}}\right) + b_1\sigma_{t-1}^2, \tag{3.29}$$

where m_t^{shift} is given by:

$$m_t^{\text{shift}} = r_t - m_t - \mu\sigma_t + \frac{\lambda}{2}\ln\frac{\alpha^2-(\beta+\sigma_t)^2}{\alpha^2-\beta^2} + \ln\frac{K_\lambda(\delta\sqrt{\alpha^2-\beta^2})}{K_\lambda(\delta\sqrt{\alpha^2-(\beta+\sigma_t)^2})}. \tag{3.30}$$

Proof. The proof follows immediately by substituting the cumulant function of the generalized hyperbolic distribution of the innovations into (2.11). □

By its construction, the mean correcting martingale measure preserves the same Generalized Hyperbolic structure. In this case, we are able to represent the return dynamic as a simple GARCH structure, making thus easier the simulation of the return process under this new measure.

The risk-neutralized return dynamics under the GLRNVR are obtained from (2.17)–(2.19) by replacing F^{-1} with the inverse cdf of the GH distribution.

In the next section we discuss the consistency between the risk neutral measures described in Sec. 2 and equilibrium measures constructed based on consumption CAPM models within our GH-GARCH framework.

3.2. *Consistency with equilibrium measures*

Using a similar framework as in Cochrane [18] we consider a representative economic agent with a strictly increasing, time separable, additive utility function $u(c_t)$, where c_t represents the aggregate consumption. The standard utility maximization problem leads to the following form of the Euler equilibrium SDF, which is also known as the Euler condition in the economic literature:

$$M_t = \rho \frac{u'(c_t)}{u'(c_{t-1})}. \tag{3.31}$$

Here $u'(c_t)$ represents the marginal utility at time t and ρ is the impatient factor. In this section we focus on only two main classes of utilities, namely, exponential and power utility functions, though extensions of our results to other general cases such as recursive utility functions of Epstein and Zin [27] can also be further explored. If the utility is of an exponential form, (i.e. $u(c_t) = (1 - \exp(-Rc_t))/R$, where R is the positive coefficient of absolute risk aversion), then the SDF from (3.31) becomes $M_t = \rho \exp(-R\Delta c_t)$, where Δ is the difference operator, (i.e. $\Delta c_t = c_t - c_{t-1}$). In the case of a power utility of consumption, (i.e. $u(c_t) = (c_t^{1-R} - 1)/(1 - R)$, where R is the coefficient of relative risk aversion and $R > 1$), the equilibrium SDF is, $M_t = \rho \exp(-R\Delta \ln c_t)$. Thus, depending on the form of the utility functions, we explore connections between changes in aggregate consumption (log aggregate consumption), Δc_t ($\Delta \ln c_t$), and the asset return process, y_t. However, since the forms of the SDFs arising from the two utility functions are similar, we focus on justifying our results only for the exponential utility case and the power utility follows immediately by essentially the same approach.

Using a similar argument as in Schroder [49], the consistency between the equilibrium SDF and the one given by the mean correcting martingale measure is satisfied if the expressions in (2.7) and (3.31) are equal. This leads to:

$$e^{-r_t} \frac{f_t^P(\varepsilon_t + \varrho_t)}{f_t^P(\varepsilon_t)} = \rho \frac{u'(c_t)}{u'(c_{t-1})}, \tag{3.32}$$

where ϱ_t represents the market price of risk from (2.10). For an exponential utility function and using the probability density function from (3.3) for the innovation process ε_t, (3.32) holds if:

$$\begin{aligned} \Delta c_t = \frac{1}{R} \Bigg[& r_t - \beta \varrho_t - \ln \frac{K_{\lambda-1/2}(\alpha\sqrt{\delta^2 + (\varepsilon_t + \varrho_t - \mu)^2})}{K_{\lambda-1/2}(\alpha\sqrt{\delta^2 + (\varepsilon_t - \mu)^2})} \\ & + \frac{1}{2}\left(\frac{1}{2} - \lambda\right) \ln \rho \frac{\delta^2 + (\varepsilon_t + \varrho_t - \mu)^2}{\delta^2 + (\varepsilon_t - \mu)^2} \Bigg]. \end{aligned} \tag{3.33}$$

Although we are able to express the change in aggregate consumption as a function of the asset return process, the above relation does not appear convenient for practical applications since y_t appears in the right hand side of (3.33) as an argument of the modified Bessel function. The same relationship can be derived for changes in log-aggregate consumptions in the power utility case. Thus, a consistency between the mean correcting martingale measure and consumption CAPM models seems difficult to be implemented in practice for the GH-GARCH settings.

Following the same argument, one can derive a relationship between aggregate consumption and asset returns for the conditional Esscher transform. Indeed, if we equate the SDF expressions from (2.14) and (3.31) for the exponential utility case, we obtain:

$$\Delta c_t = \frac{1}{R}[\ln \rho + r_t + \kappa^P_{y_t}(\theta^*_t) - \theta^*_t y_t]. \tag{3.34}$$

According to the above equation, the conditional Esscher transform coincides with the equilibrium pricing measure if changes in aggregate consumptions have the above linear dependence of the asset returns, $\Delta c_t = k_{1t} + k_{2t} y_t$, where k_{1t} and k_{2t} are two $\{\mathcal{F}_t\}$-predictable processes such that $k_{1t} = (\ln \rho + r_t + \kappa^P_{y_t}(\theta^*_t))/R$ and $k_{2t} = -\theta^*_t y_t/R$. Although this above representation is appealing, we argue that the linearity assumption can be relaxed, such that we can construct another risk neutral measure based on equilibrium arguments which is also consistent with the conditional Esscher transform. More specifically, along the line of Duan's [19] local risk neutral valuation relationship, we define a martingale measure based on a conditional bivariate assumption of the returns and changes in aggregate consumption or changes in log-aggregate consumption. Following Schmidt *et al.* [47] we first define the notion of a bivariate affine Generalized Hyperbolic (BAGH) random variable.

Definition 3.1. A 2-dimensional random variable $X = (X_1, X_2)^T$ is said to have a bivariate affine Generalized Hyperbolic (BAGH) distribution with location vector $\mu \in \mathbb{R}^2$ and scaling matrix $\Gamma \in \mathbb{R}^{2\times 2}$ if it has the following stochastic representation:

$$X = \Sigma^T Y + \mu, \tag{3.35}$$

where Σ is a lower triangular matrix such that $\Gamma = \Sigma^T \Sigma$ is positive definite and the vector components of Y are mutually independent univariate Generalized Hyperbolic random variables with $Y_i \sim \mathbf{GH}(\lambda_i, \alpha_i, \beta_i, 1, 0)$.

The BAGH distribution was introduced by Schmidt *et al.* [47] to capture some phenomena that cannot be modelled with the standard bivariate Generalized Hyperbolic (BGH) distribution. For example, one disadvantage of the BGH is that its marginal distributions cannot be mutually independent for any choice of the scaling matrix Γ, while in the BAGH the marginal distributions are for example independent when the scaling matrix is equal to the identity matrix. Moreover, the consistency between the conditional Esscher transform and equilibrium measures cannot be obtained under BGH distributional assumptions.

Theorem 3.3. *Suppose the asset return process $y := \{y_t\}_{t\in\mathcal{T}}$ satisfies equations* (3.9)–(3.11)*. Assume either one of the following conditions holds*:

(a) *The utility function is of exponential form and for any $t = 1, \ldots, T$ changes in aggregate consumption and asset returns are BAGH distributed conditional on $\mathcal{F}_{t-1}$.*
(b) *The utility function is of power form and for any $t = 1, \ldots, T$ changes in logarithm aggregate and asset returns are BAGH distributed conditional on $\mathcal{F}_{t-1}$.*

Then the risk neutral measure based on the equilibrium SDF from (3.31) is consistent with the SDF arising from the conditional Esscher transform.

