Università
Ca'Foscari
Venezia

Corso di Dottorato di ricerca

in : Economia

ciclo : **P H D, T H E S I S**

Tesi di Ricerca

in cotutela con **University of Paris 1 Pantheon-Sorbonne**

Integration of VIX information in GARCH option pricing models

SSD: SECS – P/05

Coordinatore del Dottorato :
ch. Maître de conférences HDR. Christophe CHORRO

Supervisore :
ch. Maître de conférences HDR. Christophe CHORRO
Université Paris 1 Panthéon-Sorbonne

Supervisore cotutela:
ch. Professor. Monica BILLIO
Università Ca' Foscari Venezia

Dottorandoo :
Fanirisoa Zazaravaka RAHANTAMIALISOA
Matricola : 956107

Acknowledgments

I thank God for giving me the wisdom and courage and for guiding me throughout my life for without Him I would not have come this far. I would also like to acknowledge the following for their contributions which facilitated the completion of this project.

I would like to express profound gratitude to Professor Christophe Chorro and Professor Monica Billio, my advisors, for providing unlimited, invaluable and active guidance throughout the study. Professor Christophe Chorro immense command and knowledge of the subject matter enabled me to shape this research project to the product that it is now.

I also thank my family for letting me steal their valuable time to work on this project. It is my hope that their sacrifice has finally paid off.

Finally, I owe my gratitude to a number of people who in one way or another contributed towards completion of this project especially my fellow colleagues at work.

Contents

Chapter 1

Introduction

Contents

The theory of option pricing is a topic of great interest in financial literature. Investors and traders face difficulties attempting to obtain the correct value of an option. Empirical research demonstrates that future prices changes are difficult to predict using mathematical models. However, option theory has improved in leaps and bounds since 1972, when Black and Scholes published the first analytical approach to pricing European options ([Black & Scholes (1973)]).

An option[1] entitles the holder to buy or sell a specified amount of an underlying asset at a set price before or on the expiration date of the option[2] ([Markham & Sharpe (2002)]).

[1]In 1973 the Chicago Board of Trade (CBOT) began to trade what are known in the trade as so-called options ([Markham & Sharpe (2002)]), futures and other financial derivatives.

[2]European options may only be exercised on the expiration date, while American options can be exercised at any time up to the expiration date. European call and put options are widely referred to as plain vanilla options because they are so simple, with the more advanced options commonly known as exotic.

Because it is viewed as a right rather than an obligation, the holder is able to choose not to exercise his right and let the option expire. There are two types of options: call options and put options. The former is an agreement that enables the holder to buy a bond, stock, or commodity (defined as an underlying asset) for a specific price on a specific date, whereas the latter enables them to sell the underlying asset for a specific price on a specific date. An option value is defined by several variables linked to the underlying asset and the financial markets ([Hull (2011)], pp. 214-218) such as:

- The Current Value (**S**) of the Underlying Asset : Options are assets which derive their value from an underlying asset. Consequently, the value of the asset is influenced by changes in the value of the underlying asset. As calls entitle you to buy the underlying asset at a set price, there is generally an increase in the call option price and a decrease in that of the put option as the underlying asset's price increases.

- The Strike Price (**K**) of Option: an options's price generally increases as the option gets closer to being ITM (in-the-money). This is because the strike price becomes incresingly favorable in regard to the current price of the underlying asset. Likewise, an option's price decreases as the option moves towards OTM (out-of-the-money), as the strike price is less favorable in regard to the price of the underlying asset.

- The time left before expiration (**T**) on the option: the longer an option has before the expiration date, the greater the odds are of it becoming profitable (in-the-money). This is because the more time we have before expiration, the more time there is for the value of the underlying asset to change, which increases the call and the put options' value. Moreover, the time value is affected by the volatility of the underlying asset because when you have a volatile underlying asset, you can expect plentiful price movements.

- The Volatility (σ) of the underlying asset: volatility is a measure of the uncertainty concerning future asset price movements. A greater level of volatility suggests the asset value is able to reach a greater range of values, while a lower volatility suggests the asset value is only subject to minor fluctuations. Greater variance in the value of the underlying asset will increase make the option value. While it seems counter-intuitive that an increase in a risk measure (variance) should increase value, options differ from other securities because options' buyers are not able to lose more than the price they initially; in fact, there is the potential for good yields from large price movements.

- Dividends Paid (**d**) on the Underlying Asset: the underlying asset's value is expected to go down if there are dividend payments made on the asset during the option's life. Therefore, the value of a call on the asset is a decreasing function of the size of projected dividend payments, and the value of a put is an increasing function of projected dividend payments, whereas a put option's price is an increasing function of expected dividend payments.

- The interest rate (**r**): because the option buyer has to pay the up-front option price, this entails an opportunity cost. This cost will change depending on the level of interest rates and the option's expiry date. Additionally, the risk-free interest rate is also involved when valuing options while calculating the current value of the exercise price. This is because the exercise price need not be paid (received) before the expiration of call (put) options.

The financial literature seems to believe that the option pricing theory was started by the seminal research of Black, Scholes and Merton ([Black & Scholes (1973)]). Here they suggest a mathematical option pricing model which specifies the fair market value of European options, taking in to account the probability of constant volatility. The model was based on the copy of a portfolio, made up of underlying assets and risk-free assets with equal cash flows. The model also argues that stock prices follow a log-normal distribution as a result of several factors. Firstly, asset prices cannot be negative and there are no transaction costs or taxes. Secondly, information is freely available to everyone. Finally, there is a consistent risk-free interest rate for all terms and market participants can both borrow and lend at this rate. Short selling of securities using the proceeds is allowed in the absence of risk-free arbitrage opportunities. Stocks do not pay dividends; there is a consistent variance in yield over the term of the option contract which is known to market participants. The Black-Scholes model is most frequently used as an easy and relatively effective approximation.

However, it has long been criticized due to its over-simplified and over-realistic model assumptions about European call option prices. As the financial literature shows us, many assumptions in the Black-Scholes model could prove incorrect in reality, for instance the following:

- The log-returns' normality assumption[3]. Empirical studies demonstrate that logarithmic yields have empirical distributions which are leptokurtic regarding the normal distribution and which are, in various cases, skewed (e.g., [Mandelbrot (1963)], [Fama (1965)]). This was the case for [Bollerslev (1986)] who found leptokurtosis in the monthly S&P500 returns, while [French *et al.* (1987)] reported skewness in daily S&P500 returns. [Engle & Gonzalez-Rivera (1991)] discovers excessive skewness and kurtosis in small stocks. Additionally , recent research confirms skewed distribution of log-returns, with a peak around the mean distribution and the heavy tail (see [Bollen & Inder (2002)], [Carr *et al.* (2002)]).

- The same cannot be said about homoskedastic volatility, as volatility fluctuates with the level of supply and demand. Thus the theoretical values can often be inaccurate (see [Mandelbrot (1963)]). The existing literature intends to model and predict financial volatility, and can be separated into two clear groups: parametric and non-parametric models. The former group assumes a specific functional form for volatility, and models it according to the function of observable variables, such as ARCH or GARCH models, [Engle (1982)]; [Bollerslev (1986)] while in the latter financial volatility is determined without imposing any parametric assumptions, which is the reason they known as "realized volatility models" [Andersen *et al.*(2003)]. While multiple different approaches

[3] Another crucial point of the Black-Scholes framework is its requirement that continuous trading is possible; it has a tendency to overvalue deep out-of-the-money calls and undervalue deep in-the-money calls.

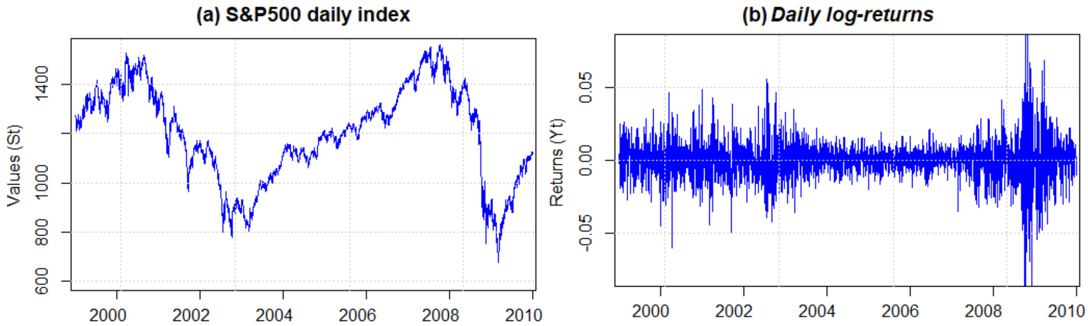
exist in the literature for volatility modeling, we currently possess no single model which explains all the stylized facts simultaneously, even with long memory modelling.

In order to clarify the ideas of leptokurtism, leverage effect, non-normality, and clustering phenomena, we will be gathering empirical evidence. We will firstly review the S&P500 daily closing indices[4] from January 1999 to December 2010. Then, we will be defining a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$. This aims to model the uncertainty where $\mathbb{P}$ can describe the physical distribution of the nature states, and $\{\mathcal{F}_t\}_{t\in\{0,\cdots,T\}}$ takes a form of information filtration. Therefore this will represent the resolution of uncertainty based on the information that has been generated by the market prices up to and including time t. We suppose that $\mathcal{F}_0 = \sigma\{\emptyset, \Omega\}$ and $\mathcal{F}_T = \mathcal{F}$. Instead of using the asset price, we will use the return. In practical analysis, the return is generally described as a log price change, similar to the relative price change. Let's denote by S_t the price of the S&P500 at time t and the price S_{t-1} at $t-1$ by considering the logarithm difference of the two consecutive prices:

$$Y_t = \log(S_t) - \log(S_{t-1}). \tag{1.1}$$

It should be mentioned that a log return is the logarithm of a gross return $\frac{S_t}{S_{t-1}}$ and $\log(S_t)$ is defined as the log price.

Figure 1.1: Time series plots of the daily prices S_t, the daily log returns Y_t, of the S&P500 in the period of January 1999 - December 2010.



The first plot (a) in Figure 1.1 is the time series plot for the daily closing indexes of S&P500. During the "dot-com bubble", the index reached an all-time high on March 24, 2000, before consequently losing roughly 50% of its value in the stock market decline of 2002. On October 9, 2007, it again attained a historic high before suffering from the subprime mortgage credit crisis between 2008-2010. The second panel (b) in Figure 1.1 shows the index's daily log returns, the daily yield curve with more volatile fluctuations. In particular, the high volatilities during the 2008-2010 are better demonstrated in the daily return plot. Returns, unlike prices, vary around a constant level close to 0, with high oscillations tending to cluster, which reflects more volatile market periods.

[4]The S&P500 is a value-weighted index of the prices of the 500 large-cap common stocks actively traded in the United States.

Figure 1.2: Histograms and Q-Q plots of the daily, log returns of the S&P500 in the period of January 1999 - December 2010. The normal density with the same mean and variance are superimposed on the histogram plots.

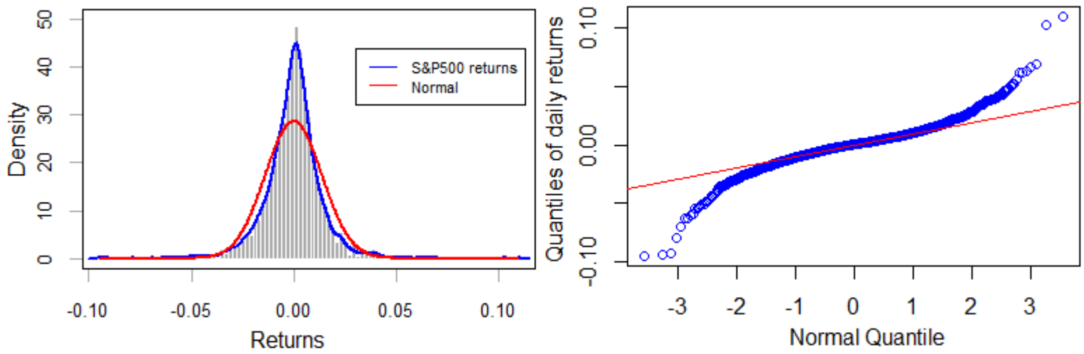


Figure 1.2 displays the probability density function for the daily returns on the S&P500 index since January 1999, followed by a zero mean and a standard deviation of 1. Added to this is the normal density function with the same mean and variance. Furthermore, in this diagram the $Q - Q$ (quantile-quantile) plots for the returns are shown. The empirical peak is greater than the normal distribution and the tails are both thicker.

Figure 1.3: (a) Time series plot of VIX (blue) and (b) Annualised Historical Volatility using $\left[\frac{1}{n-1}\sum_{k=1}^{n}(Y_t-\mu)^2\right]^{\frac{1}{2}}$ of the S&P500 in the period of January 1999 - December 2010.

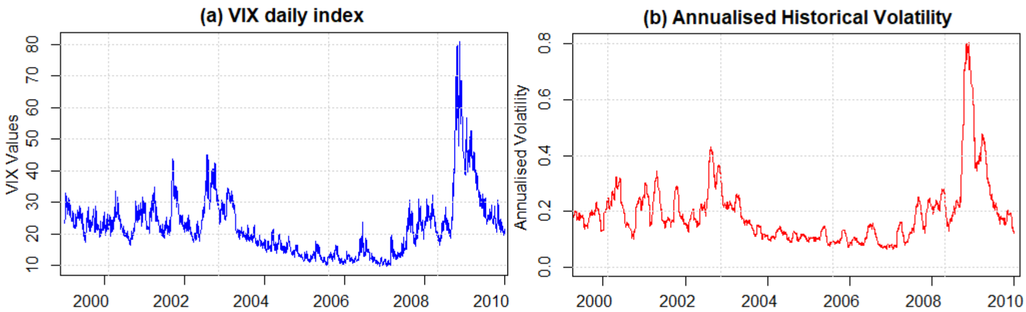


Figure 1.3 displays the VIX time series and Annualized Historical Volatility of the S&P500 from January 1999 to December 2010. The Annualized Historical Volatility is calculated using the expression $\left[\frac{1}{n-1}\sum_{k=1}^{n}(Y_t-\mu)^2\right]^{\frac{1}{2}}$ where μ is the mean of Y_t. The VIX which is a proxy of the implied volatility was used in order to examine the leverage effect:

- Stationarity: The prices of an asset recorded over times are rarely static. However, their returns generally fluctuate around a constant level, suggesting a constant mean over time. See Figure 1.1, in fact, with the majority of return sequences, we can model them as a stochastic process with at least time-invariant first two moments.

- Asymmetry: The distribution of return Y_t is frequently negatively asymmetrical (see Figure 1.2), which reflects the fact that, in financial markets, the downturns are often much steeper than the recoveries. The skewness measures[5] the degree of asymmetry of a distribution around its mean. It is a pure number that characterises the shape of the distribution. Negative skewness indicates a distribution with an asymmetric tail extending toward more negative values
- Volatility clustering: this term means that large price changes (i.e. returns with high absolute values) occur in clusters. See Figures 1.1 and 1.3. Indeed, great changes in price are generally followed by large movements in price, and periods of tranquility and high volatility periods alternate.
- Heavy tails: the probability distribution of return Y_t generally possesses heavier tails than those of a normal distribution. Figure 1.2 provides the quantile-quantile plot for normality graphical checking.
- Leverage effect: Asset returns are negatively correlated with volatility changes. [Black & Scholes (1976)]. When asset prices fall, companies become highly leveraged and riskier. And so, the volatility of their stock prices increases. On the other hand, when the volatility of stock prices increases, investors demand high returns, allowing the stock prices to decrease. Volatilities triggered by price decline are generally larger than the gains from declined volatilities. The leverage effect is powerful, even if VIX is an imperfect measure of the volatility of the S&P500 index, involving the volatility risk premium (see [Aït-Sahalia *et al.*(2013)]).

Although the fundamental mathematical model has several limitations, the Black-Scholes methodology was a critical first step for option pricing. A large body of work has sought to deal with the imperfections of the Black-Scholes model. For example, [Rubinstein (1976)] and [Brennan (1979)] gave us the basis of the discrete time approach. Because the Black-Scholes model was developed for the continuous pricing of European style options, a simpler model known as Cox-Rubinstein binomial model presents a discrete pricing model. This model, created in 1979, is widely referred to as the Binomial Option Pricing Model or the Binomial Model. However, it was quickly apparent that the binomial model is more applicable as a pricing model for American Style Options. It is centred around a basic formulation for the asset price process in which the asset can move to one of two possible prices at any time. Multiple differing models have since been developed, involving increasingly complex volatility models. However, it's simplicity ensures that, despite its limitations, the original Black-Scholes model is still the most widely used by the options traders of today.

1.1 GARCH framework

A growing body of work on time series analysis is emerging focused on the difficulties of modelling volatility as input in option pricing.

[5] Although the mean, standard deviation, and average deviation are dimensional quantities, the skewness is generally defined in such a way as to make it non-dimensional. The kurtosis is also a nondimensional quantity.

1.1.1 GARCH-in-mean

[Engle (1982)] laid out an early theoretical attempt to show volatility as a time varying process with the conditional variance and endogenous parametric specification with the class of autoregressive conditional heteroskedasticity models (ARCH). He modelled the conditional variance as a linear function of p lagged squared of squared returns Y_t^2 to reach a direct estimation of the joint dynamics of the volatility and returns. This model is widely known as the $ARCH(p)$ model and is defined as:

Définition 1 *The process $(Y_t)_{t\in\{1,\cdots,T\}}$ called the $ARCH(p)$ process is of the form:*

$$\begin{cases} Y_t &= \sqrt{h_t} z_t \\ h_t &= a_0 + a_1 Y_{t-1}^2 + \cdots + a_p Y_{t-p}^2 \end{cases} \tag{1.2}$$

where $a_i > 0$ for all i, the $(z_t)_{t\in\{1,\cdots,T\}}$ are i.i.d random variables[6] *with mean $\mathbb{E}[z_t] = 0$ and variance $Var[z_t] = 1$.*

Under zero covariance and zero mean, the process is covariance stationary if and only if the sum of the positive autoregressive parameters is less than one $\sum_{i=1}^p a_i < 1$, in which case the unconditional variance equals

$$Var[Y_t] = \frac{a_0}{1 - a_1 - a_2 - \cdots - a_p}.$$

Furthermore, the equation 1.2, identifies how the conditional variance h_t is determined by the available information $\mathcal{F}_{t-1}$. h_t possesses the property of time-varying conditional variance, which means it can capture the volatility clustering. [Bollerslev (1986)] offered us a broader picture of the ARCH model to ensure greater realism. His celebrated[7] Generalized ARCH model, models the conditional heteroskedasticity. The GARCH models are discrete-time and parametric models tasked with tracking correlation and volatility changes over time, in a manner where the conditional variance is also a function of its own lags of all order up to q.

Définition 2 *The process $(Y_t)_{t\in\{1,\cdots,T\}}$ is a generalized autoregressive conditional heteroskedasticity process* GARCH$(p;q)$ *of order $(p,q) \in (\mathbb{N}^*)^2$ if :*

$$\begin{cases} Y_t &= \sqrt{h_t} z_t \\ h_t &= a_0 + \sum_{i=1}^p a_i Y_{t-i}^2 + \sum_{j=1}^q b_i h_{t-j} \end{cases} \tag{1.3}$$

where the $(z_t)_{t\in\{1,\cdots,T\}}$ are i.i.d random variable with $\mathbb{E}[z_t] = 0$, $Var[z_t] = 1$ and $(a_i)_{i\in\{1,\cdots,p\}}$, $(b_i)_{j\in\{1,\cdots,q\}}$ are non negative constants such that $a_0 > 0$.

To put it differently, the conditional variance is a function of past returns and past conditional variances which gives us a predictable measure. The lagged conditional variance h_{t-j} present in the system causes volatility clustering. The non-negativity conditions of the

[6] In particular, if $\forall t \in t \in \{1,\cdots,T\}$, $z_t \sim N(0,h_t)$ then $Y_t \mid_{\mathcal{F}_{t-1}} \sim N(0,h_t)$ given information set available at time $t-1$.

[7] [Bollerslev *et al.* (1992)] already listed a variety of applications of these models in their survey.

coefficients ensure that h_t is strictly positive. The process $(Y_t)_{t\in\{1,\cdots,T\}}$ in equation 1.3 which is a GARCH$(p;q)$ only gives a strictly stationary solution with finite variance when

$$\sum_{i=1}^{p} a_i + \sum_{j=1}^{q} b_i < 1$$

Moreover, this strictly stationary solution is also unique, to ensure the existence of the unconditional variance. To obtain this constant unconditional variance, we observe that:

$$Var\left[Y_t\right] = \mathbb{E}\left[Y_t^2\right] = \mathbb{E}\left[h_t\right] = a_0 + \mathbb{E}\left[\sum_{i=1}^{p} a_i Y_{t-i}^2 + \sum_{j=1}^{q} b_i h_{t-j}\right],$$

$$Var\left[Y_t\right] = a_0 + \sum_{i=1}^{p} a_i \mathbb{E}\left[Y_{t-i}^2\right] + \sum_{j=1}^{q} b_i \mathbb{E}\left[h_{t-j}\right] = a_0 + Var\left[Y_t\right]\left(\sum_{i=1}^{p} a_i + \sum_{j=1}^{q} b_i\right)$$

$$Var\left[Y_t\right] = \frac{a_0}{1 - \sum_{i=1}^{p} a_i - \sum_{j=1}^{q} b_i}$$

It is worth noticing that, for $p = q = 0$, the model is reduced to the standard homoskedastic log normal process, assumed in the Black model. GARCH is probably the most common financial time series model used and has been followed by a large number of models based on more sophisticated models [Bera & Higgins (1993)], [Berkes *et al.* (2003)] and [Giraitis *et al.* (2005)].

This time-varying volatility structure in equation 1.3 is both compatible with the volatility clustering effect, and contains fat tails from the volatility data and leptokurtosis in series. The GARCH process's volatility modelling ability was documented by [Hansen & Lunde (2005)]. Despite comparing more than 300 time series models they were unable to find conclusive evidence that any of them outperform the GARCH. Both the capabilities and the limitations of GARCH models with regard to option pricing have recently been discussed at some length (see [Hardle & Hafner (2000)], [Christoffersen *et al.* (2004)], [Chorro *et al.* (2015)]) :

- GARCH models make the assumption that the magnitude alone, not the positivity or the negativity of unanticipated excess returns, determines h_t. Under[8] $\mathbb{E}\left[z_t^3\right] = 0$, the change in variance tomorrow and excess returns today are conditionally uncorrelated:

$$h_t = a_0 + \sum_{i=1}^{p} a_i h_{t-i} z_{t-i}^2 + \sum_{j=1}^{q} b_i h_{t-j}$$

  writing h_t as a function of lagged h_t and lagged z_t^2 where $Y_t^2 = h_t z_t$, it becomes clear that the conditional variance is invariant to changes in sign of the innovation z_t.

- Another frequently seen phenomenon in financial time series is known as the leverage effect. This happens when there is a negative correlation between changes in stock prices and changes in volatility. GARCH models capture volatility clustering and leptokurtosis. However, when assuming a symmetric distribution, it is impossible for them

[8]The distribution z_t is symmetric.

to account for the leverage effect. The leverage effect denotes the negative correlation is found between the asset return innovations and volatility innovations, this required the development of new and extended models over GARCH that resulted in new models. Particularly with the GARCH $(1,1)$, it can be described by the quantity:

$$Cov\left[Y_t - Y_{t-1}, h_{t+1} - h_t \mid \mathcal{F}_{t-1}\right] = Cov\left[Y_t, h_{t+1} \mid \mathcal{F}_{t-1}\right] = a_1 h_t^{\frac{3}{2}} \mathbb{E}\left[z_t^3\right]$$

measuring the impact of current variations of the log-returns Y_t on future variations of the conditional variance h_t. The leverage is equal to zero, if z_t is symmetric and compatible with empirical observations if $\mathbb{E}\left[z_t^3\right] < 0$.

As the GARCH $(1,1)$ model and the GARCH (p,q) model frequently perform as well as each other, we will focus in on the case when $p = q = 1$. It should be noted that, [Hansen & Lunde (2005)] provided convincing evidence finding a volatility model that performs better than the simple GARCH $(1,1)$ is a difficult task. Particularly the GARCH $(1,1)$ model solely comprised of three parameters in the conditional variance equation is adequate for capturing the volatility clustering. We focus on the GARCH-in-mean which includes an additional term m_t in the conditional mean equation. This model simultaneously characterizes the variance of a time series and the mean's evolution. Time varying conditional expectation is a crucial aspect of GARCH-in-mean models. Financial modeling and econometric study rely heavily upon the GARCH-in-mean model, for instance [Duan (1995)], [Chorro *et al.* (2015)] and [Christoffersen *et al.* (2012)], just to list a few. We have chosen this model due to its simultaneous characterisation of the evolution of the mean and the variance of a time series.

Furthermore, this particular model capable of explaining the excessive return (risk premium[9]) which is unable to be explained through traditional GARCH models as the condition expectation $\mathbb{E}\left[Y_t\right]$ remains at zero throughout the timeframe. The conditional mean specification is able to take different forms in practice and $h_t = F\left(z_{t-1}, h_{t-1}, \theta^V\right)$ can be characterized by the conditional distribution of the innovations process z_t with the parameters θ^D, and the structure of the conditional variance of the log-returns. The function $F(\cdot)$, known as the news impact curve, describes the impact of random shock of return z_t on the conditional variance h_t.

Définition 3 $(Y_t)_{t\in\{1,\cdots,T\}}$ *follows a general* GARCH-*in-mean order* $(p,q) \in (\mathbb{N}^*)^2$ *if :*

$$\begin{cases} Y_t &= r + m_t + \sqrt{h_t} z_t \\ h_t &= F\left(z_{t-1}, h_{t-1}, \theta^V\right) \end{cases} \tag{1.4}$$

where the $(z_t)_{t\in\{1,\cdots,T\}}$ *are i.i.d random variable,* $\mathbb{E}\left[z_t\right] = 0$*,* $Var\left[z_t\right] = 1$ *and* θ^V *is the set of parameters associated to the volatility.*

[9] In the financial market, [Engle *et al.* (1987)] pointed out: as the degree of uncertainty in asset returns varying over time, the compensation required by risk averse economic agents for holding these assets, must also be varying.

We take conditional mean[10] return m_t to be an $\mathcal{F}_{t-1}$-predictable process with respect to the information filtraton. Depending on which function $F(.)$ is chosen, and the distribution of z_t, the previous period innovation has differing effects on the current variance h_t. In this manner, it is possible for there to be multiple different extensions of GARCH models. The choice of the mean m_t effects the manner in which conditional return is dependent on the conditional volatility.

For instance, for the standard GARCH first suggested by [Bollerslev (1986)], z_{t-1} has a symmetric effect on the conditional variance with

$$h_t = F\left(z_{t-1}, h_{t-1}, \theta^V\right) = a_0 + a_1 h_{t-1} z_{t-1}^2 + b_1 h_{t-1}$$

where $\theta^V = (a_0, a_1, b_1)$, $a_0 > 0$, a_1 and b_1 are nonnegative constants, these parameters ensure the strict positivity of h_t. This basic GARCH model often acts as a fairly effective system for the analysis of financial time series and estimation of conditional volatility. Nevertheless, a few characteristics are unable to be captured by the standard GARCH model. We will be looking at different GARCH specifications HN-GARCH by [Heston & Nandi (2000)], GJR-GARCH by [Glosten *et al.* (1993)], NGARCH and IG-GARCH by [Christoffersen *et al.* (2013)], these models offer superior flexibility to the standard GARCH model.

1.1.2 Asymmetric GARCH Models

Over time, a vast body of work has been geared towards discussing the performance of option pricing for different GARCH specifications. An alternative issue was proposed by [Heston (1993)], who used stochastic volatility models which need information from the volatility structure to estimate the parameters of the model. The most celebrated theoretical framework based on GARCH processes for option pricing model was proposed by [Duan (1995)]. [Heston & Nandi (2000)] developed HN-GARCH model to capture the leptokurtosis exhibited by financial returns. That model generates a closed-form solution for European options; generally, under the GARCH framework, there are no closed-form solutions to obtain the no-arbitrage price. Regrettably, previous studies suggest that GARCH models seldom fully capture the thick tails property of the conditional distribution (for instance [Chorro *et al.* (2015)]). The introduction of more flexible distribution specifications has enabled this constraint to be overcome, which brings about the use of non normal distributions to model this excess kurtosis more accurately.

In recent years, a new set of GARCH frameworks are able to capture leverage effects, asymmetry, time varying skewness and kurtosis. In this set of GARCH models, [Nelson (1991)] with the exponential EGARCH, the GJR-GARCH [Glosten *et al.* (1993)], [Zakoian (1994)] with the threshold GARCH (TGARCH) and the asymmetric affine-GARCH by [Heston & Nandi (2000)] we can see GARCH models which take the asymmetric variance

[10] In many studies, m_t is assumed to be a function of the conditional variance h_t of the return at time t.

effects into account. It is established through these frameworks that asymmetric GARCH models perform better than classic GARCH. A new type of ARCH model known as the Asymmetric Power ARCH model (APARCH) was introduced by [Ding *et al.* (1993)], this model allows the estimation of the optimal power term.

The exponential GARCH model introduced by [Nelson (1991)] incorporates the leverage effect and specifies the conditional variance in a logarithmic form.

Définition 4 *An* EGARCH *model is expressed as:*

$$\log(h_t) = \log\left(F\left(z_{t-1}, h_{t-1}, \theta^V\right)\right) = a_0 + a_1\left(\mid z_{t-1} \mid - \gamma z_{t-1}\right) + b_1 \log(h_{t-1}) \tag{1.5}$$

where $\theta^V = (a_0, a_1, b_1, \gamma)$*, and* λ *are real parameters.*

It is crucial to grasp that the volatility dynamics have a multiplicative form instead of a linear one:

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

And, due to the presence of the exponential function, we do not have any restrictions to the parameters to guarantee that the conditional variance is positive. Furthermore, if $z_{t-1} > 0$, which corresponds to good news, the total effect of z_{t-1} is $a_1(1-\gamma)$, if $z_{t-1} < 0$ which corresponds to bad news, the total effect of z_{t-1} is $-a_1(1+\gamma)$. Thus, in case $\gamma \neq 0$ the volatility reacts asymmetrically to the rising and falling of stock prices and when $\mid \gamma \mid < 1$ this asymmetry will be compatible with empirical leverage effects.

The GJR-GARCH model, introduced by [Glosten *et al.* (1993)], is another iteration of an asymmetric GARCH model, which takes into account the dependence of a coefficient of the volatility structure for one particular event:

Définition 5 *The* GJR-GARCH *models are defined by:*

$$F\left(z_{t-1}, h_{t-1}, \theta^V\right) = a_0 + h_{t-1}\left(a_1 + \gamma \mathbb{1}_{\{z_{t-1}<0\}}\right) z_{t-1}^2 + b_1 h_{t-1} \tag{1.6}$$

where $\theta^V = (a_0, a_1, b_1, \gamma)$ *with* $a_0 > 0$*,* (γ, a_1, b_1) *are nonnegative, and* $\mathbb{1}_{\{z_{t-1}<0\}}$ *is the indicator function of the event* $\{z_{t-1} < 0\}$*.*

When the distribution of z_t is symmetric, second order stationarity condition requires $a_1 + b_1 + \frac{\gamma}{2} < 1$. Dependant on whether z_{t-1} is above or below 0, z_{t-1}^2 will have a different effect on the conditional variance h_t. If there is bad news $z_{t-1} < 0$, then $\mathbb{1}_{\{z_{t-1}<0\}} = 1$ and the complete effect on next period of conditional variance is $(a_1 + \gamma) z_{t-1}^2$. In this model, bad news will also have a larger impact on the conditional variance. If $\gamma > 0$, the leverage effect exhibits and suggests that negative shocks will have a larger impact on conditional variance than positive shocks. The TGARCH (Threshold GARCH) created by [Zakoian (1994)] is a relatively comparable version of the GJR-GARCH model where the volatility dynamics are specified in terms of conditional standard deviation instead of conditional variance.

The nonlinear NGARCH model brought in by [Engle & Ng (1993)] allows for asymmetric behavior in the volatility so that good news i.e. positive returns yield a subsequent decrease in volatility, while bad news or negative returns yields a subsequently higher volatility.

Définition 6 *The* NGARCH *model allows for asymmetric behavior in the volatility by setting :*

$$F\left(z_{t-1}, h_{t-1}, \theta^V\right) = a_0 + a_1 h_{t-1} \left(z_{t-1} - \gamma\right)^2 + b_1 h_{t-1} \quad (1.7)$$

where $\gamma > 0$, $\theta^V = (a_0, a_1, b_1, \gamma)$, $a_0 > 0$, a_1 *and* b_1 *are nonnegative constants.*

The essential point here is that this model takes the negative news into account $z_{t-1} < 0$, which yields a greater impact on variance than positive news $z_{t-1} > 0$ provided $\gamma > 0$. The persistence of variance in this model is $\Psi = a_1\left(1+\gamma^2\right) + b_1$ and the long-run unconditional variance is $h_0 = \dfrac{a_0}{1-\Psi}$.

Another problem with using GARCH models is that they have difficulty fully embracing the thick tails property of the asset returns Y_t. This could be related to the asymmetry characteristic of error term distribution. The asymmetry may lead to skewed returns. We can attribute the asymmetries seen in the implied volatility smile to the skewness of the underlying asset returns. It has been demonstrated that a high level of excess kurtosis is seen in asset returns, which results in a larger peak than the curvature found in the Gaussian distribution. [Byun & Cho (2013)] used reams of data on S&P500 index options to carry out a comparison of the empirical performances of multiple GARCH option pricing models with non-normal innovations, providing us with evidence that stocks and indices usually have negative skewness. Simply put this is because the decline rate of stock prices tends to be higher than the growth rate.

It has been demonstrated that GARCH models do not only capture volatility clustering, but can also accommodate some of the leptokurtosis in thick tails. However, GARCH models with conditionally normal errors rarely succeed to adequately capture the leptokurtosis which manifests in asset returns (To list a few well known articles: [Bollerslev (1987)], [Hsieh (1989)], [Baillie & DeGennaro (1990)], [Christoffersen *et al.* (2013)] and [Chorro *et al.* (2015)]). [Pagan *et al.* (1990)], [Brailsford & Faff (1996)] and [Loudon *et al.* (2000)] have thoroughly covered the topic of forecasting conditional variance with asymmetric GARCH models. A comparison of normal densities with non-normal ones was carried out by [McMillan *et al.* (2000)], [Yu (2002)] and [Siourounis (2002)]. [Barone-adesi *et al.* (2008)] proposed a method to price options centred on GARCH models with filtered historical non-normal innovations.

1.1.3 Non-Gaussian GARCH model

As mentioned above, while modelling time varying volatility, the main features of the GARCH framework are the volatility specification and the form of the return distribution with the set of parameters θ^D. The dynamics of the conditional volatility are generally combined with the assumption of conditionally Gaussian innovations. Several empirical studies demonstrate that these assumptions are unable to capture the fat tail and asymmetry in the distribution of daily log returns. For example, [Chorro *et al.* (2015)] demonstrates empirically that it is impossible to entirely capture the skewness and excess kurtosis that typify the mass in the tails and the asymmetry of financial time series by using the regular specifications of the GARCH framework.

When seeking a better way to reflect asymmetry, excess kurtosis and excess skewness in the GARCH models, the first major stride was the investigation of GARCH models with non-normal conditional innovations distributions: see [Boothe & Glassman (1987)], [Koedijk *et al.* (1992)] and [Huisman *et al.* (1998)]. Excess kurtosis can be accounted for with some of the heavier-tailed distributions like Student-t ([Bollerslev (1987)], [Baillie & Bollerslev (1989)], and [Beine *et al.* (2002)]) or the GED distribution ([Nelson (1991)]), however these attempts could not explain excess skewness. Indeed, while the Student-t distribution models tales which are thicker than the norm, it does not permit for skewness. Popular leptokurtic distributions, like Student-t, are not flexible enough to capture the high peakedness and the fat-tails of exchange rate returns simultaneously. [Liu & Brorsen (1995)] used an asymmetric stable density to capture skewness in a similar attempt, [Fernández & Steel (1998)] and [Lambert & Laurent (2001)] employed skewed Student's t-distribution for the purposes of modelling both skewness and kurtosis.

A number of interesting methods have been suggested to give a clearer description of this deviation from normality. To show the conditional excess kurtosis, [Bollerslev (1987)] switched the normality of the innovation with students' heavier-tailed distributions. [Glosten *et al.* (1993)] tried to deal with the asymmetry problem through the use of skewed innovation densities. [Christoffersen *et al.* (2006)], [Stentoft (2008)], [Chorro *et al.* (2015)] among others have considered the merits of less rigid innovation distributions, like [Barndorff-Nielsen (1998)]'s Normal-Inverse-Gaussian (NIG),or the Generalized Error Distribution (GED), in an attempt to describe skewness and fat-tails more clearly, [Christoffersen *et al.* (2006)] examine multiple different GARCH option models to gain better insight into the leverage effect. [Christoffersen *et al.* (2013)] have invented an affine discrete-time model with the goal of obtaining a close-form option valuation formula by using the conditional moment-generating function.

The one-dimensional Generalized Hyperbolic $(\mathrm{GH}(\lambda, \alpha, \beta, \delta, \mu))$ distribution of [Barndorff-Nielsen (1977)] is expressed as the following density function:

$$\forall z \in \mathbb{R}, \quad d_{\mathrm{GH}}(z, \lambda, \alpha, \beta, \delta, \mu) = \frac{\left(\frac{\sqrt{\alpha^2-\beta^2}}{\delta}\right)^{\lambda}}{\sqrt{2\pi} K_\lambda\left(\delta\sqrt{\alpha^2-\beta^2}\right)} \frac{K_{\lambda-\frac{1}{2}}\left(\alpha\sqrt{\delta^2+(z-\mu)^2}\right)}{\left(\frac{\sqrt{\delta^2+(z-\mu)^2}}{\alpha}\right)^{\frac{1}{2}-\lambda}} e^{\beta(z-\mu)}$$

where $\delta > 0$, $\alpha > |\beta| > 1$ and K_λ is the modified Bessel function of the third kind $K_\lambda(z) = \dfrac{1}{2}\int_0^{+\infty} y^{\lambda-1} e^{-\frac{z}{2}\left(y+\frac{1}{y}\right)} dy$ for $z > 0$. The parameters μ and δ describe the location and the scale, β describes the skewness (when $\beta = 0$ the distribution is symmetric) and α drives the kurtosis. In particular, when $\alpha^* = \alpha\delta$ and $\beta^* = \beta\delta$, if Z follows $\mathrm{GH}(\lambda, \alpha^*, \beta^*, \delta, \mu)$, then

$$\frac{z-\mu}{\delta} \hookrightarrow \mathrm{GH}(\lambda, \alpha, \beta, 1, 0).$$

A noteworthy feature of the GH distribution is that it facilitates easy computation[11] of the moment generating function provided by :

$$\mathbb{G}_{\mathrm{GH}}(u) = e^{\mu u}\left(\frac{\alpha^2-\beta^2}{\alpha^2-(\beta+u)^2}\right)^{\frac{\lambda}{2}} \frac{K_\lambda\left(\delta\sqrt{\alpha^2-(\beta+u)^2}\right)}{K_\lambda\left(\delta\sqrt{\alpha^2-\beta^2}\right)} \quad \text{with} \quad \mid \beta+u \mid < \alpha.$$

Furthermore, the GH distribution contains, as limiting cases, some distributions that are frequently made use of for financial applications.