Proof. We need only to show the consistency is realized when assumption (a) holds, since under assumption (b) the proof follows in a similar way. For any $t = 1, \dots, T$ we denote by $\mu_t = (\mu_t^1, \mu_t^2)^T$ and $\Sigma_t = (\Sigma_t^{ij})_{1 \le i,j \le 2}$ be two $\mathcal{F}_t$-predictable processes representing the location vector and the lower triangular matrix from the scaling matrix decomposition, respectively. The stochastic representation from the BAGH assumption leads to:

$$\Delta c_t = \Sigma_t^{11} X_1 + \Sigma_t^{21} X_2 + \mu_t^1 \tag{3.36}$$

$$y_t = \Sigma_t^{22} X_2 + \mu_t^2. \tag{3.37}$$

Here X_1 and X_2 are two independent, unit scale and zero location random variables GH distributed. Moreover, to be consistent with the asset return dynamics under P given in (3.9) we let $X_2 \sim \mathbf{GH}(\lambda, \delta\alpha, \delta\beta, 1, 0)$, $\Sigma_t^{22} = \delta\sigma_t$, and $\mu_t^2 = m_t + \mu\sigma_t$. We define $\tilde{Q}$ the risk neutral measure induced by the equilibrium SDF from (3.31) for exponential utility functions with the Radon-Nikodym derivative given by:

$$\frac{d\tilde{Q}}{dP} = \prod_{t=1}^{T} \rho e^{r_t - R\Delta c_t}$$

Using the Bayes' rule, we evaluate the cumulant generating function of the asset returns under $\tilde{Q}$ conditional on $\mathcal{F}_{t-1}$:

$$\kappa_{y_t}^{\tilde{Q}}(u) = \ln E^{\tilde{Q}}[e^{uy_t} | \mathcal{F}_{t-1}] = \ln E^P[\rho e^{uy_t} e^{r_t - R\Delta c_t} | \mathcal{F}_{t-1}] \tag{3.38}$$

$$= \ln E^P[\rho e^{r_t - R\Sigma_t^{11} X_1 - R\mu_t^1 + R\Sigma_t^{21} \frac{m_t + \mu\sigma_t}{\delta\sigma_t}} e^{(u - R\frac{\Sigma_t^{21}}{\delta\sigma_t}) y_t} | \mathcal{F}_{t-1}] \tag{3.39}$$

$$= r_t + \ln\rho - R\mu_t^1 + R\Sigma_t^{21} \frac{m_t + \mu\sigma_t}{\delta\sigma_t} + \kappa_{X_1}^P(-R\Sigma_t^{11})$$

$$+ \kappa_{y_t}^P\left(u - R\frac{\Sigma_t^{21}}{\delta\sigma_t}\right). \tag{3.40}$$

Imposing the condition that $\kappa_{y_t}^{\tilde{Q}}(0) = 0$ we have:

$$\kappa_{y_t}^P\left(-R\frac{\Sigma_t^{21}}{\delta\sigma_t}\right) = - r_t - \ln\rho + R\mu_t^1 - R\Sigma_t^{21} \frac{m_t + \mu\sigma_t}{\delta\sigma_t} - \kappa_{X_1}^P(-R\Sigma_t^{11}). \tag{3.41}$$

Plugging this into the above equation for the cumulant generating function under $\tilde{Q}$, we get:

$$\kappa_{y_t}^{\tilde{Q}}(u) = \kappa_{y_t}^P\left(u - R\frac{\Sigma_t^{21}}{\delta\sigma_t}\right) - \kappa_{y_t}^P\left(-R\frac{\Sigma_t^{21}}{\delta\sigma_t}\right). \tag{3.42}$$

The above representation corresponds to the cumulant generating function of the asset returns under the conditional Esscher transform where the Esscher parameter is given by:

$$\theta_t^* = -R\frac{\Sigma_t^{21}}{\delta\sigma_t}. \tag{3.43}$$

The martingale condition is then verified by solving for θ_t^* from $\kappa_{y_t}^{\tilde{Q}}(1) = r_t$. The fact that $\Sigma_t^T\Sigma_t$ is positive definite is obtained if we require that $\Sigma_t^{11} > 0$, for any $t = 1, \ldots, T$. □

We have shown that the conditional Esscher transform is consistent with a risk neutral measure constructed based on equilibrium arguments assuming that the changes in aggregate or log-aggregate consumption and asset returns follow a conditional BAGH distribution.

In the current literature, there are various subclasses and limits of the GH distributions which are frequently used for modelling financial returns. Using the results from Theorems 3.1 and 3.2 we derive risk neutralized dynamics for three special distributions, namely, the Hyperbolic distribution (HYP), the Normal Inverse Gaussian (NIG) distribution and the Variance Gamma (VG). The results are stated in Appendix A.

4. Empirical Analysis of GH-GARCH Models

In this section we examine the pricing performance of the GH-GARCH models described in the previous section using a two stage procedure. Firstly, instead of using market option prices data for model calibration, we consider a canonical approach and use historical prices of the $S\&P500$ index to estimate the unknown parameters in the proposed option valuation models. In general, this asset price information may not be sufficient for estimating the option valuation models, since option valuation is a forward-looking problem and the past may not smoothly pass over to the future. However, our results indicate very small price prediction errors for the non-normal GARCH models considered here. In the second stage we compute European Call option prices written on $S\&P500$ based on the parameter estimates obtained in the previous stage for HYP, NIG and VG-GARCH models. The pricing performance of our proposed models is analyzed relative to observed market quotes using the risk neutral measures derived in Appendix A.

4.1. *Estimation results*

The parameters for the various GARCH models are estimated under the physical measure using the maximum likelihood estimation. We use the daily closing prices of $S\&P500$ from January 02, 1988 to April 17, 2002, for a total of 3,606 observations. The data used here were taken from Yahoo Finance.

We assume the returns have a GARCH-in-mean structure under the physical measure P, with mean given by $m_t = r + \nu\sigma_t$. In a recent study, Christoffersen and Jacobs [13] found that an asymmetric model constructed by including a leverage effect in the variance equation performs the best in terms of pricing European options when the innovation is normally distributed. Thus, in our simulation study we use a threshold GARCH structure. More specifically we use the TGARCH model introduced by Glosten *et al.* [29] for modelling the conditional variance[5]:

$$\sigma_t^2 = a_0 + a_1\varepsilon_{t-1}^2\sigma_{t-1}^2 + \gamma I(\varepsilon_{t-1} < 0)\varepsilon_{t-1}^2\sigma_{t-1}^2 + b_1\sigma_{t-1}^2, \tag{4.1}$$

where ε_t are i.i.d. with common probability density function from equation (3.3), (i.e., GH-distributed), and I is the indicator function of the event $(\varepsilon_{t-1} < 0)$.

Estimation is done using the maximum likelihood method (MLE) for each probability density function of the innovation, conditional on the observed value at time zero and substituting the initial conditional variance at time 0 by the stationary variance. The notation $l(\hat{\theta})$ denotes the log likelihood evaluated at the estimate $\hat{\theta}$, and θ is the generic symbol used to represent the vector-valued parameter of the model considered. An alternative estimation method would be to use a less efficient quasi-maximum likelihood estimation based on a normal density, but the more efficient maximum likelihood is used here. Suppose that we have an observed time series $y_t, t = -s+1, \ldots, T$, where s denotes the required number of initial values. Conditional on the initial values $y_{-s+1}, \ldots, y_0$, the log-likelihood function of the TGARCH-in-mean model with a conditional GH distribution is

$$\begin{aligned} l(\theta) &= \sum_{t=1}^{T} \log f_{y_t}^P(y_t) \qquad (4.2)\\ &= T \log \frac{(\alpha^2 - \beta^2)^{\frac{\lambda}{2}}}{\sqrt{2\pi}\sigma_t\delta^\lambda K_\lambda(\delta\sqrt{\alpha^2 - \beta^2})} \\ &\quad + \left(\lambda - \frac{1}{2}\right)\sum_{t=1}^{T} \log\left(\sqrt{\delta^2 + \left(\frac{y_t - m_t}{\sigma_t} - \mu\right)^2} \Big/ \alpha\right) \\ &\quad + \sum_{t=1}^{T} \log K_{\lambda-\frac{1}{2}}\left(\alpha\sqrt{\delta^2 + \left(\frac{y_t - m_t}{\sigma_t} - \mu\right)^2}\right) + \beta\sum_{t=1}^{T}\left(\frac{y_t - m_t}{\sigma_t} - \mu\right). \qquad (4.3) \end{aligned}$$

We first discuss the estimation results for the TGARCH(1,1)-in-Mean model with normally distributed innovations. The results are presented in the first column of the Table 1. We note that all the parameters are significant and the leverage parameter γ has an expected positive sign; b_1 is approximately 0.95 which is consistent with most of the empirical results in the literature on fitting a normal TGARCH model on stock index data.

The estimation results for the TGARCH(1,1)-in-Mean models with the conditional distributions NIG, HYP and VG are presented in the remaining columns

[5] We have also tested the pricing performance of some other specifications of conditional volatility such as standard GARCH, NGARCH and EGARCH. Our results which are not reported in the current study suggest that the GJR model gives the lowest pricing error compared to the observed market quotes.

Table 1. Estimation results for normal and normal variance mean mixture GARCH models.