There are multiple possible options for parametrizing the NIG-distribution. One could most easily characterize it by thinking of it as a case of GH distribution with $\lambda = \frac{1}{2}$. Using numerical transformation as a basis, [Badescu *et al.*(2015)] achieved the integration of a centered version with unit variance of the NIG parametrization using only two parameters. Form the expression of the mean and the variance of GH $\left(\frac{1}{2}, \alpha, \beta, \delta, \mu\right)$:

$$m = \mu + \frac{\delta\beta}{\sqrt{\alpha^2-\beta^2}}, \qquad \sigma^2 = \frac{\delta\alpha^2}{\sqrt{\alpha^2-\beta^2}}. \tag{1.8}$$

they expressed (δ, μ) in terms of (α, β) by setting $\bar{\alpha} = \delta\alpha$, $\bar{\beta} = \delta\beta$ with $m = 0$ and $\sigma^2 = 1$. Together with 1.8, they were able to solve these equations thus:

$$\sigma^2 = \frac{\delta\,\frac{\bar{\alpha}^2}{\delta^2}}{\left(\sqrt{\frac{\bar{\alpha}^2}{\delta^2}-\frac{\bar{\beta}^2}{\delta^2}}\right)^3} = \frac{\bar{\alpha}^2}{\frac{1}{\delta^2}\left(\sqrt{\bar{\alpha}^2-\bar{\beta}^2}\right)^3} = 1 \implies \delta = \frac{\left(\sqrt{\bar{\alpha}^2-\bar{\beta}^2}\right)^{\frac{3}{2}}}{\bar{\alpha}}$$

the same for the mean :

$$m = \mu + \frac{\delta\frac{\bar{\beta}}{\delta}}{\sqrt{\frac{\bar{\alpha}^2}{\delta^2}-\frac{\bar{\beta}^2}{\delta^2}}} = \mu + \frac{\bar{\beta}}{\sqrt{\bar{\alpha}^2-\bar{\beta}^2}}\frac{\left(\sqrt{\bar{\alpha}^2-\bar{\beta}^2}\right)^{\frac{3}{2}}}{\bar{\alpha}} = 0 \implies \mu = -\frac{\bar{\beta}}{\bar{\alpha}}\left(\sqrt{\bar{\alpha}^2-\bar{\beta}^2}\right)^{\frac{1}{2}}.$$

The NIG distribution is especially capable of providing an explicit and simple log moment generating function

$$\kappa_{\mathrm{NIG}}(z) = \mu z + \sqrt{\bar{\alpha}^2-\bar{\beta}^2} - \delta\sqrt{\bar{\alpha}^2-\left(\bar{\beta}+z\delta\right)^2}$$

which makes the NIG distribution the natural choice to enable the GARCH-type models to use the conditional Esscher transform. Moreover, the NIG distribution is compatible with non zero skewness, large kurtosis which is in accordance with the existence of an implied volatility smile. The reason is that it gives rise to higher probabilities of extreme events in

[11] As a consequence, moments of all orders are finite.

contrast to the Gaussian distribution. Asset return series can also display considerable[12] skewness.

In fact, asymmetry of the distribution can be incorporated using leverage effects ([Nelson (1991)] and [Glosten *et al.* (1993)]), or by assuming skewed innovation densities such as normal inverse Gaussian distribution ([Forsberg & Bollerslev (2002)]) and inverse Gaussian density ([Christoffersen *et al.* (2006)]). Inverse Gaussian distribution has one mode in the interior of the range of possible values and it is skewed to the right. In this section, basic properties of the inverse Gaussian distribution are presented following [Johnson *et al.* (1994)] parameterization. The Inverse Gaussian distribution is an exponential distribution with support on $]0, \infty[$ which is a two-parameters family of continuous probability distributions. The probability density function is :

$$\forall z \in]0, \infty[, \quad d_{\text{IG}}(z, \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi z^3}} e^{\dfrac{-\lambda (x-\mu)^2}{2\mu^2 z}}$$

described by two characteristics, $\mu > 0$ is the mean and $\lambda > 0$ is the shape parameter. As λ tends to infinity, the inverse Gaussian distribution converge to a normal distribution. The log moment generating function κ_{IG} of the IG distribution can be expressed as:

$$\kappa_{\text{IG}}(z) = \frac{\lambda}{\mu}\left(1 - \sqrt{1 - \frac{2\mu^2 z}{\lambda}}\right)$$

Using this log moment generating function $\kappa_{IG}(z)$, first four raw moments of the IG distribution can be computed. The variance, the skewness and the kurtosis of this parameterization of the inverse Gaussian distribution are, respectively,

$$Var(z) = \frac{\mu^3}{\lambda}, \qquad Skew(z) = 3\left(\frac{\mu}{\lambda}\right)^{\frac{1}{2}} \text{ and } Kurt(z) = \frac{15\mu}{\lambda}$$

which tells us that the IG probability density is always positively skewed and the excess kurtosis is always positive. The idea is to have a distribution that can "reach up high" and admit some extreme values. It is pretty easy to estimate μ and λ by maximum likelihood. [Christoffersen *et al.* (2006)] used anather parametrization of the IG distribution with single parameter δ. The mean and the variance of the distribution are equal and the probability density[13] function is given by the following :

$$\forall z \in]0, \infty[, \quad d_{\text{IG}}\left(z; \delta, \delta^2\right) = \frac{\delta}{\sqrt{2\pi} z^{3/2}} \exp\left\{-\frac{1}{2}\frac{(z-\delta)^2}{z}\right\}.$$

and the log moment generating[14] function can be simplified as $\kappa_{IG}(z) = \delta(1 - \sqrt{1-2z})$. Moreover, the cumulative distribution function of the single parameter Inverse Gaussian

[12] Asset return series might exhibit episodes of sharp depreciation (appreciation) not offset by subsequent sharp appreciation (depreciation). Two reasons for skewness are: first, permanent shocks that lead to changes in the equilibrium exchange rate may be asymmetric; rapid improvements in Japanese productivity over the past thirty years is such an example; and second, speculative attacks against a currency tend to be one-sided.

[13] The standard form of inverse Gaussian distribution is $d_{\text{IG}}(z; 1, 1) = \frac{1}{\sqrt{2\pi} x^{3/2}} \exp\left\{-\frac{1}{2}\frac{(x-1)^2}{x}\right\}$.

[14] and the moments $Var(z) = \delta$, $Skew(z) = \dfrac{3}{\sqrt{\delta}}$ and $Kurt(z) = \dfrac{15}{\delta}$

distribution is related to the standard normal distribution by:

$$\begin{cases} \mathbb{P}(Z < z) & = \Phi(z_1) + e^{2\delta}\Phi(z_2), \quad \text{for} \quad 0 < z \leq \delta, \\ \mathbb{P}(Z > z) & = \Phi(-z_1) - e^{2\delta}\Phi(z_2), \quad \text{for} \quad z \geq \delta. \end{cases} \tag{1.9}$$

where $z_1 = \frac{\delta}{z^{1/2}} - z^{1/2}$and $z_2 = \frac{\delta}{z^{1/2}} + z^{1/2}$, where the Φ is cumulative distribution function (CDF) of the standard normal distribution. The variables z_1 and z_2 are related to each other by the identity $z_2^2 = z_1^2 + 4\delta$. As we can see from the cumulative distribution function of the IG (equation 1.9), the function is related to the normal distribution. Its cumulant generating function is the inverse of the cumulant generating function of a Gaussian random variable. The IG-GARCH developed by ([Christoffersen *et al.* (2006)]) consists of combining an Inverse Gaussian distribution with a GARCH type volatility model. Moreover, the idea of explicitly modeling pricing options based on IG-GARCH models has a long history in empirical finance, [Christoffersen *et al.* (2006)] illustrated with an extensive empirical test of the model using S&P500 index options that the Inverse Gaussian GARCH models performance is superior to a standard existing nested model.

An alternative derivation of the Black-Scholes equation and formula involves a risk-neutral measure, under which, as its name suggests, all agents in the economy are neutral to risks, so that they are indifferent between investments with different risks as long as these investments have the same expected return. It can be shown that, in the absence of arbitrage opportunities, there exists a unique risk-neutral measure in a complete market, where all tradable assets can be replicated by a set of fundamental assets (the fundamental theorem of arbitrage). In the literature, it is well known that option prices are derived from the risk-neutral measure. The connection between physical and risk-neutral measure known as pricing kernel, stochastic discount factor or state price density, was developed by [Rubinstein (1976)] and [Brennan (1979)]. The pricing kernel or stochastic discount factor is a key component of any option pricing model. It is a state dependent function[15] that discounts payoffs using time and risk preferences. The fundamental theorem of option pricing suggests that the price is its discounted expected value of future payoff specifically under risk-neutral measure or valuation. [Rubinstein (1976)] set the pricing kernel as a monotonic function of return and [Hansen & Singleton (1982)] postulate that the pricing kernel is a power function of the returns. [Gerber & Shiu (1994)] introduced exponential-affine stochastic discount factor. The choice of an exponential affine pricing kernel often leads to tractable computations, and provides results which are easy to compare with the standard Black-Scholes formula.

1.2 Risk Neutral Valuation for Option Pricing

There are many papers in the financial literature dealing with possible choices of a risk neutral measure, some of the widely used risk neutral measures are identified for general discrete time models. The RNVR has proved to be an important tool in the pricing and

[15]It summarizes investor preferences for payoffs over different states of the world.

hedging of financial derivatives (see [Rubinstein (1976)]). It consists of evaluating the price of contracts as the expected value of the discounted payoff function under a martingale measure.The construction of this new probability measure allows us to price derivatives in hypothetical markets where the economic agents are risk neutral. The minimal martingale measure (MMM) constructed by [Follmer & Schweizer (1991)] was also studied in the financial literature. Two other well known tools are the conditional Esscher transform, which was first applied to option pricing by [Gerber & Shiu (1994)], and the extended Girsanov principle (EGP) introduced by [Elliott & Madan (1998)].

The objective of this section is to present different risk neutral probability measures which are equivalent to the physical measure $\mathbb{P}$. We reminde that $(\mathcal{F}_t)_{t\in\{0,\cdots,T\}}$ represente the information filtration associated to $(z_t)_{t\in\{1,\cdots,T\}}$ with $\mathcal{F}_0 = \{\emptyset,\Omega\}$ and $(\mathcal{F}_t = \sigma(z_u; 1\leqslant u\leqslant t))_{t\in\{0,\cdots,T\}}$. We denote by $\mathbb{G}^{\mathbb{P}}_{Y_t|\mathcal{F}_{t-1}}(u)$ the conditional moment generating function of Y_t given $\mathcal{F}_{t-1}$ is defined by

$$\mathbb{G}^{\mathbb{P}}_{Y_t|\mathcal{F}_{t-1}}(u) = \mathbb{E}^{\mathbb{P}}\left[e^{uY_t}\middle|\mathcal{F}_{t-1}\right] \tag{1.10}$$

where $\mathcal{F}_{t-1}$ denotes the information set prior to time $t-1$ and the notation $\mathbb{E}^{\mathbb{P}}[\cdot|\mathcal{F}_{t-1}]$ denotes the conditional expectation given $\mathcal{F}_{t-1}$ under the dynamic $\mathbb{P}$ measure. We assume that under the historical probability $\mathbb{P}$, the dynamics of the bond price process $(B_t)_{t\in\{0,\cdots,T\}}$ and the discounted price process $(\bar{S}_t)_{t\in\{0,\cdots,T\}}$ are given by

$$B_t = B_{t-1}e^r,\ B_0 = 1 \quad \text{and} \quad \bar{S}_t = \frac{S_t}{B_t},$$

where r is the constant risk-free rate expressed on a daily basis.

Définition 7 *A probability measure $\mathbb{Q}$ is an equivalent martingale measure (EMM) with respect to $\mathbb{P}$ if:*

- $\mathbb{Q}\sim\mathbb{P}$, $\forall X\in\mathcal{F}$, $\mathbb{Q}(X)=0 \Longleftrightarrow \mathbb{P}(X)=0$,
- $(\bar{S}_t)_{t\in\{0,\cdots,T\}}$ *is a martingale under $\mathbb{Q}$ with respect to the information filtration $(\mathcal{F}_t)_{t\in\{0,\cdots,T\}}$, that is*
$$\mathbb{E}^{\mathbb{Q}}\left[\bar{S}_t\middle|\mathcal{F}_{t-1}\right] = \bar{S}_{t-1}.$$

In particular we have $\forall t\in\{0,\cdots,T-1\}$, $\mathbb{E}^{\mathbb{Q}}[S_t|\mathcal{F}_{t-1}] = e^r S_{t-1}$ and we denoted by $\frac{d\mathbb{Q}}{d\mathbb{P}}$ it's Radon-Nikodym derivative. This definition established the links between the existence of an EMM and the absence of arbitrage opportunities. First we state a proposition which is a very useful tool for constructing equivalent martingale measures.

Proposition 1.2.1 *Let $\mathbb{P}$ and $\mathbb{Q}$ be equivalent measures defined on the measurable space $(\Omega,\mathcal{F})$. Then there exists an almost surely positive random variable Z_t such that :*

$$\text{for all } A\in\mathcal{F}_t \qquad \begin{cases} \mathbb{E}^{\mathbb{P}}[Z_t|\mathcal{F}_{t-1}] = 1 \\ \mathbb{Q}(A) = \mathbb{E}^{\mathbb{Q}}[I_t Z_t|\mathcal{F}_{t-1}] \end{cases} \tag{1.11}$$

where $\mathcal{F}_t$ is a finite sub-σ-algebra of $\mathcal{F}$.

The Radon-Nikodym process has the following properties $\forall t \in \{0, \cdots, T\}$ and $s \geqslant t$:

$$Z_t := \frac{d\mathbb{Q}}{d\mathbb{P}} \Bigg| \mathcal{F}_t = \mathbb{E}^{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\Bigg|\mathcal{F}_t\right], \quad \mathbb{E}^{\mathbb{Q}}\left[g\, |\mathcal{F}_t\right] = \frac{\mathbb{E}^{\mathbb{P}}\left[Z_s g|\mathcal{F}_t\right]}{Z_t} \quad \text{and} \quad \mathbb{E}^{\mathbb{Q}}\left[S_T\right] = \mathbb{E}^{\mathbb{P}}\left[S_T \frac{d\mathbb{Q}}{d\mathbb{P}}\Bigg|\mathcal{F}_T\right]$$

where g is a $\mathbb{Q}$-integrable measurable function and Z_T is called the Radon-Nikodym derivative on $\mathcal{F}_T$. Assuming continuously compounded returns, the martingale condition for the discounted stock price can be replaced by:

$$\mathbb{E}^{\mathbb{Q}}\left[e^{Y_t}\big|\mathcal{F}_{t-1}\right] = e^r,$$

which is the immediate consequence of the second condition in Definition 7 as follows :

$$\mathbb{E}^{\mathbb{Q}}\left[\bar{S}_t\big|\mathcal{F}_{t-1}\right] = \bar{S}_{t-1} \Longrightarrow \mathbb{E}^{\mathbb{Q}}\left[e^{-rt}S_t\big|\mathcal{F}_{t-1}\right] = e^{-r(t-1)}S_{t-1} \Longrightarrow \mathbb{E}^{\mathbb{Q}}\left[e^{Y_t}\big|\mathcal{F}_{t-1}\right] = e^r.$$

Another approach allowing to build a martingale measure is based on the well-known stochastic discount factor approach. This approach can be linked with the risk neutral valuation relationship (RNVR) principle using an equilibrium argument.

Définition 8 *A positive process $(M_t)_{t\in\{0,\cdots,T\}}$ adapted to the information filtration $\mathcal{F}$ is call a one periode stochastic discount factor (SDF) process if the following relations hold :*

$$\forall t \in \{0, \cdots, T-1\}, \quad \mathbb{E}^{\mathbb{P}}\left[\frac{B_{t+1}}{B_t}M_{t+1}\Bigg|\mathcal{F}_t\right] = 1 \quad \textit{and} \quad \mathbb{E}^{\mathbb{P}}\left[\frac{S_{t+1}}{S_t}M_{t+1}\Bigg|\mathcal{F}_t\right] = 1. \tag{1.12}$$

In particular, the pricing relations in the previous definition give the following restrictions for the parameters :

$$\begin{cases} \mathbb{E}^{\mathbb{P}} & \left[e^r M_{t+1}|\mathcal{F}_t\right] = 1 \\ \mathbb{E}^{\mathbb{P}} & \left[e^{Y_t} M_{t+1}\big|\mathcal{F}_{t-1}\right] = 1 \end{cases} \tag{1.13}$$

Moreover, following both definition 7 and 8, the specification of an EMM is equivalent to the characterisation of a one-period stochastic discount factor process, through the relation:

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = e^{rT}\prod_{i=1}^{T} M_i. \tag{1.14}$$

where $\mathbb{Q}$ is the EMM and $(M_t)_{t\in\{0,\cdots,T-1\}}$ represents the SDF. In fact, in one hand using tower property[16] of the conditional expection, we have the following

$$\mathbb{E}^{\mathbb{P}}\left[\frac{S_{t+1}}{S_t}M_{t+1}\Bigg|\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{P}}\left[\frac{S_{t+1}}{S_t}M_{t+1}\mathbb{E}^{\mathbb{P}}\left[\frac{S_{t+2}}{S_{t+1}}M_{t+2}\Bigg|\mathcal{F}_{t+1}\right]\Bigg|\mathcal{F}_t\right]$$
$$\mathbb{E}^{\mathbb{P}}\left[\frac{S_{t+1}}{S_t}M_{t+1}\Bigg|\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{P}}\left[\frac{S_{t+1}}{S_t}M_{t+1}\frac{S_{t+2}}{S_{t+1}}M_{t+2}\Bigg|\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{P}}\left[\frac{S_{t+2}}{S_t}M_{t+1}M_{t+2}\Bigg|\mathcal{F}_t\right]$$

and by iteration from $t+2$ to T we have :

$$\mathbb{E}^{\mathbb{P}}\left[\frac{S_{t+1}}{S_t}M_{t+1}\Bigg|\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{P}}\left[\frac{S_T}{S_t}\prod_{k=t+1}^{T} M_k\Bigg|\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{P}}\left[\frac{S_T}{S_t}M_{t,T}\Bigg|\mathcal{F}_t\right] = \frac{\mathbb{E}^{\mathbb{P}}\left[\frac{S_T}{S_t}M_{0,T}\Bigg|\mathcal{F}_t\right]}{M_{0,t}} = 1$$

[16] $\mathcal{H}_1 \subset \mathcal{H}_2 \subset \mathcal{F}$ we have $E(E(X \mid \mathcal{H}_2) \mid \mathcal{H}_1) = E(X \mid \mathcal{H}_1)$.

where $M_{t,T} = \prod_{k=t+1}^{T} M_k$. In the other hand using the risk neutral measure we obtain that :

$$\mathbb{E}^{\mathbb{Q}}\left[\frac{\bar{S}_{t+1}}{\bar{S}_t}\middle|\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{Q}}\left[\frac{S_{t+1}}{S_t}e^{-r}\middle|\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{Q}}\left[\frac{S_{t+1}}{S_t}e^{-r}\mathbb{E}^{\mathbb{Q}}\left[\frac{S_{t+2}}{S_{t+1}}e^{-r}\middle|\mathcal{F}_{t+1}\right]\middle|\mathcal{F}_t\right]$$

$$\mathbb{E}^{\mathbb{Q}}\left[\frac{\bar{S}_{t+1}}{\bar{S}_t}\middle|\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{Q}}\left[\frac{S_{t+1}}{S_t}e^{-r}\frac{S_{t+2}}{S_{t+1}}e^{-r}\middle|\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{Q}}\left[\frac{S_{t+2}}{S_t}e^{-2r}\middle|\mathcal{F}_t\right]$$

and by iteration from $t+2$ to T we have :

$$\mathbb{E}^{\mathbb{Q}}\left[\frac{\bar{S}_{t+1}}{\bar{S}_t}\middle|\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{Q}}\left[\frac{S_T}{S_t}e^{-r(T-t)}\middle|\mathcal{F}_t\right] = \frac{\mathbb{E}^{\mathbb{P}}\left[\frac{S_T}{S_t}e^{-r(T-t)}\mathbb{E}^{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\middle|\mathcal{F}_T\right]\middle|\mathcal{F}_t\right]}{\mathbb{E}^{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\middle|\mathcal{F}_t\right]}$$

$$\mathbb{E}^{\mathbb{Q}}\left[\frac{\bar{S}_{t+1}}{\bar{S}_t}\middle|\mathcal{F}_t\right] = \frac{\mathbb{E}^{\mathbb{P}}\left[\frac{S_T}{S_t}e^{-r(T)}\mathbb{E}^{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\middle|\mathcal{F}_T\right]\middle|\mathcal{F}_t\right]}{e^{-r(t)}\mathbb{E}^{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\middle|\mathcal{F}_t\right]} = 1$$

when comparing both expressions, we can see clearly the link between the notion of EMM and the SDF

$$\mathbb{E}^{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\middle|\mathcal{F}_T\right] = \frac{d\mathbb{Q}}{d\mathbb{P}} = e^{rT}M_{0,T} = e^{rT}\prod_{i=1}^{T} M_i \quad \text{and} \quad \mathbb{E}^{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\middle|\mathcal{F}_t\right] = e^{rt}M_{0,t} = e^{rt}\prod_{i=1}^{t} M_i.$$

which provid the result $\frac{d\mathbb{Q}}{d\mathbb{P}} = e^{rT}\prod_{i=1}^{T} M_i$.

1.2.1 Local Risk Neutral Valuation Relationship

For the pricing of options in a GARCH volatility framework in the financial literature, most of the studies use the local risk neutral valuation principle (or LRNVR for short) introduced by [Duan (1995)] to deal with the choice of a risk neutral measure. One of the important properties of the LRNVR framework of [Duan (1995)] is that this approach provides an economic[17] argument to choose a particular equivalent martingale measure under the GARCH model with conditionally normal stock innovation.

Définition 9 *A probability measure $\mathbb{Q}$ is said to satisfy the local risk neutral valuation relationship if the following conditions are satisfied:*

- *$\mathbb{Q}$ is an equivalent martingale measure equivalent to $\mathbb{P}$,*
- *Given $\mathcal{F}_{t-1}$, Y_t follows a Gaussian distribution under $\mathbb{Q}$,*
- $Var^{\mathbb{Q}}\left[Y_t|\mathcal{F}_{t-1}\right] = Var^{\mathbb{P}}\left[Y_t|\mathcal{F}_{t-1}\right]$.

[17][Duan (1995)] provided a rigorous theoretical foundation and economic justification of the validity of LRNVR. For details, see Theorem 2.1 of [Duan (1995)].

The third condition required in the precedent definition is that the one period conditional variance of the returns are invariant almost surely under the equivalent measures. To show the application of the LRNVR, we consider a GARCH-type process for the log return process Y_t with normal innovations to simplify the illustration:

$$\begin{cases} Y_t &= \left(r + \lambda\sqrt{h_t} - \frac{1}{2}h_t\right) + \sqrt{h_t}z_t \qquad z_t \sim_{\mathbb{P}} N(0,1) \\ h_t &= F\left(z_{t-1}, h_{t-1}, \theta^V\right) \end{cases} \tag{1.15}$$

where r is the one-period risk-free rate and $\mathbb{E}^{\mathbb{P}}\left[\frac{S_t}{S_{t-1}}\middle|\mathcal{F}_{t-1}\right] = e^{r+\lambda\sqrt{h_t}}$.

Generalizing the results of [Rubinstein (1976)] and [Brennan (1979)], [Duan (1995)] derived the locally risk-neutral valuation relationship when dealing with GARCH volatility dynamics. This is satisfied by a risk-neutral measure

$$\mathbb{E}^{\mathbb{Q}}\left[\frac{S_t}{S_{t-1}}\middle|\mathcal{F}_{t-1}\right] = e^r \quad \text{and} \quad h_t^* = Var^{\mathbb{Q}}\left[Y_t|\mathcal{F}_{t-1}\right] = Var^{\mathbb{P}}\left[Y_t|\mathcal{F}_{t-1}\right] = h_t, \tag{1.16}$$

the conditional variance of the logarithmic return is invariant under the change of probability measure from $\mathbb{P}$ to $\mathbb{Q}$ almost surely. Under the measure $\mathbb{Q}$, we can derive the risk-neutral asset return process as:

$$\begin{cases} Y_t &= \left(r - \frac{1}{2}h_t^*\right) + \sqrt{h_t}z_t^* \qquad z_t^* \sim_{\mathbb{Q}} N(0,1) \\ h_t^* &= F\left(z_{t-1}^*, h_{t-1}^*, \theta^V\right) \end{cases} \tag{1.17}$$

where $z_t^* = z_t + \lambda$ is a standard normal random variable under the locally risk-neutral measure $\mathbb{Q}$. In fact, from the definition 9, Y_t follows a Gaussian distribution under $\mathbb{Q}$ and $e^{Y_t} = e^{\upsilon_t + \sqrt{h_t}z_t^*}$. Moreover, we obtain in one hand:

$$\mathbb{E}^{\mathbb{Q}}\left[\frac{S_t}{S_{t-1}}\middle|\mathcal{F}_{t-1}\right] = \mathbb{E}^{\mathbb{Q}}\left[e^{\upsilon_t+\sqrt{h_t}z_t^*}\middle|\mathcal{F}_{t-1}\right] = e^{\upsilon_t}\mathbb{E}^{\mathbb{Q}}\left[e^{\sqrt{h_t}z_t^*}\middle|\mathcal{F}_{t-1}\right]$$

and in the other hand using the moment generating function of a Gaussian random variable we have :

$$\mathbb{E}^{\mathbb{Q}}\left[\frac{S_t}{S_{t-1}}\middle|\mathcal{F}_{t-1}\right] = e^{\left(\upsilon_t + \frac{1}{2}Var^{\mathbb{Q}}[\log(Y_t)|\mathcal{F}_{t-1}]\right)}\mathbb{E}^{\mathbb{Q}}\left[1|\mathcal{F}_{t-1}\right] = e^{\left(\upsilon_t+\frac{1}{2}h_t\right)}.$$

Therefore, we deduce from the martingale condition $\mathbb{E}^{\mathbb{Q}}\left[\frac{S_t}{S_{t-1}} \mid \mathcal{F}_{t-1}\right] = e^r$ that $\upsilon_t = r - \frac{h_t}{2}$ and the dynamics of the conditional volatility may be explicitly expressed as a function of z_t^* :

$$\log\frac{S_t}{S_{t-1}} = r - \frac{1}{2}h_t^* + \sqrt{h_t^*}z_t^* \qquad \text{under} \qquad \mathbb{Q},$$

using again the condition in equation 1.16 with $h_t^* = h_t$ we can write :

$$r - \frac{1}{2}h_t + \sqrt{h_t}z_t^* = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t + \sqrt{h_t}z_t,$$
$$z_t^* = z_t + \lambda,$$

substituting this result into the NGARCH process, yields

$$h_t^* = a_0 + a_1 h_{t-1}^* \left(z_{t-1} - \gamma - \lambda\right)^2 + b_1 h_{t-1}^*.$$

The LRNVR can be applied only when the driving noise is normally distributed. However, many empirical studies show that the normality assumption should be relaxed to allow for leptokurtic and skewed densities. [Duan (1999)] extends LRNVR to be a generalized LRNVR (GLRNVR) to deal with non-Gaussian GARCH with skewed and leptokurtic distributions. Nevertheless, the GLRNVR requires intensive calculation for each time t, thus largely impact the efficiency of the implementation.

The following sections will introduce three well known alternative risk neutral measures that can be applied for non-Gaussian innovations. The conditional Esscher transform, which was first applied to option pricing by [Gerber & Shiu (1994)], the extended Girsanov principle introduced by [Elliott & Madan (1998)]. The last one is the second order Esscher transform by [Monfort & Pégoraro (2012)] which including a quadratic term in the pricing kernel to modify not only the conditional mean but also the conditional variance.

1.2.2 The Esscher transform

The Esscher transform was first introduced in actuarial science by [Esscher (1932)]. [Gerber & Shiu (1994)] show that it can be applied to price derivative securities if the log return process has stationary and independent increments. The conditional Esscher transform has been adapted by [Bühlmann *et al.* (1996)] to price option in discret time financial models. The conditional version is related to a utility maximization problem for some specific form of the utility function. In contrast to [Duan (1995)] approach, this latter framework allow a wide variety of return innovations to be chosen within the class of infinite divisible distribution.

Définition 10 *(Conditional Esscher transform) Let θ_t be a $\mathcal{F}_{t-1}$ measurable random variables. The probability measure $\mathbb{Q}^{ess}$ is called the conditional Esscher transformed measure of $\mathbb{P}$ if conditional moment generating functions exist:*

$$\left.\frac{d\mathbb{Q}^{ess}}{d\mathbb{P}}\right| \mathcal{F}_{t-1} = \prod_{i=1}^{t} \frac{e^{\theta_i Y_i}}{\mathbb{G}^{\mathbb{P}}_{Y_i|\mathcal{F}_{i-1}}\left(\theta_i\right)}$$

where θ_t is denoted as the Esscher parameter with respect to the filtration $\mathcal{F}_t$.

Moreover, following the definition 8 and the equation 1.14 that link stochastic discount factor and equivalent martingale measure, the SDF associated to $\frac{d\mathbb{Q}^{ess}}{d\mathbb{P}}$ have a parametric form characterized by an exponential affine of the log-returns :

$$\forall t \in \{1, \cdots, T\}, \qquad M_t = e^{\theta_t Y_t + \xi_t}$$

where $\xi_t = rT - \log\left(\mathbb{G}^{\mathbb{P}}_{Y_t|\mathcal{F}_{t-1}}(\theta_t)\right)$ are $\mathcal{F}_{t-1}$ measurable random variables.

Based on the description introduced in [Gerber & Shiu (1994)], we can characterize the Esscher risk neutralized measure by the following proposition.

Proposition 1.2.2 *Let the process Z_t defined by:*

$$Z_t = \frac{d\mathbb{Q}^{ess}}{d\mathbb{P}} \bigg| \mathcal{F}_{t-1} = \prod_{i=1}^{t} \frac{e^{\theta_i^* Y_i}}{\mathbb{G}^{\mathbb{P}}_{Y_i|\mathcal{F}_{i-1}}(\theta_i^*)}$$

where θ_i^* is a predictable process defines as the unique solution of the equation:

$$\mathbb{G}^{\mathbb{P}}_{Y_i|\mathcal{F}_{i-1}}(1+\theta_i^*) = e^r \mathbb{G}^{\mathbb{P}}_{Y_i|\mathcal{F}_{i-1}}(\theta_i^*). \tag{1.18}$$

Proof Under $\mathbb{Q}^{ess}$, the moment generating function of Y_t given $\mathcal{F}_{t-1}$ is given by

$$\mathbb{G}^{\mathbb{Q}^{ess}}_{Y_t|\mathcal{F}_{t-1}}(u) = \frac{\mathbb{G}^{\mathbb{P}}_{Y_t|\mathcal{F}_{t-1}}(u+\theta_t)}{\mathbb{G}^{\mathbb{P}}_{Y_t|\mathcal{F}_{t-1}}(\theta_t)} \tag{1.19}$$

which can be obtained by the following :

$$\begin{aligned}
\mathbb{G}^{\mathbb{Q}^{ess}}_{Y_t|\mathcal{F}_{t-1}}(u) &= \mathbb{E}^{\mathbb{Q}^{ess}}\left[e^{uY_t}\middle|\mathcal{F}_{t-1}\right] = \mathbb{E}^{\mathbb{P}}\left[e^{uY_t}\frac{Z_t}{Z_{t-1}}\middle|\mathcal{F}_{t-1}\right] = \mathbb{E}^{\mathbb{P}}\left[e^{uY_t}\frac{e^{\theta_t^* Y_t}}{\mathbb{G}^{\mathbb{P}}_{Y_t|\mathcal{F}_{t-1}}(\theta_t^*)}\middle|\mathcal{F}_{t-1}\right] \\
&= \frac{\mathbb{E}^{\mathbb{P}}\left[e^{uY_t}e^{\theta_t^* Y_t}\middle|\mathcal{F}_{t-1}\right]}{\mathbb{G}^{\mathbb{P}}_{Y_t|\mathcal{F}_{t-1}}(\theta_t^*)} = \frac{\mathbb{E}^{\mathbb{P}}\left[e^{(u+\theta_t)Y_t}\middle|\mathcal{F}_{t-1}\right]}{\mathbb{G}^{\mathbb{P}}_{Y_t|\mathcal{F}_{t-1}}(\theta_t^*)} = \frac{\mathbb{G}^{\mathbb{P}}_{Y_t|\mathcal{F}_{t-1}}(u+\theta_t)}{\mathbb{G}^{\mathbb{P}}_{Y_t|\mathcal{F}_{t-1}}(\theta_t)}.
\end{aligned}$$

When $u = 1$, we obtain $\mathbb{G}^{\mathbb{Q}^{ess}}_{Y_t|\mathcal{F}_{t-1}}(1) = e^r$ then the martingale equation 1.18 holds. ■
It is important to notice that the solution of θ_t^* always exists, given that the moment generating function exists and is twice differentiable. Subsequently, the equation 1.18 ensures the unicity and the martingale property of the conditional Esscher transform risk neutral measure is satisfied. Then $\mathbb{Q}^{ess}$ is called the conditional Esscher transform with respect to $\mathbb{P}$ generated by the process Y_t and the Esscher parameter θ_t^*. Under the conditional Esscher transform, the Radon-Nikodym derivative $\frac{d\mathbb{Q}^{ess}}{d\mathbb{P}}$ can be describe by:

$$\frac{d\mathbb{Q}^{ess}}{d\mathbb{P}} \bigg| \mathcal{F}_T = Z_T = \prod_{i=1}^{T} \frac{e^{\theta_i^* Y_i}}{\mathbb{G}^{\mathbb{P}}_{Y_i|\mathcal{F}_{i-1}}(\theta_i^*)}. \tag{1.20}$$

If we consider a Gaussian innovation under the Esscher transform measure $\mathbb{Q}^{ess}$ since for all $t \in \{1, \cdots, T\}$ we have $Y_t \sim N(m_t, h_t)$ (GARCH model in equation 1.4) under measure $\mathbb{P}$, then the conditional moment generating function is given by :

$$\mathbb{G}^{\mathbb{P}}_{Y_i|\mathcal{F}_{i-1}}(u) = e^{um_t + u^2\frac{h_t}{2}}$$

$$\mathbb{G}^{\mathbb{Q}^{ess}}_{Y_t|\mathcal{F}_{t-1}}(u) = \frac{e^{\left((u+\theta_t^*)m_t + (u+\theta_t^*)^2\frac{h_t}{2}\right)}}{e^{\left((\theta_t^*)m_t + (\theta_t^*)^2\frac{h_t}{2}\right)}} = e^{\left(um_t + u^2\frac{h_t}{2} + uh_t\theta_t^*\right)}$$

In the particular if $m_t = \left(r + \lambda\sqrt{h_t} - \frac{1}{2}h_t\right)$, for $u = 1$ in the equation 1.19, the Esscher parameter θ_t^* can be expressed as :

$$\theta_t^* = \frac{1}{h_t}\left(r - m_t - \frac{h_t}{2}\right) = \frac{-\lambda}{\sqrt{h_t}}.$$

Therefore, by equation 1.20 we have :

$$\frac{d\mathbb{Q}^{ess}}{d\mathbb{P}} = \prod_{i=1}^{T} \frac{e^{\theta_i^* Y_i}}{\mathbb{G}^{\mathbb{P}}_{Y_i|\mathcal{F}_{i-1}}(\theta_i^*)} = \prod_{i=1}^{T} e^{\theta_i^* Y_i - \theta_i^* m_i - \frac{h_i}{2}(\theta_i^*)^2}$$

$$\frac{d\mathbb{Q}^{ess}}{d\mathbb{P}} = \prod_{i=1}^{T} e^{\left(-\frac{\lambda}{\sqrt{h_i}}Y_i + \frac{r\lambda}{\sqrt{h_i}} + \lambda - \frac{\lambda\sqrt{h_i}}{2} - \frac{\lambda^2}{2}\right)}$$

$$\frac{d\mathbb{Q}^{ess}}{d\mathbb{P}} = \prod_{i=1}^{T} e^{-\frac{1}{2h_i}\left(\lambda^2 h_i + 2\lambda\sqrt{h_i}\left(Y_i - r - \lambda\sqrt{h_i} + \frac{1}{2h_i}\right)\right)}.$$

Moreover, the asset returns are conditionally Gaussian $Y_t \sim_{\mathbb{Q}^{ess}} N(m_t + \theta_t^* h_t, h_t)$ under $\mathbb{Q}^{ess}$, and the dynamics for Y_t is again a Gaussian model the same dynamics characterization as Duan's RNVR. The conditional Esscher transform is thus a convenient tool for derivative valuation when the distributions of asset returns are non-normal. However, solving for θ_t for each time step $t \in \{1, \cdots, T\}$ may be computationally demanding.

1.2.3 The Extended Girsanov Principle

The extended Girsanov Principle was proposed by [Elliott & Madan (1998)] and provides another approach in choosing a risk neutral measures under the discrete time framework. The construction of the new measure is based[18] on a multiplicative Doob decomposition of the discounted stock price $\left(\tilde{S}_t\right)_{t\in\{0,\cdots,T\}}$ as a product of a predictable process and a martingale:

$$\tilde{S}_t = \tilde{S}_0 A_t N_t \qquad \text{with} \qquad A_t = \prod_{i=0}^{t-1} \mathbb{E}^{\mathbb{P}}\left[\frac{\tilde{S}_{i+1}}{\tilde{S}_i}\middle|\mathcal{F}_i\right]$$

where N_t is a $\mathcal{F}_t$ martingale, A_t is a $\mathcal{F}_t$ predictable process[19] with respect to $\mathcal{F}_t$.

Définition 11 *A probability $\mathbb{Q}$ with respect to $\mathcal{F}$ is said to satisfy the Extended Girsanov Principle if the conditional law of the discounted stock price under the new measure is equal to the conditional law where their martingale component from the multiplicative Doob decomposition prior to the change of measure :*

$$\mathcal{L}^{\mathbb{Q}}\left[\frac{\tilde{S}_t}{\tilde{S}_{t-1}}\middle|\mathcal{F}_{t-1}\right] = \mathcal{L}^{\mathbb{P}}\left[\frac{N_t}{N_{t-1}}\middle|\mathcal{F}_{t-1}\right]. \tag{1.21}$$

[18]The economic foundation of the extended Girsanov principle is to minimize the adjusted hedging capital of investors hedging portfolio. For details, the reader is referred to [Elliott & Madan (1998)]

[19]Immediate consequence of $\frac{\tilde{S}_{i+1}}{\tilde{S}_i} = \frac{A_{i+1}N_{i+1}}{A_i N_i} \Longrightarrow \mathbb{E}^{\mathbb{P}}\left[\frac{\tilde{S}_{i+1}}{\tilde{S}_i}\middle|\mathcal{F}_i\right] = \frac{A_{i+1}}{A_i}$.

In fact, the dynamics of the discounted stock price process under $\mathbb{P}$ has the following representation :

$$\tilde{S}_t = \tilde{S}_{t-1} e^{\nu_t} \frac{N_{t+1}}{N_t} \qquad \text{with} \qquad e^{\nu_t} = \frac{A_t}{A_{t-1}} = e^{-r} \mathbb{E}^{\mathbb{P}} \left[e^{Y_t} \middle| \mathcal{F}_{t-1} \right].$$

where ν_t is the one period discounted excess returns between $t-1$ et t.

Définition 12 *Define a change of measure density process* $(L_t)_{t \in \{1, \cdots, T\}}$ *by*

$$\bar{L}_t = \prod_{i=1}^{t} \frac{g_i^{\mathbb{P}} \left(\frac{\tilde{S}_i}{\tilde{S}_{i-1}} \right) e^{\nu_i}}{g_i^{\mathbb{P}} \left(e^{-\nu_i} \frac{\tilde{S}_i}{\tilde{S}_{i-1}} \right)}.$$

which is a $\mathcal{F}$-martingale under $\mathbb{P}$ where $g_t^{\mathbb{P}}$ is the conditional density function of $\frac{N_t}{N_{t-1}}$ given $\mathcal{F}_{t-1}$ under $\mathbb{P}$.

Proposition 1.2.3 *([Elliott & Madan (1998)]) Let $\mathbb{Q}^{egp}$ be the probability defined by the density $\bar{L}_T$ with respect to $\mathbb{P}$ then $\mathbb{Q}^{egp}$ is the unique equivalent probability measure that satisfies the definition 11. Under the extended Girsanov change of measure $\mathbb{Q}^{egp}$ is then defined:*

$$\frac{d\mathbb{Q}^{egp}}{d\mathbb{P}} = \bar{L}_T,$$

with $\mathbb{Q}^{egp}$ is the unique equivalent probability measure that satisfies the equation 1.21.