	Normal	NIG	HYP	VG
ν	0.0443	0.042	0.0421	0.0422
	(0.0165)	(0.0167)	(0.0159)	(0.0166)
a_0	$1.11 \cdot 10^{-6}$	$8.4 \cdot 10^{-7}$	$8.4 \cdot 10^{-7}$	$8.6 \cdot 10^{-7}$
	$(3.1 \cdot 10^{-7})$	$(2.5 \cdot 10^{-7})$	$(2.6 \cdot 10^{-7})$	$(2.8 \cdot 10^{-7})$
a_1	0.0076	0.0093	0.009	0.0087
	(0.0057)	(0.0068)	(0.0067)	(0.0068)
b_1	0.9428	0.944	0.9439	0.9439
	(0.0097)	(0.0099)	(0.01)	(0.0111)
γ	0.0721	0.0738	0.0735	0.0735
	(0.0141)	(0.0160)	(0.0159)	(0.0167)
$\tilde{\alpha}$		1.6893	1.3992	
		(0.2375)	(0.2536)	
$\tilde{\beta}$		−0.1916	−0.1139	
		(0.0682)	(0.0462)	
ρ				−0.0654
				(0.0221)
λ				2.1145
				(0.2165)
Skw	0	−0.263	−0.219	−0.190
Kts	0	1.879	1.592	1.441
$l(\hat{\theta})$	−4697.3	−4551.5	−4554.5	−4557.5
AIC	9404.6	9117.0	9123.0	9129.0
BIC	9435.6	9160.3	9166.3	9172.3
χ^2	0.0000	0.3336	0.2717	0.2948
$Q(20)$	34.14	34.60	34.59	34.55
$Q^2(20)$	4.96	5.67	5.60	5.54

Note: AIC and BIC are the Akaike and Bayes information criteria. χ^2 is a p-value of a chi-square goodness-of-fit test for standardized residuals against the theoretical density. $Q(20)$ and $Q^2(20)$ denote the Ljung-Box statistic for serial correlation in the standardized residuals and squared standardized residuals for up to lag 20.

of Table 1. It is first observed that the estimates of the conditional mean and variance parameters are only moderately affected by the introduction of the skewed innovations' distributions. In other words, these estimates appear to be approximately the same in all the estimated models. There is also little change in the accuracy of these parameter estimators. The shape parameters are statistically significant for each model with skewed innovation's distribution. For each of these models, the GARCH parameter a_1 in (4.1) is not statistically significant.

The estimated shape parameters of the skewed distributions show that all the fitted distributions have negative skewness. This follows since for NIG and HYP the estimate of $\widetilde{\beta}$ is significantly negative, and for VG the estimate of ρ is negative.

This can be tested by a LR-test, for example $\widetilde{\alpha}-\widetilde{\beta}=0$ (symmetry) versus $\widetilde{\alpha}-\widetilde{\beta}<0$. The null hypothesis of symmetry is clearly rejected with the p-value of 0.004.

Model comparison may be made by either the Akaike or Bayes, information criteria

$$\mathrm{AIC} = -2l(\hat{\theta}) + 2d, \tag{4.4}$$

$$\mathrm{BIC} = -2l(\hat{\theta}) + d\log N, \tag{4.5}$$

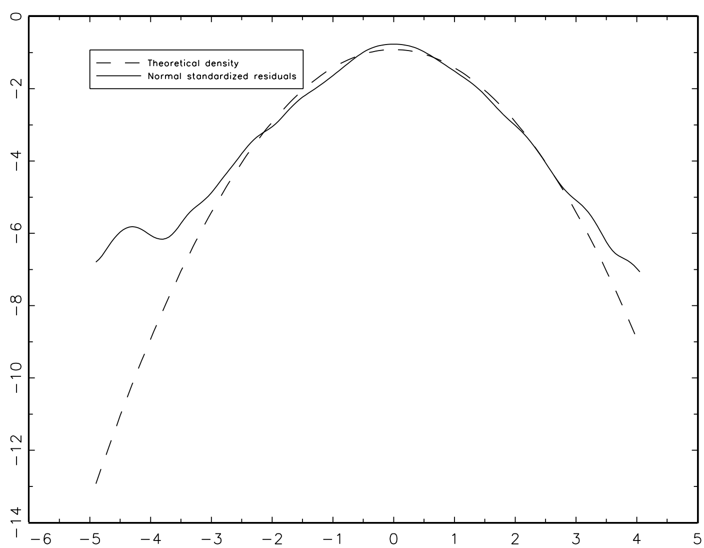
where $l(\hat{\theta})$ is the log likelihood evaluated at the MLE $\hat{\theta}$, d denotes the number of parameters of the model and N the number of observations. A smaller AIC or BIC, indicates a better model. From Table 1 we find a significant improvement over the normal model by any of these skewed models. Of all the models, the NIG model is the best, according to both the AIC and BIC, while the VG model is the least favorable.

Table 1 also gives the point estimates of the skewness (skw) and kurtosis (kts) of the standardized residuals for the fitted models, that are calculated for a given model using the MLE $\hat{\theta}$ substituted for the *true* parameter. These observed values of skewness vary from -0.19 (the VG model) to -0.26 (the NIG model). The observed kurtoses vary from 1.44 to 1.88, respectively. These values give evidence for the departure from normality. One may also compute the Jarque-Bera (JB) test for normality, but the normal model for the innovation has already been rejected. When fitting the normal TGARCH model, the observed value of the JB test applied to the standardized residuals is 4406.9 which provides sufficient evidence against the normality assumption for the innovation.

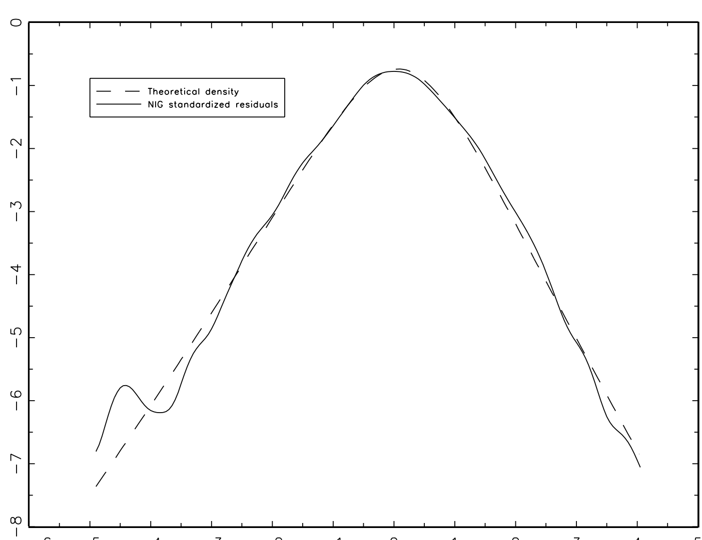
The improvement in the empirical performance of the model achieved from the skewed distributions is illustrated in Fig. 1. Here the estimated log densities of the standardized residuals of the normal and GH-TGARCH models, obtained using a kernel density estimator in R with the default bandwith and kernel, are plotted against the theoretical log densities of these models based on the estimated parameters in Table 1. Plot (a) shows that the normal model is unable to capture either the skewness or the kurtosis. The other non-Gaussian models capture both the skewness and the kurtosis much better. All of them are similar, with relatively small differences. This is consistent with the conclusion based on the AIC and BIC selection methods reported in Table 1. They do fail to capture the behaviour in the lower or left tail. However this is less serious than failing to capture the upper or right tail which has a greater effect on the expected value of payoff functions, especially call options. See for example Table 2 below or Badescu and Kulperger [2] which supports this observation.

4.2. *Empirical performance of the GH-GARCH models*

The pricing performance of our models is tested relative to 54 European Call options on the $S\&P500$ index at the close of the market on April 18, 2002. The data were taken from [48]. On April 18, 2002, the closing price was $S_0 = \$1124.47$, the annual risk free rate was $r = 1.9\%$, and the dividend yield was $d = 1.2\%$. The strike prices range from \$975 to \$1325 and we consider options with maturities $T = 22$, 46, 109, 173, and 234 days. The average option price is \$56.94.

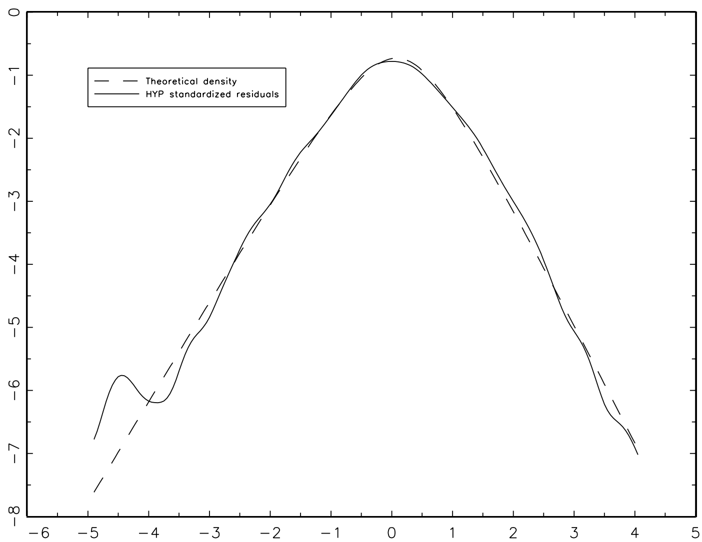


(a) Normal TGARCH vs theoretical

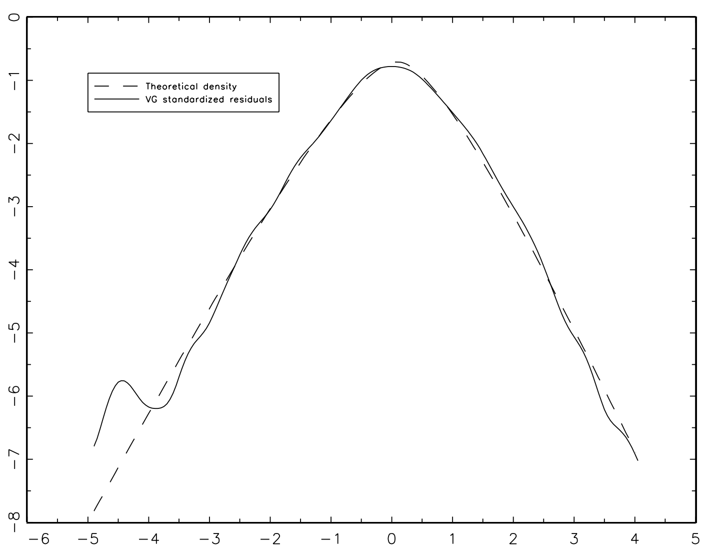


(b) NIG-TGARCH vs. theoretical

Fig. 1. Logarithmic densities of the Normal, NIG, HYP and VG standardized residuals versus the theoretical densities based on MLE estimates using returns from January 04, 1988 to April 17, 2002. The dashed line is used to represent the theoretical log densities.