In the Gaussian case, the risk-neutral dynamics implied by the EGP coincide with those given by the conditional Esscher transform. However, this is no longer the case for the NIG distribution case. Under the extended Girsanov principle measure, the Radon-Nikodym derivative is given by

$$\frac{d\mathbb{Q}^{egp}}{d\mathbb{P}} = \prod_{i=1}^{T} \frac{f_i^{\mathbb{P}} \left(Y_t - r + m_t + \frac{h_t}{2} \right)}{f_i^{\mathbb{P}} (Y_t)} = \prod_{i=1}^{T} e^{-\frac{1}{2h_i} \left(\lambda^2 h_i + 2\lambda\sqrt{h_i} \left(Y_i - r - \lambda\sqrt{h_i} + \frac{1}{2h_i} \right) \right)}.$$

where $f_i^{\mathbb{P}}$ is the conditional pdf of Y_t given $\mathcal{F}_{t-1}$. Thus,

$$\frac{d\mathbb{Q}^{egp}}{d\mathbb{P}} = \prod_{i=1}^{t} \frac{f_i^{\mathbb{P}} \left(Y_i - r + m_i + \frac{h_i}{2} \right)}{f_i^{\mathbb{P}} (Y_i)} = \prod_{i=1}^{t} \frac{e^{-\frac{1}{2h_i} \left(Y_i - r + \frac{h_i}{2} \right)^2}}{e^{-\frac{(Y_i - m_i)^2}{2h_i}}}$$

$$\frac{d\mathbb{Q}^{egp}}{d\mathbb{P}} = \prod_{i=1}^{t} e^{-\frac{1}{2h_i} \left(\left(Y_i - r + \frac{h_i}{2} \right)^2 - (Y_i - m_i)^2 \right)} = \prod_{i=1}^{t} e^{-\frac{1}{2h_i} \left(\left(Y_i - r + \frac{h_t}{2} \right)^2 - \left(Y_i - \left(r + \lambda\sqrt{h_i} - \frac{1}{2} h_i \right) \right)^2 \right)}$$

$$\frac{d\mathbb{Q}^{egp}}{d\mathbb{P}} = \prod_{i=1}^{t} e^{-\frac{1}{2h_i} \left(\lambda^2 h_i + 2\lambda\sqrt{h_i} \left(Y_i - r - \lambda\sqrt{h_i} + \frac{1}{2h_i} \right) \right)}.$$

It is important to notice that concerning the Model 1.15 the Radon-Nikodym derivative $\frac{d\mathbb{Q}^{ess}}{d\mathbb{P}}$ and $\frac{d\mathbb{Q}^{egp}}{d\mathbb{P}}$ derived by the Esscher transform and the extended Girsanov principle, respectively, satisfy the same representation :

$$\frac{d\mathbb{Q}^{ess}}{d\mathbb{P}} = \frac{d\mathbb{Q}^{egp}}{d\mathbb{P}} = \prod_{i=1}^{t} e^{-\frac{1}{2h_i}\left(\lambda^2 h_i + 2\lambda\sqrt{h_i}\left(Y_i - r - \lambda\sqrt{h_i} + \frac{1}{2h_i}\right)\right)}$$

Consequently, the equivalent martingale measures $\mathbb{Q}^{ess}$ and $\mathbb{Q}^{egp}$ obtained from the two approaches are the same. Subsequently, the risk-neutral dynamics of Model 1.15 under measure $\mathbb{Q}^{ess}$ and $\mathbb{Q}^{egp}$ is written as follows:

$$\begin{cases} Y_t &= \left(r - \frac{1}{2}h_t^*\right) + \sqrt{h_t}z_t^* \\ h_t^* &= F\left(z_{t-1}^*, h_{t-1}^*, \theta^V\right) \end{cases} \tag{1.22}$$

with $z_t^* \sim_{\mathbb{Q}} N(0,1)$ and $z_t^* = z_{t-1} - \lambda$. Duan's method is specific to normal innovations and does not naturally extend to non-normal GARCH. One advantage of the EGP is that it does not require any distributional assumption about the returns. Thus this principle can be applied to investigate pricing and hedging for various types of discrete time models.

In another point of view, few papers have investigated certain aspects of the empirical performance of monotonically declining kernels. For instance [Bakshi *et al.*(1997)], illustrate that the prices of S&P500 calls are inconsistent with monotonically declining kernels and motivate U-shaped pricing kernels. The first empirical evidence of nonmonotone pricing kernels was reported in [Aït-Sahalia & Lo (1998)] and [Jackwerth (2000)]. Subsequent empirical assessments of pricing kernel monotonicity include [Barone-adesi *et al.* (2008)], [Chabi-Yo *et al.* (2008)], [Barone-adesi *et al.* (2012)], [Christoffersen *et al.* (2013)], [Beare & Schmidt (2016)]. Each of these studies produced a comprehensive assertion on the importance to leave out the idea of the monotonicity of the stochastic discount factor. We mention, among many others, [Monfort & Pégoraro (2012)] introduce the notion of Second-Order GARCH Option Pricing Model using the Second-Order Esscher Transform as a U-shaped function stochastic discount factor with the exponential quadratic fonction the log-returns.

1.2.4 Second Order Esscher Transform

A variety of alternatives for pricing kernel have been developed to provide an answer to the nonmonotone pricing kernels problems. As a natural alternative of the of monotonically declining kernels, [Monfort & Pégoraro (2012)] derived from the Esscher Transform the notion of Second Order Esscher Transform for GARCH Option Pricing Model. They introduced an extension of the classical Essher transform including a quadratic term in the pricing kernel. This approache propose to modify not only the conditional mean but also the conditional variance. Providing the intuition of a U-shaped stochastic discount factor, they specified the exponential quadratic function of log-returns by using first-order and second-order stochastic risk-sensitivity coefficients as follows :

$$\forall t \in \{1, \cdots, T\}, \quad M_t = e^{\theta_{2,t} Y_t^2 + \theta_{1,t} Y_t + \xi_t} \tag{1.23}$$

where ξ_t, $\theta_{1,t}$ and $\theta_{2,t}$ are $\mathcal{F}_{t-1}$ measurable random variables. Moreover, the pricing relation described in definition 1.12 can be written in this case as:

$$\begin{cases} \xi_t &= -r - \log\left(\mathbb{G}^{\mathbb{P}}_{(Y_t, Y_t^2)|\mathcal{F}_{t-1}}(\theta_{1,t}, \theta_{2,t})\right) \\ e^r &= \dfrac{\mathbb{G}^{\mathbb{P}}_{(Y_t, Y_t^2)|\mathcal{F}_{t-1}}(\theta_{1,t}+1, \theta_{2,t})}{\mathbb{G}^{\mathbb{P}}_{(Y_t, Y_t^2)|\mathcal{F}_{t-1}}(\theta_{1,t}, \theta_{2,t})}. \end{cases}$$

Furthermore, the dynamics of the log-returns Y_t can be describe using the moment generation function under the equivalent martingale measure $\mathbb{Q}^{Qua}$ associated to the exponential quadratic stochastic discount factor M_t in 1.23:

$$\mathbb{G}^{\mathbb{Q}^{Qua}}_{Y_t|\mathcal{F}_{t-1}}(u) = \frac{\mathbb{G}^{\mathbb{P}}_{(Y_t, Y_t^2)|\mathcal{F}_{t-1}}(u+\theta_{1,t}, \theta_{2,t})}{\mathbb{G}^{\mathbb{P}}_{(Y_t, Y_t^2)|\mathcal{F}_{t-1}}(\theta_{1,t}, \theta_{2,t})}$$

However, unlike the Esscher transform, where the pricing equations have a unique solution (proposition 1.2.2), the previous system has in general an infinite number of solutions.

In particular, if $\forall t \in \{1, \cdots, T\}$, $\theta_{2,t} = 0$ we obtain the risk neutral dynamics associated to the Esscher transform in system 1.22 with $h_t = h_t^*$. Considering the Gaussian GARCH model (in equation 1.15) where $Y_t \sim N(m_t, h_t)$, the logarithm of the conditional moment generating function $\mathbb{G}^{\mathbb{P}}_{(Y_t, Y_t^2)|\mathcal{F}_{t-1}}(u, v)$ of $\left(Y_t, Y_t^2\right)$ under $\mathbb{P}$ can be express in the following way:

$$\log\left(\mathbb{G}^{\mathbb{P}}_{(Y_t, Y_t^2)|\mathcal{F}_{t-1}}(u, v)\right) = \log\left(\mathbb{E}^{\mathbb{P}}\left[e^{uY_t + vY_t^2}\middle|\mathcal{F}_{t-1}\right]\right)$$

$$\log\left(\mathbb{G}^{\mathbb{P}}_{(Y_t, Y_t^2)|\mathcal{F}_{t-1}}(u, v)\right) = \frac{h_t u^2 + 2m_t u + 2vm_t^2}{2(1-2vh_t)} - \frac{1}{2}\log(1-2vh_t)$$

for $(u, v) \in \mathbb{R}^2$ with $v < \dfrac{1}{2h_t}$. This is due to the very particular form of the Gaussian density function (see [Chorro *et al.* (2015)] for more details).

Following [Monfort & Pégoraro (2012)], assuming $\theta_{2,t} < \dfrac{1}{h_t}$ and considering $\pi = \dfrac{h_t^*}{h_t}$ as a constant proportional wedge between h_t and h_t^*, in the case of NGARCH, we obtain the dynamics of Y_t under $\mathbb{Q}^{Qua}$:

$$\begin{cases} Y_t &= \left(r - \frac{1}{2}h_t^*\right) + \sqrt{h_t^*} z_t^* \qquad z_t^* \sim_{\mathbb{Q}} N(0,1) \\ h_t^* &= \pi a_0 + \pi a_1 h_{t-1}^* \left(z_{t-1}^* - \frac{1}{\sqrt{\pi}}(\lambda+\gamma) - \frac{1}{\pi}(\pi - 1)\right)^2 + b_1 h_{t-1}^* \end{cases}$$

with $\pi = 1 + 2\theta_{2,t} h_t^*$ and $\pi > 0$ implying a U-shape for 1.23, the U-shape quadratic Esscher transform induces a risk neutral variance strictly greater than the historical one. If $\forall t \in \{1, \cdots, T\}$, $\theta_{2,t} = 0$ then $\pi = 1$ which implies that $h_t = h_t^*$. In this case we recover the risk neutral dynamics associated to the Esscher transform in system 1.22 under $\mathbb{Q}^{ess}$.

1.3 Estimation

After selecting a reasonable model, we need to estimate the parameters to fit the data. In this section, our aim is to fit the GARCH models we discussed in section 1.1. As such, we focus our discussion on a brief review of the maximum likelihood estimation (MLE) and the quasi-maximum likelihood (QML) strategies to estimate the parameter $\theta = \left(\lambda, \theta^V, \theta^D\right)$ based on i.i.d observations.

1.3.1 Maximum Likelihood Estimator

The maximum likelihood estimator remains the preferred estimator to estimate the parameters. The probability density function for a random variable Y of the model 1.4, conditioned on a set of parameters $\theta = \left(\lambda, \theta^V, \theta^D\right)$, is denoted by $f\left(Y|\theta\right)$. In this setting, the joint density of n-observation independent and identically distributed observations $(Y_1, \cdots, Y_n)$ from this process is the product of the individual densities:

$$L\left(Y|\theta\right) = f\left(Y_1, \cdots, Y_n|\theta\right) = \prod_{i=1}^{n} f\left(Y_i|Y_1, \cdots, Y_{i-1}, \theta\right)$$

This joint density is the likelihood function, defined as a function of the unknown parameters vector, θ, where Y is used to indicate the collection of sample data. The conditional log-likelihood based on the observations $(Y_1, \cdots, Y_n)$ is

$$l\left(Y|Y_1, \cdots, Y_n, \theta\right) = \log L\left(Y|\theta\right) = \sum_{i=1}^{n} \log\left[f\left(Y_i|Y_1, \cdots, Y_{i-1}, \theta\right)\right], \tag{1.24}$$

In practice it is often more convenient to work with the logarithm of the likelihood function. The best estimator, $\hat{\theta}$, is the value of θ that maximizes $L\left(Y|\theta\right)$ (which is equivalent to maximize $l\left(Y|Y_1, \cdots, Y_n, \theta\right)$). We are looking for the $\hat{\theta}$, that maximizes the likelihood of observing our sample, when it exists, by

$$\hat{\theta}^* = \arg\underset{\theta}{\mathrm{Max}} \quad f\left(Y_1, \cdots, Y_n|\theta\right) \tag{1.25}$$

or equivalently as the solution of the vectorial score equation:

$$\sum_{i=1}^{n} \frac{\partial \log\left[f\left(Y_1, \cdots, Y_n|\theta\right)\right]}{\partial \theta} = 0$$

which is the necessary condition for maximizing $l\left(Y_1, \cdots, Y_n|\theta\right)$. Let's work through an example analytically for the Inverse Gaussian distributions. The IG distribution, which is often used as the underlying distribution to capture asymmetry with GARCH option pricing, the conditional density function of Y_n given $(Y_1, ..., Y_{n-1})$ is given by :

$$f\left(Y_n \mid Y_1, ..., Y_{n-1}, \eta\right) = \prod_{i=1}^{n} \frac{\eta}{\sqrt{2\pi} Y_i^{3/2}} \exp\left\{-\frac{1}{2}\frac{(Y_i - \eta)^2}{Y_i}\right\}.$$

and the conditional log-likelihood is given by

$$l\left(Y_1,\cdots,Y_n|\eta\right)=\sum_{i=1}^{n}\log\left[f\left(Y_i|Y_1,\cdots,Y_{i-1},\eta\right)\right],$$
$$l\left(Y_1,\cdots,Y_n|\eta\right)=\sum_{i=1}^{n}\log\left[\frac{\eta}{\sqrt{2\pi}Y_i^{3/2}}\exp\left\{-\frac{1}{2}\frac{\left(Y_i-\eta\right)^2}{Y_i}\right\}\right],$$
$$l\left(Y_1,\cdots,Y_n|\eta\right)=n\log\left(\eta\right)-\frac{n}{2}\log\left(2\pi\right)-\frac{3n}{2}\log\left(\left(Y_i\right)\right)-\frac{1}{2}\sum_{i=1}^{n}\left[\frac{\left(Y_i-\eta\right)^2}{Y_i}\right]$$

and we calculate derivatives of the log-likelihood function with respect the parameter η, we obtain the likelihood equations for η :

$$\frac{\partial l\left(Y_1,\cdots,Y_n|\eta\right)}{\partial\eta}=0\iff\eta^2\sum_{i=1}^{n}\left[\frac{1}{Y_i}\right]+n\eta+n=0\quad\text{where}\quad\eta>0$$

The maximum likelihood estimator is efficient, and it achieves Cramér-Rao lower bound when the sample size tends to infinity. However, this method may lead to inconsistent estimates if the distribution of the innovation is misspecified. Alternatively, the Gaussian MLE, regarded as a quasi-maximum likelihood estimator (QMLE) may be consistent and asymptotically normal, provided that the innovation has a finite fourth moment.

1.3.2 Quasi-Maximum Likelihood Estimation

The Quasi Maximum Likelihood[20]as brought in by [Wedderburn (1974)],which is sometimes referred to as pseudo-likelihood estimate, this estimator is possibly the most well-known estimation strategy for conditional heteroskedasticity time series. The function that is maximized to form a QMLE a more simple structure of the actual log likelihood function defined in 1.24. This simple function is frequently formed with the log-likelihood function of an unspecified model. On the other hand, the original ML method assumes that the specified density function is the true density function. As a consequence, the findings in the ML method are merely special cases of the QML method. [Francq & Zakoian (2004)] has suggested that the conditional heteroskedasticity time series is specifically applicable to Quasi-maximum likelihood (QML) method.

Is this case, we can use the dynamic GARCH in mean model to show the method of Quasi-maximum likelihood. When we remembering the dynamic GARCH model in the function 1.15 without specific distribution for the innovation z_t, the observations $(Y_1,\cdots,Y_n)$ follow the dynamic:

$$\begin{cases} Y_t &= \left(r+\lambda\sqrt{h_t}-\frac{1}{2}h_t\right)+\sqrt{h_t}z_t \\ h_t &= F\left(z_{t-1},h_{t-1},\theta^V\right)\end{cases} \tag{1.26}$$

where $(z_t)_{t\in\{1,\cdots,n\}}$ is a sequence of i.i.d. random variables of variance one, mean zero. In this setting of the QML approach, the conditional log-likelihood based on the observations $(Y_1,\cdots,Y_n)$:

[20] w

$$l\left(Y_1,\cdots,Y_n \mid \theta\right)=\sum_{t=1}^{n}-\frac{\log(h_t)}{2}+\log\left[d_{\theta^D}\left(\frac{Y_t-r-\lambda\sqrt{h_t}+\frac{1}{2}h_t}{\sqrt{h_t}}\right)\right] \tag{1.27}$$

with the set of the parameters $\theta=\left(\lambda,\theta^V,\theta^D\right)$.The true value of the parameter is not known, and is shown by $\theta=\left(\lambda,\theta^V,\theta^D\right)$ which denotes the value under the true model. The innovations are i.i.d random variables with mean 0, variance 1, and unknown density $f(\cdot)$, we do not make any assumption on the distribution for QML.

- We need the conditions for h_0 and θ to begin with. For a given value of parameters, under the second-order stationarity assumption, the unconditional variance is an acceptable [21] selection for the unidentified initial values.
- The Gaussian GARCH framework, which we apply, corresponds to z_t are habitually distributed normally where the conditional Gaussian quasi-likelihood are shown as :

$$l\left(Y_1,\cdots,Y_n \mid \lambda,\theta^V\right)=\sum_{t=1}^{n}-\frac{\log(h_t)}{2}-\frac{\log(2\pi)}{2}-\frac{\left(Y_t-r-\lambda\sqrt{h_t}+\frac{1}{2}h_t\right)}{2h_t}$$

  and $\left(\lambda,\theta^V\right)$ are estimated by maximizing the conditional Gaussian quasi-likelihood as follows :

$$\hat{\theta}=\left(\hat{\lambda},\hat{\theta}^V\right)=\arg\underset{(\lambda,\theta^V)}{\text{Max}}\; l\left(Y_1,\cdots,Y_T \mid \lambda,\theta^V\right) \tag{1.28}$$

- θ^D is obtained by maximizing the conditional log-ikelihood in the function 1.27 with residuals that can be assessed from the step before:

$$\left(z_i\left(\hat{\lambda}_{0,n},\hat{\theta}_n^V\right)\right)_{i\in\{1,\cdots,n\}}=\left(\frac{Y_i-r-\lambda\sqrt{h_i}+\frac{1}{2}h_i}{\sqrt{h_i}\left(\hat{\theta}_n^V\right)}\right)_{i\in\{1,\cdots,n\}}$$

By contrast, less consideration is given to inference using a non-Gaussian QMLE. Normally a non-Gaussian QMLE does not show a consistent estimation when the true error distribution strays from the likelihood. A non-Gaussian QMLE method which is tough on error misidentification, is more effective than Gaussian QMLE, and it needs to choose the correct innovation distribution.

1.3.3 Other estimation strategies

However, calibration on option prices can directly value the GARCH model parameters. This method finds parameters using the non-linear least squares estimation (NLS) that reduces a loss function that identifies the error between the model prices and the market ones. A

[21] In reality what is important is the choice of initial values

wide breadth of work has already been done showing where stochastic volatility models are valued based on empirical information on option prices using the non-linear least squares (NLS) approach (for instance: [Christoffersen *et al.* (2004)], [Barone-adesi *et al.* (2008)]). There are two conceivable approaches, one that estimates a different set of parameters for each cross-section of options, and one which is used to estimate all cross-sections in the sample. Many employ NLS, using loss functions to minimize the pricing error of the daily cross-section of options, including but not exclusively [Bakshi *et al.*(1997)]. [Duan (1995)] suggests a GARCH model where we would use a single cross-section. The latter method is used in [Heston & Nandi (2000)] and [Christoffersen *et al.* (2004)], for example.

Moreover, the importance of filtering volatility from fundamental returns to approximating model parameters using several cross-sections of options means the subsequent estimations are consistent with returns. This is in the sense that the volatility used to value options is consistent with the models' risk neutralization and the fundamental return data. Yet while returns are used in the filtering, the loss function does not explicitly cover an element based on returns. This situation needs to combine loss functions for options data and returns data. The works that exist on option pricing with GARCH specifications use returns to filter conditional spot volatility. Let us deliberate on the specification of the GARCH in mean model in equation1.4. In this case, h_t can be extracted as observable using the volatility updating rule (see [Christoffersen *et al.* (2004)]):

$$\begin{cases} h_t &= F\left(z_{t-1}, h_{t-1}, \theta^V\right), \\ z_t &= \dfrac{Y_t - (r + m_t)}{\sqrt{h_t}}. \end{cases} \tag{1.29}$$

The updating from h_t to h_{t+1} can then only be expressed in terms of observables and parameters by substituting z_t in the variance dynamic. We can attain a returns based proxy for spot variances $\left(h_t^R\right)$ for the structural parameters that are given $-$. [Heston & Nandi (2000)] and [Christoffersen *et al.* (2004)] set initial spot volatility at time zero to equal unconditional volatility 250 days before the first date included in the sample. Therefore we can note any loss function of interest as only a function of the parameters and observables.

In [Kanniainen *et al.* (2014)] it is suggested that the root mean square error between model and market option prices is a justifiable loss[22] θ^* shows that they find estimates of the set of the risk neutral parameters, minimizing the Implied Volatility Root Mean Square Error (IVRMSE)[23]:

$$\hat{\theta}^* = \arg\underset{\theta}{\text{Min}}\; IVRMSE\,(\theta) = \arg\underset{\theta}{\text{Min}} \sqrt{\frac{1}{N_{T_{Op}}} \sum_{t,i} \left(\frac{c_{i,t}\left(h_t^R; \theta\right) - \hat{c}_{i,t}}{\hat{V}_{i,t}} \right)^2}. \tag{1.30}$$

In this equation, n_t is the amount of option contracts in the model at time t and $N_{T_{Op}} = \sum_{t=1}^{T_{Op}} n_t$ where T_{Op} is the number of days in the options sample. $c_{i,t}\left(h_t^*; \theta\right)$ shows

[22]Note that financial works on option theory do not propose the appropriate loss function. Because of this loss functions are mostly selected out of econometric function.

[23]The Implied Volatility Root Mean Square Error (IVRMSE). This will be used in the empirical study to assess and compare the performances of the models, in terms of pricing

the price of the $i-$th option at time t given by the model[24] while $\hat{c}_{i,t}$ is the price observed in the market. $\hat{V}_{i,t}$ is the Vega associated to $\hat{c}_{i,t}$ that is calculated using the implied Black-Scholes volatility $\sigma_{i,t}$ attained from the market price.

[Kanniainen *et al.* (2014)] proposed a difference method to make spot volatility observable. They suggested extracting the daily spot volatilities from the series of VIX, instead of calculating spot volatilities from return sequence with the volatility updating rule VIX can be theoretically worked out by using a formula given by the CBOE :

$$\text{VIX}_t^2 = \frac{2}{\tau}\sum_i \frac{\triangle K_i}{K_i^2} Q(K_i) - \frac{1}{\tau}\left(\frac{F_t(t+\tau)}{K_0} - 1\right)^2$$

where $\tau = \dfrac{30}{365}$, r is the risk-free interest rate to expiration, F_t is the forward index level computed bt the index option prices, K_0 denoted the first strike below the the forward index level F, K_i is the strike of i^{th} out-of-money option[25] $Q(K_i)$ is the midpoint of the bid-ask spread for each option with strike K_i, $\triangle K_i$ is the interval between strike prices - half the distance between the strike on either side of K_i. The VIX index is quoted as percentage rather than a dollar amount.

In fact, VIX represents the market's expectation of the movements in the S&P500 about volatility 30 day ahead and estimates expected volatility from the prices of stock index options in a wide range of strike prices:

$$\frac{1}{\tau}\left(\frac{\text{VIX}_t}{100}\right)^2 \cong \mathbb{E}^{\mathbb{Q}}\left[\sum_{j=1}^{T} h_{t+j}\right] \tag{1.31}$$

where $T = 30$, $\mathbb{E}^{\mathbb{Q}}[.]$ is an expectation under the risk-neutral measure. [Kanniainen *et al.* (2014)] extablixhed close forme expression for the VIX for the case of the affine model by [Heston & Nandi (2000)] (HN-GARCH), the nonaffine models GJR-GARCH by[Glosten *et al.* (1993)] and the NGARCH by[Engle & Ng (1993)] under Gaussian innovation. They proposed to calibrate the models on option prices with a proxy for the conditional variance can be extracted from the series of VIX using the following expression:

$$h_t^{\text{VIX}} = \frac{(1-\bar{\Psi})^T}{1-\bar{\Psi}^T}\left[\frac{\tau}{T}\left(\frac{\text{VIX}_t}{100}\right)^2 - h_0\left(1 - \frac{1-\bar{\Psi}^T}{(1-\bar{\Psi})}\right)\right] \tag{1.32}$$

where $\bar{\Psi}$ denote the volatility persistence under the risk-neutral measures and the h_0 unconditional long-term variance.

$$\hat{\theta}^* = \arg\underset{\theta}{\text{Min}} \sqrt{\frac{1}{N_{T_{Op}}}\sum_{t,i}\left(\frac{c_{i,t}\left(h_t^{\text{VIX}};\theta\right) - \hat{c}_{i,t}}{\hat{V}_{i,t}}\right)^2}. \tag{1.33}$$

[24]This price has been calculated using the FFT methodology shown in section 2.3 and it depends on the risk-neutral conditional volatility at time t, h_t^*, that is found from the log-returns and the risk-neutral GARCH updating rule prepared at its unconditional level.

[25]a call if $K_i > K_0$; a put if $K_i < K_0$; both call and put if $K_i = K_0$

It may be preferable to use the NLS approach rather than MLE for the purpose of option pricing, i.e. to estimate the parameters directly using information on option prices.To calibrate a model with the nonlinear leastsquares approach using option data, a large set of option contracts must be valued repeatedly to minimize the pricing error. If closed-form solutions are available for option prices and VIX index, computation is not necessarily a major problem. However, if a GARCH specification requires Monte Carlo methods, the NLS approach becomes computationally very demanding compared to Returns-MLE. NLS as it is computationally more expensive than MLE, especially if no closed-form solutions are available for option prices. Therefore, fast global optimization algorithms are crucial when calibrating volatility models using option data. moreover, the IVRMSE loss function may cause problems when using large option datasets because it is computationally intensive, as it requires inversion of the Black-Scholes formula at each step of the numerical search procedure.

The next subsection reviews several tools we need for option pricing valuation approachs: the Fast Fourier transform that can be used with closed formulas like in the case of GARCH setting in the spirit of [Heston & Nandi (2000)], the Monte Carlo Simulation method to simulate sample paths of the asset as a simple alternative to the parametric specification of the SDF, performance metric for option pricing to analyze the pricing error gap between various option pricing models.

1.4 Option pricing Valuation

Various techniques have been provided in the financial literature to answer the problem of valuing a European call under different assumptions of the underlying asset's model. In our subsequent discussion, the current stock price is set to be S_t and the price of a derivative instrument with terminal payoff $g(S_T)$ at the maturity T can be expressed in term of its payoff function by the following conditional[26] expectation :

$$C_t(g) = \mathbb{E}_{\mathbb{Q}}\left[e^{-r(T-t)}g(S_T)\Big|F_t\right], \tag{1.34}$$

for instance, the pay off of the European call option at the terminal condition is given by $(S_T - K)_+$.

In particular, if the characteristic function of the underlying is tractable in closed form, option prices can also be obtained by the power and versatility of Fourier Analysis using Fast Fourier transform. Once the dynamics of the log-returns under the risk-neutral measure $\mathbb{Q}$ of the GARCH type model is available with closed form formulas obtain from the characteristic function, the option valuation can be calculated by Fast Fourier transform introduced by [Carr & Madan (1999)]. They provided an efficient tool in producing call prices for European

[26] In this case, the Black and Scholes value of C_t for the call options on a non-dividend-paying stock is simply given by $C_t = S_t N(d_1) - Ke^{-r(T-t)}N(d_2)$ where $d_1 = \dfrac{\log\left(\frac{S_t}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}$, $d_2 = d_1 - \sigma\sqrt{T-t}$ and $N(d_1)$ is the cumulative probability distribution function for a standardized normal distribution.

option, this approach can manage complicated models with accuracy and without loss of efficiency. This next subsection describes the FFT algorithm.

1.4.1 The Fast Fourier transform (FFT)

First, we present the definition of the Fourier transform of a function and review some of its properties. The Fourier transform is a linear operator which transforms a function into a continuous range of its frequency components. Let f be a piecewise continuous real function over $]-\infty, +\infty[$ which satisfies the integrability condition:

$$\int_{-\infty}^{+\infty} |f(x)|\, dx < +\infty.$$

Définition 13 *The Fourier transform of f is defined by the following expression:*

$$\hat{f}(w) = \int_{-\infty}^{+\infty} f(t)e^{iwt}dt \qquad \forall w \in \mathbb{R}. \tag{1.35}$$

Given $\hat{f}(w)$, the function f can be recovered by the following Fourier inversion formula:

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \hat{f}(w)e^{-iwt}dw \qquad \forall w \in \mathbb{R}. \tag{1.36}$$

Then if f is a square integrable function, the Fourier transform of f and its inverse transform are well defined.

This section describes the fair value of the European call option. Let k be the log of the strike price K and s_t the log of the price S_t of the underlying asset and let $q_T(s)$ be the risk neutral density of the log price s of the underlying asset associated to the EMM $\mathbb{Q}$. The fair value $C_{T-t}(k)$ of the option at time t is related to the risk-neutral density $q_T(s)$ by :

$$\begin{aligned}
C_{T-t}(K) &= e^{-r(T-t)}\mathbb{E}_{\mathbb{Q}}\left[(S_T - K)_+ \middle| F_t\right] \\
C_{T-t}(K) &= e^{-r(T-t)}\mathbb{E}_{\mathbb{Q}}\left[\left(e^{\log(S_T)} - e^{\log(K)}\right)_+ \middle| \mathcal{F}_t\right] \\
C_{T-t}(K) &= \int_k^{+\infty} e^{-r(T-t)}(e^{\log(S_T)} - e^{\log(K)})q_{T-t}(s)ds \\
C_{T-t}(k) &= \int_k^{+\infty} e^{-r(T-t)}(e^{s_T} - e^k)q_{T-t}(s)ds
\end{aligned}$$

and using those equations, the initial call value of the option at time $t = 0$ of a European call option with strike K and maturity T is given by the following expression, as we see in [Chorro *et al.* (2015)] :

$$C_T(k) = \int_k^{+\infty} e^{-rT}\left(e^s - e^k\right) q_T(s)ds.$$

The square integrability of the call value function is required since we will compute the Fourier transform and its inverse. Unfortunately such expression is not integrable, condition that is necessary to apply the FFT. When the log strike price k converges toward to $-\infty$, we have:

$$\lim_{k\longrightarrow-\infty} C_T(k) = \int_{+\infty}^{-\infty} e^{-rT}\left(e^s\right) q_T(s)ds$$
$$\lim_{k\longrightarrow-\infty} C_T(k) = e^{-rT}\mathbb{E}_{\mathbb{Q}}\left[S_T|\mathcal{F}_0\right]$$
$$\lim_{k\longrightarrow-\infty} C_T(k) = S_0.$$

which is non zero and therefore $C_T(k)$ is not integrable. According to [Carr & Madan (1999)], in order to have a square integrable function, we are going to consider the modified call price :

$$c_T(k) = e^{\alpha k} C_T(k)$$

where $\alpha > 0$ is chosen in order to make $c_T(k)$ square-integrable. The Fourier transform of $c_T(k)$ can be expressed in -terms of the characteristic function of $\log(S_T)$ under $\mathbb{Q}$:

$$\Psi_T(w) = \int_{-\infty}^{+\infty} e^{iwk} c_T(k)dk$$
$$\Psi_T(w) = \int_{-\infty}^{+\infty} e^{iwk} \int_{k}^{+\infty} e^{\alpha k} e^{-rT}(e^s - e^k) q_T(s)dsdk$$
$$\Psi_T(w) = \int_{-\infty}^{+\infty} e^{-rT} q_T(s) \int_{-\infty}^{s} (e^{(\alpha+iw)k+s} - e^{(\alpha+1+iw)k})dkds$$
$$\Psi_T(w) = \int_{-\infty}^{+\infty} e^{-rT} q_T(s) e^{(iw+\alpha+1)s} \left(\frac{1}{iw+\alpha} - \frac{1}{1+iw+\alpha}\right) ds$$
$$\Psi_T(w) = \frac{e^{-rT}}{\alpha^2+\alpha+i(2\alpha+1)w-w^2} \int_{-\infty}^{+\infty} q_T(s) e^{(iw+\alpha+1)s} ds$$
$$\Psi_T(w) = \frac{\phi_T(w-i(\alpha+1))e^{-rT}}{\alpha^2+\alpha+i(2\alpha+1)w-w^2}$$

where $\phi_T(.)$ is the characteristic function of $\log(S_T)$ under the risk neutral probability $\mathbb{Q}$.

Proposition 1.4.1 *Let $\alpha > 0$, the Fourier transform of $c_T(k)$ exists if* $\mathbb{E}_{\mathbb{Q}}\left\{S_T^{\alpha+1}\right\} < +\infty$

Proof From the definition of characteristic function, we have :

$$|\phi_T(-i(\alpha+1))| = \left|\mathbb{E}_{\mathbb{Q}}\left\{e^{(-i(\alpha+1))i\log S_T}\right\}\right| = \left|\mathbb{E}_{\mathbb{Q}}\left\{e^{(\alpha+1)\log S_T}\right\}\right| = \mathbb{E}_{\mathbb{Q}}\left\{S_T^{(\alpha+1)}\right\}$$

thus we have the following equality :

$$\Psi_T(0) = \int_{-\infty}^{+\infty} c_T(k)dk = \frac{\phi_T(-i(\alpha+1))e^{-rT}}{\alpha^2+\alpha}$$
$$|\Psi_T(0)| = \frac{e^{-rT}\mathbb{E}_{\mathbb{Q}}\left\{S_T^{(\alpha+1)}\right\}}{\alpha^2+\alpha}$$

and $\mathbb{E}_{\mathbb{Q}}\left\{S_T^{(\alpha+1)}\right\} < +\infty$ implies $|\Psi_T(0)| < +\infty$. Therefore, $c_T(k)$ is well defined when the moment of order $(\alpha+1)$ of the underlying is finite. ■

The call price $C_T(k)$ can be recovered by taking the Fourier inversion transform:

$$C_T(k) = e^{-\alpha k} c_T(k) = \frac{e^{-\alpha k}}{2\pi} \int_{-\infty}^{+\infty} e^{-iuk} \Psi_T(u) du = \frac{e^{-\alpha k}}{\pi} \int_0^{+\infty} e^{-iuk} \Psi_T(u) du$$

by substitution, we obtain :

$$C_T(k) = \frac{e^{-\alpha k}}{\pi} \int_0^{+\infty} e^{-iwk} \left[\frac{\phi_T(w - i(\alpha+1)) e^{-rT}}{\alpha^2 + \alpha + i(2\alpha+1)w - w^2} \right] dw$$

Concerning the FFT implementation, we start with the choice on the number of intervals N equidistant points and the step-width Δw. A numerical Riemann approximation for $C_T(k)$ is given by :

$$C_T(k) \approx \frac{e^{-\alpha k}}{\pi} \sum_{j=0}^{N-1} e^{-iu_j k} \Psi_T(u_j) \Delta u$$

where $u_j = j\Delta u = j\delta$ and $\Delta u = \delta = \frac{a}{N}$. The FFT returns N values of k and we employ a regular spacing of size λ, so which gives us log strike levels ranging from $-b$ to b, where $\lambda = \frac{2b}{N}$. Thus

$$C_T(k_l) \approx \frac{e^{-\alpha k_l}}{\pi} \sum_{j=0}^{N-1} e^{-ij\delta(-b+\lambda l)} \Psi_T(j\delta)\delta \approx \frac{e^{-\alpha k_l}}{\pi} \sum_{j=0}^{N-1} e^{-ij\lambda\delta l} e^{ij\delta b} \Psi_T(j\delta)\delta$$

$$C_T(k_l) \approx \frac{e^{-\alpha k_l}}{\pi} \sum_{j=0}^{N-1} e^{-i\delta\lambda jl} e^{ij\delta b} \Psi_T(j\delta)\delta.$$

where $\delta\lambda = \frac{2\pi}{N}$. We have succeeded in extracting the analytical expression of the Fast Fourier transform of the faire value of the the call price of the European option. However, the choosing of the parameters and the algorithm steps need to be carefully studied in order to have accurate results. In fact, the freedom of choice of the damping coefficient and the integration path affect the accuracy of the method.

On the other hand, when characteristic functions are not avalaible in a closed expression, we can use Monte Carlo simulations of independent realizations $S_{n,t}$ of the process S_t and approximate the conditional expectation in equation 1.34. This method is particularly useful because of its ability to estimate integrals and the efficiency of Monte Carlo simulation increases with the number of paths used in the simulations.

1.4.2 Monte Carlo and Empirical Martingale Simulation Method

Theoretical foundation of Monte Carlo methods are mainly based on two fundamentals asymptotic results : the Strong Law of Large Numbers and the Central Limit Theorem.

The Strong Law of Large Numbers predicts that, under integrability conditions, the mean of a sequence of i.i.d random variables converges toward the expectation as the sample size increases.

Theorem 1.4.1 *Let $(X_i)_{i\in\mathbb{N}^*}$ be a sequence of i.i.d random variables.*

- *Suppose that $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, we have :*

$$\frac{X_i + \cdots + X_n}{n} \xrightarrow{a.s \text{ and } L^1} \mathbb{E}_{\mathbb{P}}[X]$$

- *if $\mathbb{E}_{\mathbb{P}}[|X|] = +\infty$, the sequence $\sum_{i=1}^{n} X_i$ diverges almost surely.*

When we analyse a method, there are three particularly important considerations: bias, variance and computing time.

Theorem 1.4.2 *Let $(X_i)_{i\in\mathbb{N}^*}$ be a sequence of i.i.d random variables. Suppose that $X_1 \in L^2(\Omega, \mathcal{F}, \mathbb{P})$, then*

$$\frac{(X_1 + \cdots + X_n) - nm}{\sqrt{n}\sigma} \xrightarrow{\mathcal{D}} N(0,1)$$

where $m = \mathbb{E}_{\mathbb{P}}[X]$ and $\sigma^2 = Var[X]$.

For instance, based on the expectation in equation 1.34, the Monte Carlo estimate of the call value $C_{T-t}(g)$ of the conditional distribution of the S_T given $\mathcal{F}_t$ under $\mathbb{Q}$ can be computed by the following procedure :

1. Generate $(S_{t,i})_{i\in\{1,\cdots,N\}}$ a N-sample of the conditional distribution of S_T given $\mathcal{F}_t$ under $\mathbb{Q}$,
2. Compute the estimate price by using the following expression :

$$C_{T-t}(g) \approx \frac{1}{N}\sum_{n=1}^{N} e^{-r(T-t)} g(S_{t,n})$$
$$C_{T-t}(g) \approx \frac{1}{N}\sum_{n=1}^{N} e^{-r(T-t)} (S_{t,n} - K)_+ \qquad \text{for} \qquad g(x) = (x-K)_+$$

and from the central limit theorem, it is possible to obtain the confidence interval associated to the option price whose length is proportional to the variance of the estimator :

$$\frac{C_{T-t}(g)}{e^{-r(T-t)}} \in \left[\frac{\sum_{n=1}^{N}(S_{t,n}-K)_+}{N} - \frac{1.96\hat{\sigma}_N}{\sqrt{N}}, \frac{\sum_{n=1}^{N}(S_{t,n}-K)_+}{N} + \frac{1.96\hat{\sigma}_N}{\sqrt{N}}\right]$$
$$\text{where} \quad \hat{\sigma}_N^2 = \frac{N}{N-1}\left(\frac{\sum_{n=1}^{N}\left((S_{t,n}-K)_+\right)^2}{N} - \left(\frac{\sum_{n=1}^{N}(S_{t,n}-K)_+}{N}\right)^2\right).$$

We can increase the efficiency of the Monte Carlo simulation by reducing the variance of the estimator. There is a efficient variance reduction technique, the so-called Empirical

Martingale Simulation Method (EMS). Moreover, as details in [Duan & Simonato (1998)]. Classical no arbitrage bounds are often violated once the classical Monte-Carlo approximation is used, in particular the martingale condition is not empirically verified :

$$\mathbb{E}_{\mathbb{Q}}\left[e^{-r(T-t)}S_T \middle| F_{t-1}\right] \approx C_{T-t}\left(Id\right) \neq S_t.$$

To overcome the preceding problem and in order to improve the efficiency of the Monte Carlo price estimators, [Duan & Simonato (1998)] have proposed the EMS method that is a powerful and simple multiplicative adjustment to Monte Carlo simulation. In the following, we describe the EMS procedure for option pricing:

1. Use the standard Monte Carlo simulation method to generate N independent realisation of S_T under $\mathbb{Q}$, denoted by $(S_{T,i})_{i\in\{1,\cdots,N\}}$,
2. the empirical martingale adjustement modifies the simulated sample paths as follows :

$$\tilde{S}_{T,i} = \frac{S_{T,i}}{\frac{1}{N}\sum_{i=1}^{N} S_{T,i}} S_t e^{-r(T-t)}$$

3. compute the option prices estimator $C_{t,n}^{EMS}$ of a European call option with a payoff function $g\left(x\right) = \left(x-K\right)_+$, strike K and maturity T by

$$C_{t,n}^{EMS} = \frac{e^{-r(T-t)}}{n}\sum_{i=1}^{n}\left(\tilde{S}_{T,i}-K\right)_+$$

For more details, [Chorro *et al.* (2012)] proposed an explicit study of the empirical pricing performance of the empirical martingale simulation method.