(c) HYP-TGARCH vs. theoretical



(d) VG-TGARCH vs. theoretical

Fig. 1. (*Continued*)

Based on the parameters estimated from historical returns, prices for European Calls are computed using Monte-Carlo simulation under the three price kernels considered in Sec. 2.

There are two natural ways one can calculate such prices: (i) simulate asset paths under the risk neutral measure considered and then evaluate prices as an average of payoffs corresponding to Monte-Carlo paths; (ii) simulate asset paths

Table 2. Normal and GH TGARCH pricing errors for European Call options on 18 April 2002 using three risk neutral measures.

Model	RMSE	ARPE (%)	APE (%)
Normal	4.304	7.37	5.92
NIG-MCMM	2.551	5.21	3.53
NIG-ESS	1.498	4.30	2.21
NIG-GLRNVR	2.096	6.17	3.47
HYP-MCMM	2.748	6.00	3.75
HYP-ESS	2.245	4.84	2.88
HYP-GLRNVR	3.166	6.31	4.14
VG-MCMM	3.001	6.51	4.17
VG-ESS	2.229	4.85	3.19
VG-GLRNVR	2.024	4.66	2.96

under the physical measure P and evaluate option prices as a weighted average of the payoff for each of the corresponding path, where the weights are given by the Radon-Nikodym derivative evaluated for this Monte-Carlo simulated path. This method may be viewed as a version of the importance sampling technique.

In addition, the latter is quite useful when there is an explicit form of the return dynamics under P but not tractable return dynamics under Q, thus making method (i) difficult to implement. Also, method (ii) provides a convenient way for variance reduction. Indeed, we tested this by numerical simulations and we found that the variance of the Monte-Carlo estimator is lower when simulating under P. For a more detailed discussion about the variance reduction property of the Esscher transform we refer to Badescu and Kulperger [2]. In general, one may wish to choose the more efficient method to compute option prices, even though the risk neutral return dynamics are available and (i) is convenient to implement. In our empirical study we compute option prices for the MCMM and for Duan's GLRNVR using the approach described in (i), while for the conditional Esscher transform we use (ii). For all models, the Monte-Carlo estimates for the option prices are calculated based on 50,000 simulations and the last estimated value of the conditional standard deviation from the MLE procedure is used as a starting value for updating the conditional variance process according to the TGARCH specification in (4.1). Implementing the MCMM in practice is generally faster than the other two methods. For example, the conditional Esscher transform is more time consuming for the HYP and VG specifications since at any time point we need to solve numerically for the Esscher parameter θ_t^*, from the martingale equations.

As already mentioned in Stentoft [52], the implementation of Duan's GLRNVR algorithm according to (2.17)–(2.19) relies on approximating $F^{-1}(\Phi(\xi_t - \upsilon_t))$ and evaluating $\kappa_{\eta_t}^{Q^{(g)}}(\sigma_t)$, where F is the corresponding standard GH distribution function, $\{\xi_t\}$ is a sequence of standard Gaussian random variables under $Q^{(g)}$, and υ_t is the market price of risk. Both Duan [20] and Stentoft [52] suggested an approximation of F^{-1} based on a numerical integration over a grid of points covering most of the support of the distribution and then using a linear interpolation between these grid points.

We propose below a simple Monte-Carlo method for approximating the sample quantile function which is computationally faster than the numerical integration and interpolation used in [20].

(1) Simulate a large sample, say $M = 100,000$ of random variables $\epsilon_1, \ldots, \epsilon_M$ from F.
(2) Sort these data into $\epsilon_{(1)}, \ldots, \epsilon_{(M)}$ (the order statistics of the random sample).
(3) For a given $0 < u < 1$ solve for i such that $\frac{i}{M+1} < u \leq \frac{i+1}{M+1}$.
(4) Set the sample quantile function $F^{-1}(u) = \epsilon_{(i)}$ for $i > 0$, and $F^{-1}(u) = \epsilon_{(1)}$ if $i = 0$.
(5) Repeat Steps 3 and 4 as needed for various u.

For approximating $\kappa_{\eta_t}^{Q^{(g)}}(\sigma_t)$ we use a similar Monte-Carlo approach as in [20]. For both numerical approximations we let $\upsilon_t = 0.04$ which is close to the average estimated value of 0.0421 for the risk premium parameter ν across the GH pricing models considered in our study.

The performance evaluated based on real option prices data of the normal GH-TGARCH models is measured with three indicators: (i) the dollar root mean squared error (RMSE), (ii) the average relative pricing error (ARPE) and (iii) the average absolute error (APE) given below.

$$\mathrm{RMSE}(\$) = \sqrt{\sum_{j=1}^{\mathrm{NO}} \frac{(C_j^{\mathrm{market}} - C_j^{\mathrm{model}})^2}{\mathrm{NO}}} \tag{4.6}$$

$$\mathrm{ARPE}(\%) = \frac{1}{\mathrm{NO}} \sum_{j=1}^{\mathrm{NO}} \frac{|C_j^{\mathrm{market}} - C_j^{\mathrm{model}}|}{C_j^{\mathrm{market}}} \times 100 \tag{4.7}$$

$$\mathrm{APE}(\%) = \frac{1}{\mathrm{NO} \cdot \overline{C}^{\mathrm{market}}} \sum_{j=1}^{\mathrm{NO}} |C_j^{\mathrm{market}} - C_j^{\mathrm{model}}| \times 100 \tag{4.8}$$

where NO represents the total number of options and $\overline{C}^{\mathrm{market}}$ is the average option price.

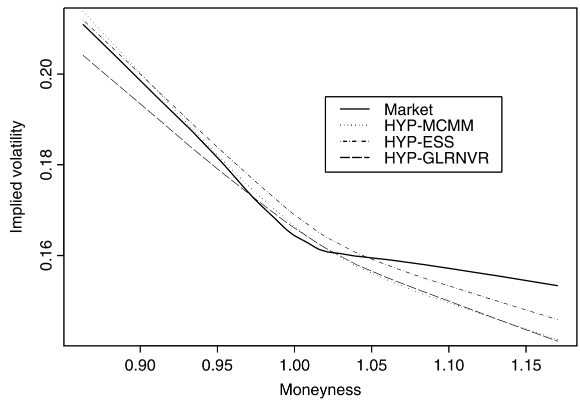
Table 2 summarizes the overall pricing errors of the various models considered here. We notice that all GH-TGARCH models outperform the Gaussian-TGARCH models with respect to all three indicators. Amongst the GH models, the option prices based on the NIG distribution provide the smallest errors for the MCMM and the Esscher transform, while the VG model provides the best fit for Duan's GLRNVR. Option prices are sensitive to the risk neutral measures used. For example, the conditional Esscher transform outperforms the other two counterparts for the NIG-TGARCH and HYP-TGARCH, while the GLRNVR is a better choice for the VG-TGARCH than the other. The NIG specification for the innovations combined with the conditional Esscher transform provides the best pricing model having the lowest RMSE $= 1.498$, compared to 2.245 for the HYP case and 2.229 for the VG driving noise case. However, the differences between the ARPE's are not dramatic (lowest ARPE is 4.30 for NIG and highest is 4.85 for VG) since the Esscher transform does not perform very well for short maturity options ($T = 22$ days) when the market quotes are lower.

Except for the HYP model, option prices computed based on the conditional Esscher transform and the GLRNVR outperform the ones calculated based on the MCMM and an argument for this might be that both former martingale measures are computed based on equilibrium arguments.

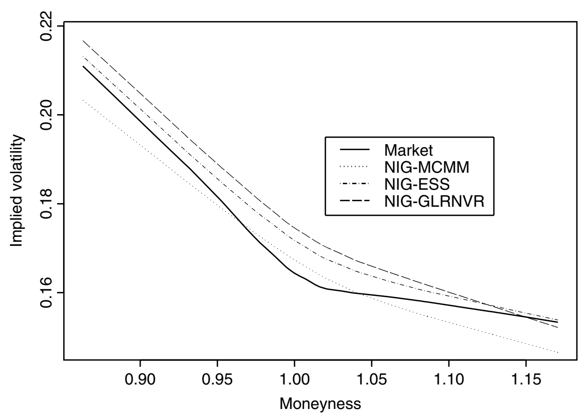
To document the behavior of different pricing measures for short and long-maturity options we give the average relative pricing errors (ARPE) for $T = 46$ days, (i.e., short-maturity options), and $T = 243$ days, (i.e., long-maturity options), in Table 3. We note that when $T = 46$ the MCMM and GLRNVR perform better than the conditional Esscher transform (except for the NIG-GLRNVR case which has the highest ARPE). For long maturity options ($T = 243$ days), the conditional Esscher transform is the best risk-neutral measure giving the lowest ARPE of 2.46% for the NIG-TGARCH model. Another interesting question to ask is how well these models describe the Black–Scholes implied volatility. For both observed market prices and for each model option price we compute the implied volatility based on the Black–Scholes formula. Figure 2 gives these for the Hyperbolic, NIG and VG models described earlier. The implied volatility for each model, and each of the price kernels is plotted. All follow the general shape and pattern for implied volatility, but we see in Fig. 2(b) the NIG innovations' distribution with the conditional Esscher transform does the best overall in terms of mimicking the empirical behavior of the implied volatility from the observed option prices data. Our empirical findings suggest that in general the conditional Esscher transform is a better pricing method for long-maturity options. To support this finding we illustrate in Fig. 3 a comparison between Monte-Carlo option prices computed based on the Normal-TGARCH and NIG-TGARCH and observed market quotes on April 18, 2002, under all three risk neutral measures used.

Table 3. ARPE (in percentage) for European Call options on 18 April 2002 under the three pricing kernels

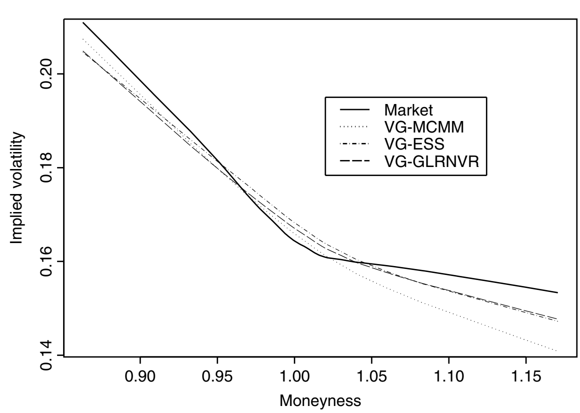
Maturity	Pricing kernel	NIG	HYP	VG
$T = 46$	Esscher	8.79	7.19	6.88
	MCMM	7.02	5.53	6.17
	GLRNVR	12.01	6.61	5.66
$T = 243$	Esscher	2.46	8.78	7.30
	MCMM	8.85	12.21	14.21
	GLRNVR	3.02	14.21	7.33



(a) HYP-TGARCH

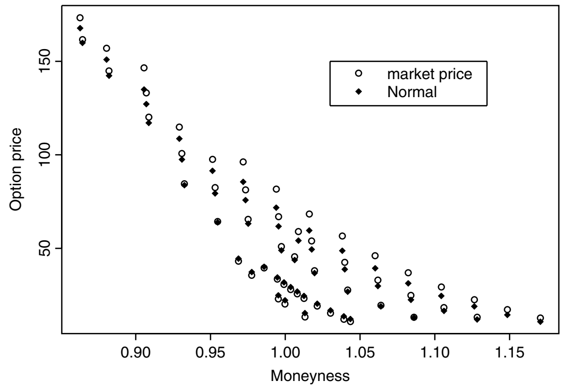


(b) NIG-TGARCH

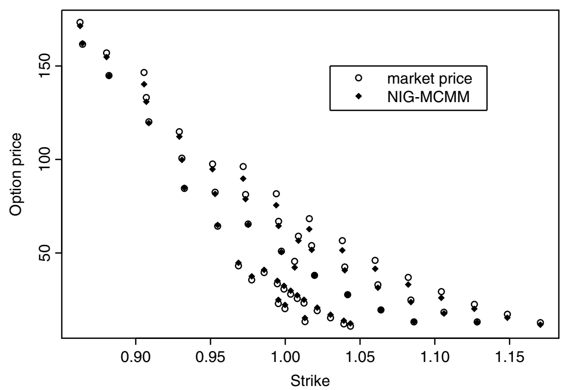


(c) VG-TGARCH

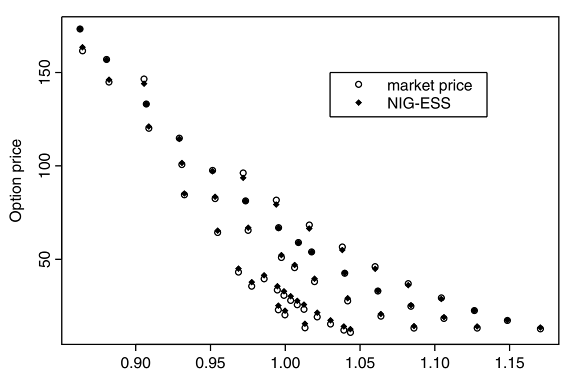
Fig. 2. Black–Scholes implied volatilities for HYP, NIG and VG models.



(a) Normal-TGARCH

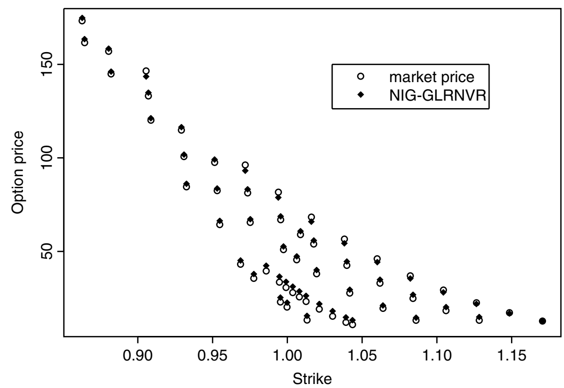


(b) NIG-TGARCH (MCMM)



(c) NIG-TGARCH (Esscher)

Fig. 3. Market prices vs. Monte-Carlo option prices for TGARCH models with Normal and NIG innovations based on the mean correcting martingale measure, the Esscher transform, and the Generalized local risk neutral valuation relationship on 18 April 2002. The maturities are $T = 22, 46, 109, 173$, and 234 days and the inner set of prices corresponds to $T = 22$.



(d) NIG-TGARCH (GLRNVR)

Fig. 3. (*Continued*)

5. Conclusions and Future Directions

This paper studies time series models of a conditional location scale structure, specifically a GARCH structure, with a GARCH-in-mean conditional mean. A normal innovation model does not adequately fit the log returns, while the three non-Gaussian models considered here give a much better fit.

However, for option pricing it is not only necessary to have good historical models, but also the choice of a price kernel plays a significant role in the prediction of real option prices. We studied three particular choices of such price kernels. Using real option data, it is shown that the conditional Esscher transform choice for a price kernel under the NIG case is best, uniformly so for three criteria of RMSE (root mean square error), APRE (average percent relative error) and APE (average absolute error). The conditional Esscher transform provides the best choice of a risk neutral measure for the HYP case as well, while Duan's LRNVR is a better choice for the VG case. Since our empirical findings are based on a relatively small option data set it would be interesting to test the out-of-sample pricing performance of such models and corresponding martingale measures on larger data sets, on various individual stocks and options written on other underlying assets, such as currencies, interest rates and commodities. Following the work of Stentoft [52] one may examine the sensitivity of the pricing kernel when applied to American or other exotic options. The impact of the risk neutral measure used can also be tested by calibrating the model parameters to observed market prices using the standard minimization algorithm.

Appendix A. Subclasses and Limits of the GH–GARCH Model

A.1. *HYP–GARCH*

The Hyperbolic distribution is a particular case of the GH distribution when $\lambda = 1$. The scale and location invariant parameters are $\tilde{\alpha} = \alpha\delta$ and $\tilde{\beta} = \beta\delta$. Firstly, we express the location and scale parameters as functions of the invariant parameters $\tilde{\alpha}$ and $\tilde{\beta}$. From (3.12) and (3.13) we obtain:

$$\tilde{\delta} = \left(\frac{1}{\tilde{\gamma}} R_1(\tilde{\gamma}) + \frac{\tilde{\beta}^2}{\tilde{\gamma}^2} S_1(\tilde{\gamma})\right)^{-\frac{1}{2}},$$

$$\tilde{\mu} = -\frac{\tilde{\delta}}{\tilde{\gamma}} \tilde{\beta} R_1(\tilde{\gamma}),$$

where $\tilde{\gamma} = \sqrt{\tilde{\alpha}^2 - \tilde{\beta}^2}$. The probability density function of the innovations becomes:

$$f_{\varepsilon_t}(x) = \frac{\tilde{\gamma}}{2\tilde{\delta}\tilde{\alpha} K_1(\tilde{\gamma})} \exp\left(-\tilde{\alpha}\sqrt{1 + \left(\frac{x - \tilde{\mu}}{\tilde{\delta}}\right)^2} + \tilde{\beta}\left(\frac{x - \tilde{\mu}}{\tilde{\delta}}\right)\right),$$

where $x \in \mathbf{R}$; $\tilde{\delta} > 0$; $|\tilde{\beta}| < \tilde{\alpha}$; $\tilde{\delta}$ and $\tilde{\mu}$ are given above. Thus ε_t follows a zero mean and unit variance Hyperbolic distribution with invariant parameters $\tilde{\alpha}$ and $\tilde{\beta}$. We write:

$$\varepsilon_t \sim \mathbf{Hyp}(\tilde{\alpha}, \tilde{\beta}, \tilde{\delta}, \tilde{\mu}), \quad t \in \mathcal{T} \backslash \{0\}.$$