After fitting a variety of forcasting models for option valuation, when more than one forecasting technique seems reasonable for a particular application, then all the models can be compared and evaluated on the basis of the pricing errors generated by them. One can subtract the forecast value from the observed value of the price and obtain a measure of error. The forecast accuracy measures can also be used to rank pricing models. Mean Error (ME), Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Square Error (RMSE), Mean Percentage Error (MPE) and Mean Absolute Percentage Error (MAPE) are used as forecast accuracy measures.

1.4.3 Performance Measures

All the models are compared on the basis of the forecasting errors generated by them. In order to evaluate the pricing forecast performance, and also to order the predictions, the literature related to the evaluation of option pricing forecast have developed several measures of accuracy. The accuracy of a forecast performance refers to how well a given forecasting technique can guess the value of the predicted attribute for new or previously unseen data :

Définition 14 *Let* $(C_i)_{i\in 1,\cdots,n}$ *be a datasetof real market observations and* $\left(\hat{C}_i\right)_{i\in 1,\cdots,n}$ *the associated forecasts obtained from a given model.*

- *Then the mean error (ME) is given by:.*

$$ME = \frac{1}{n}\sum_{i=1}^{n}\left(C_i - \hat{C}_i\right)$$

- *The average absolute differences is taken to obtain the mean absolute error (MAE).*

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|C_i - \hat{C}_i\right|$$

The mean absolute error (MAE) or mean absolute deviation (MAD) is calculated by taking the absolute value of the difference between the estimated forecast and the actual value so that the negative values do not compensate the positive values.

- *The Mean squared error (MSE) is taken to measure the variability in forecast errors:*

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(C_i - \hat{C}_i\right)^2$$

- *The root mean square error (RMSE) measures the geometric average magnitude of the square error.*

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(C_i - \hat{C}_i\right)^2}$$

- *The percentage error (MPE) is the relative of error at a particular point of time in the series.*

$$MPE = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{C_i - \hat{C}_i}{C_i}\right)$$

The average percentage error in the entire series is a general measure of fit useful in comparing the fits of different models.

- *Because the positive and negative errors may tend to cancel themselves, MPE statistic is often replaced by the mean absolute percentage error (MAPE).*

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{C_i - \hat{C}_i}{C_i}\right|$$

The closer MAPE approaches zero, the better the forecasting results.

An important aspect of the forecast metrics used for price evaluations is their capability to rank among model results. The more discriminating measure that produces higher variations in its model performance metric among different sets of model results is often the more desirable. In this regard, the MAE might be affected by a large amount of average error values without adequately reflecting some large errors. Giving higher weighting to the unfavorable conditions, the RMSE usually is better at revealing model performance differences.

The sensitivity of the RMSE to outliers is the most common concern with the use of this metric. In fact, the root mean squared error is more sensitive than other measures to the occasional large error: the squaring process gives disproportionate weight to very large errors. The RMSE gives a relatively high weight to large errors. This means the RMSE is most useful when large errors are particularly undesirable. If an occasional large error is not a problem in decision situation, then the MAE or MAPE may be a more relevant criterion. Furthermore, in the data assimilation field, the sum of squared errors is often defined as the cost function to be minimized by adjusting model parameters. In such applications, penalizing large errors through the defined least-square terms proves to be very effective in improving model performance. Under the circumstances of calculating model error sensitivities or data assimilation applications, MAE are definitely not preferred over RMSE .

1.5 Outline of the Thesis

This thesis makes the following contributions to the literatures :

1. In an important paper, [Christoffersen *et al.* (2006)] proposed an option pricing model based on an IG-GARCH process and the conditional Esscher transform to underline the importance of modelling conditional skewness. One of the main features of this approach is to provide, as in [Heston & Nandi (2000)], semi-closed form formulas for call options but for non Gaussian innovations. Recently, the monotonicity of the stochastic discount factor (often supposed to be exponential affine of the log-returns) was discussed in the literature (see for example [Christoffersen *et al.* (2006)] and [Monfort & Pégoraro (2012)]) to favor U shapes. In this first paper, we have explored an extension of [Christoffersen *et al.* (2006)] using an U-shaped pricing kernel that increases the flexibility of the link between the historical and the risk-neutral distributions while preserving the tractability of the model. Our empirical results are clear, the in and out of sample pricing performances of the IG-GARCH are improved by the choice of this new pricing kernel. What is more, we show in this framework that an estimation strategy based on returns-VIX information provides very interesting pricing errors at a low computational cost because expensive calibration on options can be bypassed.

2. The second chapter of the present thesis derives from a very simple finding: under Gaussian hypotheses, some GARCH-type models have outstanding properties (closed-form expressions for the VIX and/or option prices) that fail when NIG innovations are involved. Nevertheless, it is now well documented that Gaussian GARCH option pricing models produce poor pricing errors when compared with skewed and fat-tailed counterparts. Thus, inspired by the so-called quasi-maximum likelihood estimator, a new two-steps approach is provided to both take benefit of these remarkable features in Gaussian environment and work with more realistic distributions. This strategy

estimates separately the volatility and the distribution parameters supposing Gaussian innovations in the first step to incorporate VIX or options information in the estimation process. In a second step, the NIG distribution is fitted from the residuals obtained in the previous stage. What is more, we provide an empirical test for our new estimation methodology on a large dataset of options written on the S&P500.

3. This final chapter attempts to fill several gaps in the GARCH option pricing literature, in particular, from an empirical point of view. Firstly, in the spirit of [Christoffersen *et al.* (2004)] the aim of our study is to provide an intensive comparison analysis of empirical performances, in VIX index or options valuation, between different GARCH type models using Gaussian or non-Gaussian distributions under different classes of risk neutral measures. Furthermore, particular attention is granted on the choice of the information set (VIX, options, returns) in the estimation process. As a natural non-Gaussian alternative we favor the so-called NIG distribution not only because it is known to fit the statistical properties of asset returns remarkably but also because, combined with the Esscher and EGP SDF, the pricing equations may be solved explicitly. What is more, monotonic and non-monotonic pricing kernels ([Monfort & Pégoraro (2012)], [Chorro & Fanirisoa (2016)] are considered for Gaussian and IG distributions. To our knowledge, in the existing literature empirical studies asked about, in general, the impact of the distribution ([Christoffersen *et al.* (2006)], [Chorro *et al.* (2012)]), the choice of the SDF ([Badescu *et al.*(2011)], [Christoffersen *et al.* (2013)], [Chorro & Fanirisoa (2016)]) or the estimation strategy ([Hao & Zhang (2013)], [Kanniainen *et al.* (2014)], [Papantonis (2016)], [Lalancette & Simonato (2017)]) on pricing performances, but few of them consider all these factors at the same time. Our study is a mean of making a contribution to understand the global impact of these complementary aspects (24 combinations of GARCH/distribution/SDF/estimation are tested). Secondly, inspired by the work of [Hao & Zhang (2013)] that proposes to explain the poor pricing performances of Gaussian GARCH models by their inefficiency to capture the variance risk premium, we also explore in this chapter if it is possible to partly classify GARCH option pricing models by their ability to simply reproduce the VIX index. From purely numerical aspects, such a conclusion would be very interesting to backtest these models in an efficient way only using VIX information, when available, instead of complex option datasets.

Chapter 2

Option valuation with IG-GARCH model and a U-shaped pricing kernel.

Contents

In the financial literature, ARCH/GARCH models, introduced by [Engle (1982)] and [Bollerslev (1986)], have gained widespread acceptance over the last few decades to model the heteroscedasticity of asset returns. They emerged as one of the most popular and flexible discrete time alternatives to continuous time diffusions because the endogenous parametric specification of volatility makes it possible to estimate the joint dynamics of returns and volatility using only the time series of returns. From this seminal step, GARCH models have been extended in various directions to cope, in particular, with asymmetry properties (see e.g. [Terasvirta (2009)] for a recent survey). Recently, [Duan (1995)] was the first to provide a coherent theoretical framework, the so-called Local Risk Neutral Valuation Relationship (LRNVR), to price contingent claims when the underlying dynamics is given by a GARCH model with Gaussian innovations. While this approach outperforms the [Black & Scholes (1973)] benchmark, it is restricted to Gaussian innovations and the prices are obtained using Monte Carlo simulations. Following the preceding methodology, [Heston & Nandi (2000)] considered a new conditionally-normal GARCH-like volatility updating scheme able to cope with skews in option prices. Moreover, they derived an interesting semi-closed form expression for call option prices, making the pricing of such financial products fast and compatible with calibration estimation methods at a reasonable computational cost. Nevertheless, as this model is conditionally Gaussian, it usually fails to capture the short term behavior of equity option smiles. In fact, it is now well-documented ([Chorro *et al.* (2015)], p.41) that Gaussian innovations cannot take into account all the mass in the tails and the asymmetry that characterize the distribution of daily log-returns even if an asymmetric GARCH filter is applied.[1]

During the last decade (see [Chorro *et al.* (2015)], Chap.3 for a recent survey), researchers have intensively investigated the way to extend the Duan's option pricing model to incorporate the skewness and leptokurtosis observed in financial datasets into GARCH residuals. In general, such a choice is motivated by equilibrium arguments (see [Badescu *et al.*(2009)]) and/or by its compatibility with the myriad of possible candidates for the distribution. An important contribution in this direction was the work of [Siu *et al.* (2004)] in which the authors used for the first time, in the GARCH setting, the conditional Esscher transform introduced in [Bühlmann *et al.* (1996)] to price European options using a shifted Gamma distribution. This approach is equivalent (see [Gouriéroux & Monfort (2007)]) to considering a special parametric form for the pricing kernel (exponential-affine of the log-returns) and allows for explicit and tractable risk-neutral dynamics in many situations.[2] The flexibility of the exponential-affine parameterization is probably one of its main advantages with respect to its natural competitors as the generalized LRNVR of [Duan (1999)] (see also [Stentoft (2008)] and [Simonato & Stentoft (2015)]) or the extended Girsanov principle of [Elliott & Madan (1998)] (see also [Badescu *et al.*(2008)] and [Badescu *et al.*(2011)]). Nevertheless, in spite of their differences, all the preceding specifications coincide with the LRNVR in the Gaussian framework and depend on a single stochastic parameter related to the equity risk premium and uniquely determined by the martingale constraints.

The choice of a flexible characterization for the pricing kernel is an old topic (see [Rubinstein (1976)], [Brennan (1979)]. [Campbell *et al.* (1997)], [Cochrane (2001)], [Ross (1978)], and [Harrison& Kreps (1979)] among others) that often leads to parametric forms that are monotonic functions of the log-returns ([Rubinstein (1976)], [Hansen & Singleton (1982)], [Hansen & Singleton (1983)], [Gerber & Shiu (1994)], and [Bühlmann *et al.* (1996)]). However, many recent empirical studies suggest evidence against the monotonicity assumption ([Bates (1996)], [Bakshi *et al.*(1997)], [Ziegler (2007)], [Chabi-Yo *et al.* (2008)], and [Bakshi *et al.*(2015)]). In the GARCH setting, two approaches have been proposed to overcome this problem and take into account market and volatility risks: [Monfort & Pégoraro (2012)] introduced an extension of the classical Esscher transform, including a quadratic term in the pricing kernel while [Christoffersen *et al.* (2013)] proposed a variance dependent pricing kernel (see also [Badescu *et al.*(2015)] for a slightly

[1]Concerning asymmetric volatility responses, refer to the EGARCH model introduced in [Nelson (1991)], the GJR GARCH model of [Glosten *et al.* (1993)], the APARCH model developed in [Ding *et al.* (1993)], as well as the TGARCH studied in [Zakoian (1994)].

[2]See among others, [Christoffersen *et al.* (2006)] for the Inverse Gaussian distribution, [Badescu *et al.*(2008)] for the mixture of Gaussian distributions, [Chorro *et al.* (2012)] for the Generalized hyperbolic distribution.

different approach compatible with non-affine models).

In this chapter, we propose an extension of the so-called Inverse Gaussian GARCH (IG-GARCH) model of [Christoffersen *et al.* (2006)] where the authors provide a new particular affine GARCH structure with Inverse Gaussian innovations to take conditional skewness into account. Using the pricing kernel derived from the conditional Esscher transform, they obtained the risk neutral dynamics, depending only on historical parameters, which gave rise to a closed-form option pricing formula as in [Heston & Nandi (2000)].

The main idea is to use an extended and non-monotonic version of the exponential-affine pricing kernel, particularly well-adapted to the Inverse Gaussian distribution, in order to increase the flexibility of the link between the historical and the risk-neutral distributions while preserving the tractability of the model. In fact, even in the case of our[3] new pricing kernel, closed-form expressions remain available for European call options and the VIX index.[4] Therefore, it is possible to combine, at a reasonable computational cost, historical returns dynamics with options or VIX information in the estimation process to build more accurate joint likelihood as explained in [Christoffersen *et al.* (2012)] and [Kanniainen *et al.* (2014)].

Finally, we perform a GMM test to check the validity of each pricing kernel with respect to the martingale conditions and present a comparative analysis of in-sample and out-of-sample pricing performances of the IG-GARCH model associated with both exponential-affine and exponential U-shaped pricing kernels and estimated using options or VIX information. We compute the Implied Volatility Root Mean Square Error (IVRMSE) for each model to evaluate and compare the pricing errors. This empirical study provides strong evidences indicating that the exponential U-shaped pricing kernel is clearly superior in approximating the price of options written on the S&P500 for the concerned period. What is more, we show, in this framework that an estimation strategy based on returns-VIX information provides very interesting pricing errors at a low computational cost because expensive calibration on options can be bypassed.

The remainder of the chapter is organized as follows. The next section defines and develops the theoretical framework, giving, in particular, the risk neutral dynamics under the two different pricing kernels and the associated closed form expressions for option prices and the VIX index. We present in Section 3 the methods of estimation based on different joint maximum likelihood. The numerical results are contained in Section 4. More precisely, we describe the returns, VIX and options datasets on the S&P500 used in the chapter, we perform a GMM test to validate the martingale conditions and, finally, provide the in and out-of-sample pricing performances. Concluding remarks are given in Section 5.

[3]The new form of the pricing kernel has been inspired by the work of [Monfort & Pégoraro (2012)] which introduces a second-order Esscher transform particularly well-adapted to the Gaussian (or mixture of Gaussian) case. In the IG-GARCH setting, the idea is to replace in the pricing kernel, the quadratic term of [Monfort & Pégoraro (2012)] by a hyperbolic one that is more suitable for our choice of distribution.

[4]The VIX expresses the market expectations of the 30-day volatility implied in equity index options.

2.1 The stock price dynamics and the stochastic discount factors

This section presents the theoretical framework of the present chapter. Our study uses as a core model the inverse Gaussian GARCH (IG-GARCH) model of [Christoffersen *et al.* (2006)] known to cope with conditional skewness as well as conditional heteroskedasticity and a leverage effect. First, let us briefly review the main lines of this approach that will be used in the following as a keystone to price options written on the S&P500 index using different pricing kernels.

2.1.1 The stock price dynamics under the physical probability measure $\mathbb{P}$

We consider a discrete time economy with a time horizon $T \in \mathbb{N}^*$ consisting of a risk-free zero-coupon bond (associated with the risk-free rate r expressed on a daily basis and supposed to be constant) and a stock (the risky asset). Following [Christoffersen *et al.* (2006)], we assume that, under the physical probability measure $\mathbb{P}$, the logarithm of the returns of the stock price process $(S_t)_{t\in\{0,\ldots,T\}}$ fulfills

$$\begin{cases} Y_{t+1} = \log\left(\frac{S_{t+1}}{S_t}\right) & = \quad r + \nu h_{t+1} + \eta y_{t+1} \\ h_{t+1} & = \quad w + b h_t + c y_t + a\frac{h_t^2}{y_t} \end{cases} \tag{2.1}$$

with $a_0 > 0$, $a_1 \geq 0$, $b_1 \geq 0$ and where the $(y_t)_{t\in\{1,\ldots,T\}}$ are random variables generating an information filtration denoted by $(\mathcal{F}_t)_{t\in\{0,\ldots,T\}}$ where $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $(\mathcal{F}_t = \sigma(y_u; 1 \leq u \leq t))_{t\in\{1,\ldots,T\}}$. Moreover, we suppose that, given $\mathcal{F}_{t-1}$, y_t follows an Inverse Gaussian distribution with degree of freedom $\delta_t = \frac{h_t}{\eta^2}$.[5] Traditionally, the moment generating function[6] of the pair $(y_t, \frac{1}{y_t})$ can be expressed as:

$$\mathbb{E}\left[e^{\theta y_t + \frac{\phi}{y_t}}\right] = \frac{\delta_t}{\sqrt{\delta_t^2 - 2\phi}} e^{\left[\delta_t - \sqrt{(\delta_t^2 - 2\phi)(1-2\theta)}\right]} \tag{2.2}$$

from which we deduce that

$$E[Y_t \mid \mathcal{F}_{t-1}] = r + (\nu + \frac{1}{\eta})h_t, \; Var[Y_t \mid \mathcal{F}_{t-1}] = h_t$$

and

$$Cov[Y_t - Y_{t-1}, h_{t+1} - h_t \mid \mathcal{F}_{t-1}] = Cov[Y_t, h_{t+1} \mid \mathcal{F}_{t-1}] = (\frac{c}{\eta} - \eta^3 a)h_t.$$

In particular, h_t is the conditional variance of the log-returns and 2.1 may be considered as a GARCH-type model of conditional volatility accommodating with asymmetric volatility

[5] There exist in the literature different parameterizations of the Inverse Gaussian distribution. In this chapter, the definition and properties of the Inverse Gaussian distribution are presented along the lines of [Johnson *et al.* (1994)] and [Barndorff-Nielsen (1998)]. In particular, the associated density function is given by the one parameter family: $\mathbf{1}_{\{y>0\}} \frac{\delta}{\sqrt{2\pi y^3}} e^{-\left(\sqrt{y}-\delta/\sqrt{y}\right)^2/2}$ where $\delta \in \mathbb{R}_+^*$ and we have $\mathbb{P}(y_t = 0) = 0$.

[6] Having option pricing in mind, the existence and the simple expression of the moment generating of the Inverse Gaussian distribution will be fundamental to using the so-called Esscher transform (and the variant presented in this chapter) to specify the stochastic discount factors.

responses. We refer the reader to [Christoffersen *et al.* (2006)] for an in-depth discussion on the statistical characteristics of this process.

To conclude the presentation of the historical dynamics, let us review one of the key feature of the IG-GARCH model (that may be seen in this way as a skewed analogous of the [Heston & Nandi (2000)] model): the historical conditional moment generating function of $log(S_T)$ may be expressed using backward recursive equations.

Proposition 2.1.1 *(See [Christoffersen et al. (2006)] Appendix A) Given $\mathcal{F}_t$, the moment generating function under $\mathbb{P}$ of $log(S_T)$ is characterized by:*

$$\mathbb{G}^{\mathbb{P}}_{log(S_T)|\mathcal{F}_t}(\phi) = \mathbb{E}\left[S(T)^{\phi} \mid \mathcal{F}_t\right] = S(t)^{\phi} exp\left[A(t) + B(t)\left(w + bh_t + cy_t + a\frac{(h_t)^2}{y_t}\right)\right]$$

with $A(T) = B(T) = 0$ and

$$\begin{array}{l} A(t) = A(t+1) + \phi r + wB(t+1) - \frac{1}{2}\log(1 - 2a(\eta)^4 B(t+1)) \\ B(t) = bB(t+1) + \phi\nu + (\eta)^{-2} - (\eta)^{-2}\sqrt{(1 - 2a(\eta)^4 B(t+1))(1 - 2cB(t+1) - 2\phi\eta)}. \end{array}$$

This property of the conditional moment generating function will be used in the option pricing analysis to obtain prices using the fast Fourier transform methodology.

2.1.2 Two stochastic discount factors and the related risk-neutral dynamics

When we have option pricing in mind, conditional distributions of returns and volatility specifications are not the only issues we should pay attention to. In fact, the use of realistic discrete time volatility structures and continuous distributions gives rise to incompleteness and equivalent martingale measures are not unique in general. It is a conventional knowledge that in the discrete time setting the construction of such a probability measure is equivalent to the specification of a one-period stochastic discount factor process (see for example [Chorro *et al.* (2015)], Chap. 3.2.2). The purpose of this section is to present two approaches compatible with the dynamics introduced in 2.1 in order to obtain tractable risk-neutral processes. The first one, due to [Bühlmann *et al.* (1996)], and first applied in the GARCH setting by [Siu *et al.* (2004)], is based on the conditional extension of the [Esscher (1932)] transform used by [Gerber & Shiu (1994)] to price contingent claims in continuous time. The second and new one, inspired by the second-order Esscher transform introduced by [Monfort & Pégoraro (2012)] for Gaussian GARCH models, induces more flexibility in the definition of the stochastic discount factor and permits to obtain different realistic shapes.