Corollary A.1. *Under $Q^{(e)}$ the returns dynamics are given by:*

$$y_t = m_t + \sigma_t(\tilde{\mu} + \tilde{\delta}\vartheta_{1t}) + \sigma_t \tilde{\delta} \vartheta_{2t} \xi_t,$$

$$\xi_t | \mathcal{F}_{t-1} \sim \mathbf{Hyp}\left(\tilde{\alpha}, \tilde{\beta}_{1t}, \frac{1}{\vartheta_{2t}}, -\frac{\vartheta_{1t}}{\vartheta_{2t}}\right),$$

where $\{\tilde{\beta}_{1t}\}$, $\{\vartheta_{1t}\}$, $\{\vartheta_{2t}\}$ are some $\{\mathcal{F}_t\}$-predictable processes given by:

$$\tilde{\beta}_{1t} = \tilde{\beta} + \theta_t^* \tilde{\delta} \sigma_t,$$

$$\vartheta_{1t} = \frac{\tilde{\beta}_{1t}}{\sqrt{\tilde{\alpha}^2 - \tilde{\beta}_{1t}^2}} R_1(\sqrt{\tilde{\alpha}^2 - \tilde{\beta}_{1t}^2}),$$

$$\vartheta_{2t} = \left(\frac{1}{\sqrt{\tilde{\alpha}^2 - \tilde{\beta}_{1t}^2}} R_1\left(\sqrt{\tilde{\alpha}^2 - \tilde{\beta}_{1t}^2}\right) + \frac{\tilde{\beta}_{1t}^2}{\tilde{\alpha}^2 - \tilde{\beta}_{1t}^2} S_1\left(\sqrt{\tilde{\alpha}^2 - \tilde{\beta}_{1t}^2}\right)\right)^{\frac{1}{2}}, \quad t \in \mathcal{T} \backslash \{0\},$$

such that ξ_t has zero conditional mean and unit conditional variance given $\mathcal{F}_{t-1}$, and θ_t^* is the unique predictable solution of the following martingale equation:

$$\frac{1}{2} \ln\left(\frac{\tilde{\alpha}^2 - (\tilde{\beta} + (1 + \theta_t^*)\tilde{\delta}\sigma_t)^2}{\tilde{\alpha}^2 - (\tilde{\beta} + \theta_t^* \tilde{\delta}\sigma_t)^2}\right) + \ln\left(\frac{K_1\left(\sqrt{\tilde{\alpha}^2 - (\tilde{\beta} + \theta_t^* \tilde{\delta}\sigma_t)^2}\right)}{K_1\left(\sqrt{\tilde{\alpha}^2 - (\tilde{\beta} + (1 + \theta_t^*)\tilde{\delta}\sigma_t)^2}\right)}\right)$$
$$= \tilde{\mu}\sigma_t + m_t - r.$$

The following corollary gives the risk-neutral return dynamics under the mean-correcting martingale measure.

Corollary A.2. *Under $Q^{(m)}$ the return process y_t has the following representation*:

$$y_t = r - \tilde{\mu}\sigma_t + \frac{1}{2}\ln\left(\frac{\tilde{\alpha}^2 - (\tilde{\beta} + \sigma_t\tilde{\delta})^2}{\tilde{\alpha}^2 - \tilde{\beta}^2}\right) + \ln\left(\frac{K_1\left(\sqrt{\tilde{\alpha}^2 - \tilde{\beta}^2}\right)}{K_1\left(\sqrt{\tilde{\alpha}^2 - (\tilde{\beta} + \sigma_t\tilde{\delta})^2}\right)}\right) + \sigma_t\varepsilon'_t,$$

$$\varepsilon'_t \sim \mathbf{Hyp}(\tilde{\alpha}, \tilde{\beta}, \tilde{\delta}, \tilde{\mu}),$$

$$\sigma_t^2 = a_0 + a_1\sigma_{t-1}^2\varphi\left(\varepsilon'_{t-1} + \frac{m_{t-1}^{\text{shift}}}{\sigma_{t-1}}\right) + b_1\sigma_{t-1}^2,$$

where m_t^{shift} is given by:

$$m_t^{\text{shift}} = r - m_t - \tilde{\mu}\sigma_t + \frac{1}{2}\ln\frac{\tilde{\alpha}^2 - (\tilde{\beta} + \sigma_t\tilde{\delta})^2}{\tilde{\alpha}^2 - \tilde{\beta}^2} + \ln\frac{K_1\left(\sqrt{\tilde{\alpha}^2 - \tilde{\beta}^2}\right)}{K_1\left(\sqrt{\tilde{\alpha}^2 - (\tilde{\beta} + \sigma_t\tilde{\delta})^2}\right)},$$

$$t \in \mathcal{T}\backslash\{0\}.$$

A.2. *NIG – GARCH*

The Normal Inverse Gaussian distribution is also a particular case of the GH distribution when $\lambda = -1/2$. As in the previous case, it is convenient to work with the same scale-location invariant parameters are $\tilde{\alpha} = \alpha\delta, \tilde{\beta} = \beta\delta$. Solving for the scale and the location parameters from (3.12) and (3.13) we obtain:

$$\tilde{\delta} = \frac{\left(\sqrt{\tilde{\alpha}^2 - \tilde{\beta}^2}\right)^{3/2}}{\tilde{\alpha}},$$

$$\tilde{\mu} = -\frac{\tilde{\beta}}{\tilde{\alpha}}\left(\sqrt{\tilde{\alpha}^2 - \tilde{\beta}^2}\right)^{1/2},$$

and therefore the probability density function of the innovation ε_t is given in this case by:

$$f_{\varepsilon_t}(x) = \frac{\tilde{\alpha}}{\pi\tilde{\delta}}\exp\left[\sqrt{\tilde{\alpha}^2 - \tilde{\beta}^2} + \tilde{\beta}\left(\frac{x - \tilde{\mu}}{\tilde{\delta}}\right)\right]\frac{K_1\left(\tilde{\alpha}\sqrt{1 + (\frac{x-\tilde{\mu}}{\tilde{\delta}})^2}\right)}{\sqrt{1 + (\frac{x-\tilde{\mu}}{\tilde{\delta}})^2}}.$$

We denote this by:

$$\varepsilon_t \sim \mathbf{NIG}(\tilde{\alpha}, \tilde{\beta}, \tilde{\delta}, \tilde{\mu}).$$

Corollary A.3. *Under $Q^{(e)}$ the return dynamics are given by*:

$$y_t = m_t + \sigma_t(\tilde{\mu} + \tilde{\delta}\vartheta_{1t}) + \sigma_t\tilde{\delta}\vartheta_{2t}\xi_t,$$

$$\xi_t|\mathcal{F}_{t-1} \sim \mathbf{NIG}\left(\tilde{\alpha}, \tilde{\beta}_{1t}, \frac{1}{\vartheta_{2t}}, -\frac{\vartheta_{1t}}{\vartheta_{2t}}\right),$$

where $\{\tilde{\beta}_{1t}\}$, $\{\vartheta_{1t}\}$, $\{\vartheta_{2t}\}$ *are some* $\{\mathcal{F}_t\}$*-predictable processes given by*:

$$\tilde{\beta}_{1t} = \tilde{\beta} + \theta_t^* \tilde{\delta}\sigma_t,$$

$$\vartheta_{1t} = \frac{\tilde{\beta}_{1t}}{\sqrt{\tilde{\alpha}^2 - \tilde{\beta}_{1t}^2}},$$

$$\vartheta_{2t} = \frac{\tilde{\alpha}}{\left(\sqrt{\tilde{\alpha}^2 - \tilde{\beta}_{1t}^2}\right)^{\frac{3}{2}}},$$

such that ξ_t *has zero conditional mean and unit conditional variance given* $\mathcal{F}_{t-1}$, *and* θ_t^* *is the unique predictable solution of the following martingale equation*:

$$\sqrt{\tilde{\alpha}^2 - (\tilde{\beta} + \tilde{\delta}\sigma_t + \theta_t^* \tilde{\delta}\sigma_t)^2} - \sqrt{\tilde{\alpha}^2 - (\tilde{\beta} + \theta_t^* \tilde{\delta}\sigma_t)^2} = \tilde{\mu}\sigma_t + m_t - r, \quad t \in \mathcal{T}\backslash\{0\}.$$

In this case we can determine an analytical form for the risk-neutral Esscher parameter θ_t^*:

$$\theta_t^* = -\frac{1}{2} - \frac{\tilde{\beta}}{\sigma_t \tilde{\delta}} - \frac{1}{2}\sqrt{\frac{(r - m_t - \sigma_t \tilde{\mu})^2}{\sigma_t^2 \tilde{\delta}^2}\left(\frac{4\tilde{\alpha}^2}{\sigma_t^2 \tilde{\delta}^2 + (r - m_t - \sigma_t\tilde{\mu})^2} - 1\right)}.$$

For some choices of $\tilde{\alpha}$ and $\tilde{\beta}$, (i.e. the ones making the return dynamics a normal GARCH process), the LRNVR is equivalent to the conditional Esscher transform. Indeed if we let $\tilde{\alpha} \to \infty$ and $\tilde{\beta} = 0$ the following relations will hold:

$$\begin{aligned} m_t^{Q^{(e)}} &= m_t + \sigma_t(\tilde{\mu} + \tilde{\delta}\vartheta_{1t}) \to m_t + \theta_t^* \sigma_t^2, \\ \sigma_t^{Q^{(e)}} &= \sigma_t \tilde{\delta}\vartheta_{2t} \to \sigma_t, \\ \theta_t^* &\to \frac{1}{\sigma_t^2}\left(r - m_t - \frac{\sigma_t^2}{2}\right), \\ \xi_t &\to \mathbf{N}(0,1), \end{aligned}$$

and, therefore, we obtain the same dynamics as Duan [19] for normal GARCH models:

$$y_t | \mathcal{F}_{t-1} \sim \mathbf{N}\left(r - \frac{1}{2}\sigma_t^2, \sigma_t^2\right).$$

Corollary A.4. *Under* $Q^{(m)}$ *the return process* $\{y_t\}$ *follows*:

$$\begin{aligned} y_t &= r - \tilde{\mu}\sigma_t - \sqrt{\tilde{\alpha}^2 - \tilde{\beta}^2} + \sqrt{\tilde{\alpha}^2 - (\tilde{\beta} + \sigma_t\tilde{\delta})^2} + \sigma_t \varepsilon_t', \\ \varepsilon_t' &\sim \mathbf{NIG}(\tilde{\alpha}, \tilde{\beta}, \tilde{\delta}, \tilde{\mu}), \\ \sigma_t^2 &= a_0 + a_1 \sigma_{t-1}^2 \varphi\left(\varepsilon_{t-1}' + \frac{m_{t-1}^{\text{shift}}}{\sigma_{t-1}}\right) + b_1 \sigma_{t-1}^2, \end{aligned}$$

where m_t^{shift} *is given by*:

$$m_t^{\text{shift}} = r - m_t - \tilde{\mu}\sigma_t - \sqrt{\tilde{\alpha}^2 - \tilde{\beta}^2} + \sqrt{\tilde{\alpha}^2 - (\tilde{\beta} + \sigma_t\tilde{\delta})^2}$$

Taking the limiting case that $\tilde{\alpha} \to \infty$ and $\tilde{\beta} = 0$, we have:

$$\begin{aligned} m_t^{Q^{(m)}} &= r - \tilde{\mu}\sigma_t - \sqrt{\tilde{\alpha}^2 - \tilde{\beta}^2} + \sqrt{\tilde{\alpha}^2 - (\tilde{\beta} + \sigma_t\tilde{\delta})^2} \to r - \frac{1}{2}\sigma_t^2 \\ \varepsilon_t' &\to \mathbf{N}(0,1) \end{aligned}$$

and thus, under $Q^{(m)}$ the returns are conditionally normal distributed:

$$y_t|\mathcal{F}_{t-1} \sim \mathbf{N}\left(r - \frac{1}{2}\sigma_t^2, \sigma_t^2\right)$$

ϑ which again corresponds to the same risk neutral return dynamics obtained by Duan [19].

A.3. *VG – GARCH*

The Variance Gamma distribution is a special limit case of the GH distribution when $\delta \to 0$. Unlike the previous two cases, in order to work with invariant type parameters, it is convenient to use the second parametrization. Thus, we denote by $\rho = \beta/\alpha$ and $\xi = \delta\gamma$, so the Variance Gamma density is obtained when $\xi \to 0$ and $|\beta| < \alpha$ for positive values of λ. Solving for the location and scale parameters from (3.12) and (3.13), the probability density function becomes:

$$f_{\varepsilon_t}(x) = \frac{(1-\rho^2)^\lambda}{\tilde{\alpha}\sqrt{\pi}2^{\lambda-1/2}\Gamma(\lambda)}\left(\frac{|x-\tilde{\mu}|}{\tilde{\alpha}}\right)^{\lambda-1/2}\exp\left(\rho\frac{(x-\tilde{\mu})}{\tilde{\alpha}}\right)K_{\lambda-1/2}\left(\frac{|x-\tilde{\mu}|}{\tilde{\alpha}}\right),$$

where $\tilde{\alpha}$ and $\tilde{\mu}$ are given by:

$$\begin{aligned} \tilde{\alpha} &= \left(\frac{2\lambda}{1-\rho^2} + \frac{4\lambda\rho^2}{(1-\rho^2)^2}\right)^{-\frac{1}{2}}, \\ \tilde{\mu} &= -\frac{2\tilde{\alpha}\lambda\rho}{1-\rho^2}. \end{aligned}$$

In this case we say that $\{\varepsilon_t\}$ are standard i.i.d. VG distributed as follows:

$$\varepsilon_t \sim \mathbf{VG}(\lambda, \rho, \tilde{\alpha}, \tilde{\mu}).$$

It is easy to check that the above representation is a location-scale family with the cumulant generating function given by:

$$\kappa_{\varepsilon_t}(u) = \tilde{\mu}u + \lambda\ln\frac{1-\rho^2}{1-(\rho+\tilde{\alpha}u)^2}$$

Corollary A.5. *Under $Q^{(e)}$ the return dynamics are given by:*

$$y_t = m_t + \sigma_t(\tilde{\mu} + \tilde{\alpha}\vartheta_{1t}) + \sigma_t\tilde{\alpha}\vartheta_{2t}\xi_t,$$

$$\xi_t|\mathcal{F}_{t-1} \sim \mathbf{VG}\left(\lambda, \rho_{1t}, \frac{1}{\vartheta_{2t}}, -\frac{\vartheta_{1t}}{\vartheta_{2t}}\right),$$

where $\{\tilde{\rho}_{1t}\}$, $\{\vartheta_{1t}\}$, and $\{\vartheta_{2t}\}$ are some $\{\mathcal{F}_t\}$-predictable processes given by:

$$\begin{aligned}
\tilde{\rho}_{1t} &= \tilde{\rho} + \theta_t^*\tilde{\alpha}\sigma_t,\\
\vartheta_{1t} &= \frac{2\tilde{\alpha}\lambda\rho_{1t}}{1-\rho_{1t}^2},\\
\vartheta_{2t} &= \left(\frac{2\lambda}{1-\rho_{1t}^2} + \frac{4\lambda\rho_{1t}^2}{(1-\rho_{1t}^2)^2}\right)^{\frac{1}{2}},
\end{aligned}$$

such that ξ_t has zero conditional mean and unit conditional variance given $\mathcal{F}_{t-1}$, and θ_t^* is the unique predictable solution of the following martingale equation:

$$\lambda \ln\left(\frac{1-(\rho+(1+\theta_t^*)\tilde{\alpha}\sigma_t)^2}{1-(\rho+\theta_t^*\tilde{\alpha}\sigma_t)^2}\right) = \tilde{\mu}\sigma_t + m_t - r, \quad t \in \mathcal{T}\backslash\{0\}.$$

Corollary A.6. *Under $Q^{(m)}$ the return process $\{y_t\}$ has the following form:*

$$\begin{aligned}
y_t &= r - \tilde{\mu}\sigma_t - \lambda\ln\left(\frac{1-\rho^2}{1-(\rho+\tilde{\alpha}\sigma_t)^2}\right) + \sigma_t\varepsilon_t',\\
\varepsilon_t' &\sim \mathbf{VG}(\lambda, \rho, \tilde{\alpha}, \tilde{\mu}),\\
\sigma_t^2 &= a_0 + a_1\sigma_{t-1}^2\varphi\left(\varepsilon_{t-1}' + \frac{m_{t-1}^{\text{shift}}}{\sigma_{t-1}}\right) + b_1\sigma_{t-1}^2,
\end{aligned}$$

where m_t^{shift} is given by:

$$m_t^{\text{shift}} = r - m_t - \tilde{\mu}\sigma_t - \lambda\ln\left(\frac{1-\rho^2}{1-(\rho+\tilde{\alpha}\sigma_t)^2}\right).$$

Acknowledgments

The authors would like to thank the editor and the two anonymous referees for their helpful comments and suggestions. Alexandru Badescu wishes to thank NSERC for their support. Robert J. Elliott would like to thank ARC and SSHRC for their continued support. Reg Kulperger would like to thank NSERC for its support. Tak Kuen Siu wishes to thank ARC for its support.

References

[1] K. Aas and I. H. Haff, The generalized hyperbolic skew student's t-distribution, *Journal of Financial Econometrics* **4**(2) (2006) 275–309.

[2] A. M. Badescu and R. J. Kulperger, GARCH option pricing: A semiparametric approach, *Insurance: Mathematics and Economics* **43**(1) (2008) 69–84.

[3] A. M. Badescu, R. J. Kulperger and E. Lazar, Option valuation with normal mixture GARCH models, *Studies in Nonlinear Dynamics & Econometrics* **12**(2) (2008) Article 5.
[4] R. T. Baillie and T. Bollerslev, The message in daily exchange rate: A conditional variance tale, *Journal of Business and Economics Statistics* **7** (1989) 297–305.
[5] O. E. Barndorff-Nielsen and N. Shepard, Non-gaussian ornstein-uhlenbeck-based models and some of their uses in financial economics, *Journal of the Royal Statistics Society B* **63** (2001) 167–241.
[6] G. Barone-Adesi, R. Engle and L. Mancini, A garch option pricing model in incomplete markets, *Review of Financial Studies* **21** (2008) 1223–1258.
[7] F. Black and M. Scholes, The pricing of options and corporate liabilities, *Journal of Political Economy* **81** (1973) 637–659.
[8] T. Bollerslev, Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* **31** (1986) 307–327.
[9] T. Bollerslev, A conditionally heteroskedastic time series model for speculative prices and rates of return, *The Review of Economics and Statistics* **69** (1987) 542–547.
[10] M. J. Brennan, The pricing of contingent claims in discrete time models, *Journal of Finance* **24**(1) (1979) 53–68.
[11] H. Bühlmann, F. Delbaen, P. Embrechts and A. N. Shiryaev, No-arbitrage, change of measure and conditional Esscher transforms, *CWI Quarterly* **9**(4) (1996) 291–317.
[12] C. Chorro, D. Guegan and F. Ielpo, Option pricing for GARCH-type models with generalized hyperbolic innovations, *Quantitative Finance*, in Press (2010).
[13] P. Christoffersen and K. Jacobs, Which GARCH model for option valuation? *Management Science* **50** (2004) 1204–1221.
[14] P. Christoffersen, S. L. Heston and K. Jacobs, Option valuation with conditional skewness, *Journal of Econometrics* **131** (2006) 253–284.
[15] P. Christoffersen, C. Dorion, K. Jacobs and Y. Wang, Volatility components: Ane restrictions and Non-normal innovations, forthcoming in the *Journal of Business and Economic Statistics* (2009).
[16] P. Christoffersen, R. Elkamhi, B. Feunou and K. Jacobs, Option valuation with conditional heteroskedasticity and non-normality, *Review of Financial Studies* 23(5) (2010), 2139–2183.
[17] P. K. Clark, A subordinated stochastic process model with finite variance for speculative price, *Econometrica* **41** (1973) 135–155.
[18] J. Cochrane, *Asset Pricing* (Princeton University Press, Princeton, NJ, 2001).
[19] J.-C. Duan, The GARCH option pricing model, *Mathematical Finance* **5** (1995) 13–32.
[20] J.-C. Duan, Conditionally fat-tailed distributions and the volatility smile in options, unpublished manuscript, University of Toronto (1999).
[21] J.-C. Duan, P. Ritchken and Z. Sun, Approximating GARCH-jump models, jump-diffusion processes, and option pricing, *Mathematical Finance* **16**(1) (2006) 21–52.
[22] E. Eberlein, Application of generalized hyperbolic Lévy motions to finance, in *Lévy Processes: Theory and Applications*, O. E. Barndorff-Nielsen, T. Mikosch and S. I. Resnick, eds. (Birkauser Verlag, 2001), pp. 319–337.
[23] E. Eberlein and E. A. V. Hammerstein, Generalized hyperbolic and inverse Gaussian distributions: Limiting cases and approximation of processes, *Proceedings of the 4th Ascona Conference* (Birkhäuser Verlag, 2003).
[24] R. J. Elliott and D. G. Madan, A discrete time equivalent martingale measure, *Mathematical Finance* **9**(6) (1998) 825–841.

[25] R. F. Engle, Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation, *Econometrica* **50** (1982) 987–1008.

[26] R. F. Engle and V. K. Ng, Measuring and testing the impact of news on volatility, *Journal of Finance* **48**(5) (1993) 1749–1778.

[27] L. Epstein and S. Zin, Substitution, risk aversion, and temporal behavior of consumption and asset returns: A theoretical framework, *Econometrica* **57** (1989) 937–969.

[28] R. Garcia, E. Ghysels and E. Renault, The econometrics of option pricing, in *Handbook of Financial Econometrics*, Y. Ait-Sahalia and L. P. Hansen eds. (Elsevier-North-Holland, Amsterdam, 2003).

[29] L. R. Glosten, R. Jagannathan and D. Runkle, On the relation between the expected value and the volatility for the nominal excess return on stocks, *Journal of Finance* **48** (1993) 1779–1801.

[30] H. U. Gerber and E. S. W. Shiu, Option pricing by Esscher transforms (with discussions), *Transactions of the Society of Actuaries* **46** (1994) 99–191.

[31] C. Gourieroux and A. Monfort, Econometric specifications of stochastic discount factor models, *Journal of Econometrics* **136** (2007) 509–530.

[32] L. P. Hansen and S. Richard, The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models, *Econometrica* **55** (1987) 587–613.

[33] W. Hardle and C. M. Hafner, Discrete time option pricing with exible volatility estimation, *Finance and Stochastics* **4** (2000) 189–207.

[34] M. J. Harrison and D. M. Kreps, Martingales and arbitrage in multiperiod security markets, *Journal of Economic Theory* **20** (1979) 381–408.

[35] S. L. Heston and S. Nandi, A closed-form GARCH option pricing model, *Review of Financial Studies* **13** (2000) 585–626.

[36] D. A. Hsieh, Modeling heteroskedasticity in daily foreign-exchange rates, *Journal of Business and Economic Statistics* **7** (1989) 307–317.

[37] K. Hsieh and P. Ritchken, An empirical comparison of GARCH option pricing models, *Review of Derivative Research* **8** (2005) 129–150.

[38] M. B. Jensen and A. Lunde, The NIG-S&ARCH model: A fat-tailed, stochastic, and autoregressive conditional heteroskedastic volatility model, *Econometrics Journal* **4** (2001) 319–342.

[39] M. Lanne and P. Saikkonen, Modeling conditional skewness in stock returns, *The European Journal of Finance* **13**(8) (2007) 691–704.

[40] R. Lucas, Asset pricing in an exchange economy, *Econometrica* **46** (1978) 1429–1445.

[41] C. Menn and S. T. Rachev, A GARCH option pricing model with α — stable innovations, *European Journal of Operational Research* **163** (2005) 201–209.

[42] R. C. Merton, Rational theory of option pricing, *Bell Journals of Economics and Management Science* **4** (1973) 41–83.

[43] A. R. Pagan and G. W. Schwert, Alternative models for conditional stock volatility, *Journal of Econometrics* **45** (1990) 267–290.

[44] K. Prause, The generalized hyperbolic model: Estimation, financial derivatives, and risk measures. Dissertation zur Erlangung des Doktogrades der Mathematischen Fakultat der Albert-Ludwigs-Universitat Freiburg i. Br (1999).

[45] B. Rosenberg, The behavior of random variables with nonstationary variance and the distribution of security prices (reprinted as pages 83–108 in Shephard, 2005) (1972).

[46] M. Rubinstein, The valuation of uncertain income streams and the pricing of options, *Bell Journal of Economics* **7** (1976) 407–425.

[47] R. Schmidt, T. Hrycej and E. Stutzle, Multivariate distribution models with generalized hyperbolic margins, *Computational Statistics and Data Analysis* **50**(8) (2006) 2065–2096.
[48] W. Schoutens, *Lévy Processes in Finance: Pricing Financial Derivatives* (John Wiley and Sons, 2003).
[49] M. Schroder, Risk-neutral parameter shifts and derivatives pricing in discrete time, *Journal of Finance* **59**(5) (2004) 2375–2402.
[50] T. K. Siu, H. Tong and H. Yang, On pricing derivatives under GARCH models: A dynamic Gerber-Shiu approach, *North American Actuarial Journal* **8**(3) (2004) 17–32.
[51] T. K. Siu and H. Yang, Option pricing when the regime-switching risk is priced, *ACTA Mathematicae Applicatae Sinica (English Series)* **25**(3) (2009) 369–388.
[52] L. Stentoft, American option pricing using GARCH models and the Normal inverse gaussian distribution, *Journal of Financial Econometrics* **6** (2008) 540–582.
[53] G. Tauchen and M. Pitts, The price variability-volume relationship on speculative markets, *Econometrica* **51** (1983) 485–505.
[54] S. J. Taylor, Financial returns modelled by the product of two stochastic processes: A study of daily sugar prices, 1961–1975, in *Time Series Analysis: Theory and Practice*, O. D. Anderson (ed.), Vol. 1 (Amsterdam, North-Holland, 1982), pp. 203–226.
[55] S. J. Taylor, *Modelling Financial Time Series* (New York: John Wiley and Sons, 1986).
[56] S. J. Taylor, *Asset Price Dynamics and Prediction* (Princeton: Princeton University Press, 2005).
[57] S. J. Taylor, *Modelling Financial Time Series*, 2nd edn. (Singapore: World Scientific, 2008).
[58] S. S. Wang, A class of distortion operators for pricing financial and insurance risks, *Journal of Risk and Insurance* **67**(1) (2000) 15–36.
[59] S. S. Wang, Equilibrium pricing transforms: New results on Bühlmann's 1980 economic model, *ASTIN Bulletin* **32** (2003) 213–234.