2.1.2.1 The exponential-affine stochastic discount factor

The conditional Esscher transform introduced by [Bühlmann *et al.* (1996)] has been a major innovation in the discrete time financial literature providing a flexible framework to price European derivatives. In the GARCH setting it has been combined, with empirical successes, with various families of distributions such as Gaussian jumps in [Duan *et al.* (2005)] and [Duan *et al.* (2006)], mixture of Gaussian distributions in [Badescu *et al.*(2008)], or

Generalized Hyperbolic distributions in [Chorro *et al.* (2012)]. This approach is equivalent (see [Gouriéroux & Monfort (2007)]) to considering a stochastic discount factor that is exponential-affine of the log-returns:[7]

$$\forall t \in \{1, \cdots, T\}, \qquad M_t^{ess} = e^{\theta_t Y_t + \varepsilon_t},$$

where θ_t and ε_t are $\mathcal{F}_{t-1}$-measurable random variables that may be uniquely obtained, under mild conditions, from the pricing equations[8]

$$\begin{cases} \mathbb{E}_{\mathbb{P}}\left\{e^r M_t^{ess} \mid \mathcal{F}_{t-1}\right\} = 1 \\ \mathbb{E}_{\mathbb{P}}\left\{e^{Y_t} M_t^{ess} \mid \mathcal{F}_{t-1}\right\} = 1. \end{cases} \tag{2.3}$$

The equivalent martingale measure associated with $(M_t^{ess})_{t\in\{1,\cdots,T\}}$ is denoted by $\mathbb{Q}^{ess}$ and in the framework of the IG-GARCH model 2.1 introduced in the preceding section we obtain the following proposition that perfectly describes the risk-neutral dynamics under $\mathbb{Q}^{ess}$:

Proposition 2.1.2 *(See [Christoffersen et al. (2006)] Appendix B) Assuming that the process $(Y_t)_t$ is defined by 2.1, then,*

a) $\forall t \in \{1, \cdots, T\}$, the system (2.3) admits a unique solution $(\theta_t^*, \varepsilon_t^*)$ characterized by:

$$\theta_t^* = \theta^* = \tfrac{1}{2}\left[\eta^{-1} - \tfrac{1}{\nu^2\eta^3}\left[1 + \tfrac{\nu^2\eta^3}{2}\right]^2\right]$$
$$\varepsilon_t^* = -r(\theta^* + 1) - \theta^*\nu h_t - \left[\delta_t\left(1 - \sqrt{(1 - 2\theta^*\eta)}\right)\right].$$

b) Under $\mathbb{Q}^{ess}$, the process $(Y_t)_t$ is again an IG-GARCH model with changed parameters:

$$\begin{cases} Y_{t+1} = \log\left(\frac{S_{t+1}}{S_t}\right) &= r + \nu^* h_{t+1}^* + \eta^* y_{t+1}^* \\ h_{t+1}^* &= w^* + b h_t^* + c^* y_t^* + a^* \frac{(h_t^*)^2}{y_t^*} \end{cases} \tag{2.4}$$

$$\textit{where} \qquad \nu^* = \nu\left(\frac{\eta^*}{\eta}\right)^{-\frac{3}{2}}, \quad y_{t+1}^* = y_{t+1}\left(\frac{\eta^*}{\eta}\right)^{-1},$$
$$w^* = w\left(\frac{\eta^*}{\eta}\right)^{\frac{3}{2}}, \quad c^* = c\left(\frac{\eta^*}{\eta}\right)^{\frac{5}{2}}, \quad a^* = a\left(\frac{\eta^*}{\eta}\right)^{-\frac{5}{2}},$$

with $\eta^* = \dfrac{\eta}{1 - 2\theta^*\eta}$ and where, given $\mathcal{F}_{t-1}$, y_t^* follows an Inverse Gaussian distribution with degree of freedom $\delta_t^* = \frac{h_t^*}{(\eta^*)^2}$.

We remark, from the preceding proposition, that the conditional dynamics under $\mathbb{Q}^{ess}$ is the same as under the historical probability with changed parameters and that the risk-neutral conditional variance can be expressed as $h_{t+1}^* = (\eta^*/\eta)^{\frac{3}{2}} h_{t+1}$.[9] One important empirical consequence for the pricing of European call and put options is that proposition 2.1.1 remains valid under $\mathbb{Q}^{ess}$, thus semi-closed form formulas will be available for prices.

[7] This exponential-affine restriction of the stochastic discount factor is equivalent to the assumption (12) of [Christoffersen *et al.* (2006)].

[8] The equations are derived by applying the pricing formula to the risk-free and risky assets.

[9] Contrary to what happens for Gaussian GARCH models, the IG-GARCH framework is able to cope with the well-known stylized fact that the risk-neutral variance is in general greater than the historical one.

Even if the assumption of an exponential-affine stochastic discount factor is well theoretically justified in the literature, in particular in equilibrium pricing models (see [Badescu *et al.*(2009)]), it is not the only issue to obtain arbitrage-free price processes that derive from the pricing equations (2.3). Therefore, in the next subsection, we are going to see how to extend the exponential-affine pricing kernel M_t^{ess} in order to increase the flexibility of the link between the historical and the risk-neutral distributions while preserving the tractability of the model.

2.1.2.2 The exponential U-shaped stochastic discount factor

We derive in this subsection the risk-neutral dynamics of the IG-GARCH model using an exponential U-shaped pricing kernel that extends the classical conditional Esscher transform. Inspired by the second-order Esscher transform recently introduced by [Monfort & Pégoraro (2012)] in the Gaussian setting, we include the term $\frac{\rho_t}{y_t}$ in the specification of M_t^{ess} to be able to generate an exponential U-shaped function:

$$\forall t \in \{0, \cdots, T\}, \qquad M_t^{Ushp} = e^{\theta_t Y_t + \varepsilon_t + \frac{\rho_t}{y_t}} = e^{\theta_t Y_t + \varepsilon_t + \frac{\eta \rho_t}{Y_t - r - \nu h_{t+1}}}$$

where θ_t, ε_t and ρ_t are $\mathcal{F}_{t-1}$ measurable random variables.[10] Under the risk-neutral probability $\mathbb{Q}^{Ushp}$ associated with $(M_t^{Ushp})_{t\in\{1,\cdots,T\}}$, the overall dynamics of the log-return is, once again similar the historical one:

Proposition 2.1.3 *(See Appendix) $\forall t \in \{1, \cdots, T\}$, if we assume a constant proportional wedge between h_t and h_t^* (i.e $h_t^*/h_t = \pi$) the dynamics of Y_t under $\mathbb{Q}^{Ushp}$is of the form:*

$$\begin{cases} Y_{t+1} = \log\left(\frac{S_{t+1}}{S_t}\right) & = & r + \nu^* h_{t+1}^* + \eta^* y_{t+1}^* \\ h_{t+1}^* & = & w^* + b h_t^* + c^* y_t^* + a^* \dfrac{(h_t^*)^2}{y_t^*} \end{cases} \tag{2.5}$$

where $$\nu^* = \frac{\nu}{\pi}, \quad w^* = w\pi, \quad c^* = \frac{c\pi\eta^*}{\eta}, \quad a^* = \frac{a\eta}{\pi\eta^*},$$
$$\eta^* = \sqrt[3]{\frac{\pi^2}{\nu^2}\left(-1 + \sqrt{1 + \frac{8\nu}{27\pi}}\right)} + \sqrt[3]{\frac{\pi^2}{\nu^2}\left(-1 - \sqrt{1 + \frac{8\nu}{27\pi}}\right)},$$

and where, given $\mathcal{F}_t$, y_{t+1}^* follows an IG distribution with degree of freedom $\delta_{t+1}^* = \dfrac{h_{t+1}^*}{(\eta^*)^2}$.

As before, we obtain a similar IG-GARCH structure for the risk-neutral dynamics of the log-returns. Of course, if $\forall t \in \{1, \cdots, T\}$, we impose $\rho_t^* = 0$, we recover the result of the proposition 2.1.2. Nevertheless, the risk-neutral dynamics given by proposition 2.1.2 only depends on the initial historical set of parameters while the dynamics presented in proposition 2.1.3 introduces a risk-neutral parameter π. Therefore, the first model may

[10] From (2.1), we obtain $M_t^{Ushp} = e^{\theta_t \eta y_t + \frac{\rho_t}{y_t} + \varepsilon_t + \theta_t(r + \nu h_t)}$. In the empirical exercise performed in section 4, we obtain, independently of the estimation process, $\eta < 0$, $\theta_t < 0$ and $\rho_t > 0$. Therefore, $\lim\limits_{y_t \to 0^+} M_t^{Ushp} = \lim\limits_{y_t \to +\infty} M_t^{Ushp} = +\infty$ and M_t^{Ushp} is a U-shaped function of y_t.

be directly estimated from returns using a conditional version of the classical maximum likelihood (ML) estimation while an extra information (based on option prices) is needed for the estimation of the second one. In the next two subsections we show how to include this extra information in an efficient way in the estimation strategy. More precisely, we show that for the two risk-neutral IG-GARCH models we have numerically efficient closed form expressions not only for the price of European call options but also for the VIX index at any time.

2.1.3 Pricing European call options using Fast Fourier Transform (FFT)

It is well-known from the pioneering work of [Heston (1993)] that the price of European call options may be expressed using the risk-neutral conditional moment generating function of $log(S_T)$ (see also [Chorro *et al.* (2015)], p. 184): if $\mathbb{Q}$ is an arbitrary equivalent martingale measure, we have

$$\begin{aligned} e^{-r(T-t)}E_{\mathbb{Q}}[(S_T-K)_+ \mid \mathcal{F}_t] &= \frac{S_t}{2} + \frac{e^{-r(T-t)}}{\pi}\int_0^{+\infty} \mathcal{R}e\left[\frac{K^{-i\phi}\mathbb{G}^{\mathbb{Q}}_{log(S_T)|\mathcal{F}_t}(i\phi+1)}{i\phi}\right] d\phi \\ &- Ke^{-r(T-t)}\left(\frac{1}{2} + \frac{1}{\pi}\int_0^{+\infty} \mathcal{R}e\left[\frac{K^{-i\phi}\mathbb{G}^{\mathbb{Q}}_{log(S_T)|\mathcal{F}_t}(i\phi)}{i\phi}\right] d\phi\right). \end{aligned}$$

Even though this formula prevents to use slow Monte Carlo methods to approximate the price process, two important numerical issues stay. First, $\mathbb{G}^{\mathbb{Q}}_{log(S_T)|\mathcal{F}_t}$ has to be computed effectively, second, finding the price necessitates univariate numerical integration. For the first point, the IG-GARCH model is particularly well-designed because proposition 2.1.1 (combined with the two preceding risk-neutral dynamics) provides an interesting backward recursive approach. For the second point, the answer is given by [Carr & Madan (1999)] who offer a powerful strategy based on the Fast Fourier Transform (FFT) to compute option prices efficiently for a full range of strikes and a given maturity.[11] In the empirical part, this approach will be used to estimate parameters directly from option prices minimizing an appropriate loss function.

To conclude this section, we provide, for the IG-GARCH model and the two preceding specifications of the stochastic discount factor, a closed-form expression for the one-month risk-neutral expectation of the integrated variance to integrate information on VIX without costly computations.

2.1.4 Related VIX formulas

In a recent study, [Hao & Zhang (2013)] (see also [Qiang *et al.* (2015)]) derived VIX formulas implied by various non-affine Gaussian GARCH dynamics combined with the so-called [Duan (1995)] Local Risk Neutral Valuation Relationship. Furthermore, they proposed a new joint likelihood estimation methodology, including returns and VIX data that was used

[11] For the sake of brevity, we refer the reader to [Chorro *et al.* (2015)], p. 137, where a detailed algorithm is proposed with the associated R source code also used in the present chapter.

in [Kanniainen *et al.* (2014)] to improve pricing performances in the Gaussian GARCH setting.[12] The aim of this subsection is to simply obtain analogous formulas for the affine IG-GARCH model to implement VIX likelihood approaches in the empirical part of the chapter.

The VIX index may be seen as the fair-value strike for a 22-business days variance swap and is known as the fear index. From 2003, the VIX relies on the concept of static replication using all calls and puts with valid quotes, and thus it is independent of any underlying option pricing model. Nevertheless, in discrete time and in the absence of jumps, it can be written as

$$\frac{1}{\tau}\left(\frac{\text{VIX}_t}{100}\right)^2 = \mathbb{E}_{\mathbb{Q}}\left[\frac{1}{T_c}\int_t^{t+T_c} h_s ds \mid \mathcal{F}_t\right] \approx \frac{1}{T_c}\sum_{j=1}^{T_c}\mathbb{E}_{\mathbb{Q}}\left[h_{t+j} \mid \mathcal{F}_t\right] \tag{2.6}$$

where $\tau = 252$, $T_c = 22$, $\mathbb{Q}$ is an equivalent martingale measure[13] and h_t the conditional and historical daily variance. Concerning the IG-GARCH model, from the risk-neutral dynamics described in 2.4 and 2.5 we remark, using iterative properties of the conditional expectation, that the expected conditional variance $\mathbb{E}_{\mathbb{Q}}\left[h_{t+j} \mid \mathcal{F}_t\right]$ can be expressed (see the proof in the Appendix) as a linear combination of the conditional spot variance h_{t+1} and the unconditional variance h_0, weighted by $(\psi^*)^{j-1}$:

$$\mathbb{E}_{\mathbb{Q}}\left[h_{t+j} \mid \mathcal{F}_t\right] = h_{t+1}\left[\psi^*\right]^{j-1} + h_0\left[1 - (\psi^*)^{j-1}\right]$$

where the risk-neutral variance persistence $\psi^* = b + \dfrac{c^*}{(\eta^*)^2} + a^*(\eta^*)^2$ and $\pi h_0 = \dfrac{w^* + a^*(\eta^*)^4}{1-\psi^*}$ only depend on the risk-neutral parameters of the model.[14] Therefore, we easily obtain that:

$$\frac{1}{\tau}\left(\frac{\text{VIX}_t}{100}\right)^2 = h_{t+1}\frac{1-(\psi^*)^{T_c}}{(1-\psi^*)\,T_c} + h_0\left(1 - \frac{1-(\psi^*)^{T_c}}{(1-\psi^*)\,T_c}\right). \tag{2.7}$$

2.2 Estimation of parameters

In the literature, there exist different methods for the estimation of GARCH parameters, the most popular one being the conditional version of the classical Maximum Likelihood Estimation (MLE). In fact, once the GARCH volatility structure and the innovations' density

[12]Such a joint calibration of model parameters is also performed in [Badescu *et al.*(2017)] for the NGARCH model with non-Gaussian innovations when the risk-neutral dynamics is obtained using the so-called extended Girsanov principle of [Elliott & Madan (1998)] and in [Papantonis (2016)] for the Heston-Nandi model associated with the variance dependent pricing kernel of [Christoffersen *et al.* (2013)]. In the latter study a new bivariate normal model for log-returns and the VIX is also introduced to take into account market correlations but this approach is not a priori compatible with the conditional IG distribution of the log-returns in our setting.

[13]In this section, we implicitly suppose that $\mathbb{Q}$ derives from the one period stochastic discount factor processes defined in sections 2.2.1 and 2.2.2.

[14]For the IG-GARCH model, the risk-neutral parameters are simple functions of the historical ones in the case of an exponential-affine stochastic discount factor while they are functions of the historical parameters and π under $\mathbb{Q}^{Ushp}$.

are specified, the conditional log-likelihood based on return observations is in general easy to express and historical parameters are obtained using optimization schemes. For the IG-GARCH model, the knowledge of historical parameters is sufficient to deduce the dynamics under $\mathbb{Q}^{ess}$ because risk-neutral parameters are functions of the historical ones. For the dynamics under $\mathbb{Q}^{Ushp}$, it is not a priori possible to extract the risk-neutral parameter π from return data only. Any additional information, based, for example on options or the VIX index, has to be exploited. To make fair the comparison between the risk-neutral dynamics presented in this chapter and to better exploit the technical flexibility of the IG-GARCH framework, we favor in our study joint estimation strategies using both Returns-Option (see for example [Christoffersen *et al.* (2012)]) and Returns-VIX (see [Kanniainen *et al.* (2014)]) observations.

2.2.1 Joint MLE Estimation using option prices and asset returns

It is well-known that GARCH parameters may be efficiently extracted from option data, when semi-closed form formulas are available for call options prices, minimizing an appropriate loss function. In [Heston & Nandi (2000)] or [Christoffersen *et al.* (2006)] the authors minimize the root mean square error between model and market option prices but as argued in [Christoffersen *et al.* (2012)] this criteria places a greater weight on expensive in-the-money and long-maturity options. To overcome this problem, the linear vega-approximation of implied volatility errors is a popular approach. We obtain estimates of the set of parameters involved in the risk-neutral dynamics, denoted by θ^*, minimizing the Implied Volatility Root Mean Square Error (IVRMSE):[15]

$$\hat{\theta}^* = \arg \underset{\theta^*}{\mathrm{Min}}\; IVRMSE\left(\theta^*\right) = \arg \underset{\theta^*}{\mathrm{Min}} \sqrt{\frac{1}{N_{T_{Op}}} \sum_{t,i} \left(\frac{c_{i,t}\left(h_t^*;\theta^*\right) - \hat{c}_{i,t}}{\hat{V}_{i,t}} \right)^2}. \tag{2.8}$$

Here, n_t is the number of option contracts in the sample at time t and $N_{T_{Op}} = \sum_{t=1}^{T_{Op}} n_t$ where T_{Op} is the number of days in the sample of options. $c_{i,t}\left(h_t^*;\theta^*\right)$ denotes the price of the $i-$th option at time t given by the model[16] while $\hat{c}_{i,t}$ is the price observed in the market. $\hat{V}_{i,t}$ is the Vega associated with $\hat{c}_{i,t}$ that is computed using the implied Black-Scholes volatility $\sigma_{i,t}$ obtained from the market price.

To avoid the distortion of parameters that may appear when performing pure calibration exercises,[17] we present in this subsection a joint MLE estimation using both option prices and asset returns to estimate the parameters of the model as explained in

[15]The Implied Volatility Root Mean Square Error (IVRMSE) will be used in the empirical study to evaluate and compare the pricing performances of the models.

[16]This price is computed using the FFT methodology presented in section 2.3 and depends on the risk-neutral conditional volatility at time t, h_t^*, that is obtained from the log-returns and the risk-neutral GARCH updating rule initialized at its unconditional level.

[17]In fact, when calibrating model parameters, all the attention is focused on the minimization of the in-sample error. Therefore, it is possible to overfit the options dataset and to produce poor out-of-sample pricing errors.

[Christoffersen *et al.* (2012)]. On the one hand, we need to build the log-likelihood function associated with the log-returns $(Y_1, ..., Y_T)$. Under IG innovations, the conditional density function of Y_t given $(Y_1, ..., Y_{t-1})$ is given by:

$$f\left(Y_t \mid Y_1, ..., Y_{t-1}\right) = \frac{h(t)}{\sqrt{2\pi\left(Y_t - r - \nu h_t\right)^3}} e^{-\frac{1}{2}\left(\sqrt{\frac{Y_t - r - \nu h_t}{\eta}} - \frac{h_t}{\eta^2}\sqrt{\frac{\eta}{Y_t - r - \nu h_t}}\right)^2},$$

and the conditional log-likelihood is given by

$$\log L_R = \sum_{t=1}^{T} \log f\left(Y_t \mid Y_1, ..., Y_{t-1}\right) \tag{2.9}$$

that is a function of the historical parameters. On the other hand, in order to obtain the log-likelihood function associated with option data, we consider the Black-Scholes Vega weighted option valuation error:

$$\varepsilon_{i,t} = \left(\frac{c_{i,t}\left(h_t^*; \theta^*\right) - \hat{c}_{i,t}}{\hat{V}_{i,t}}\right)$$

that is an approximation of the implied volatility error. Moreover, assuming that the errors $(\varepsilon_{i,t})$ are independent and identically distributed centered Gaussian random variables the corresponding option log-likelihood can be written (see [Christoffersen *et al.* (2006)]) as:

$$\log L_{Op} = -\frac{1}{2}\sum_{i,t}\left[log\left(\frac{1}{N_{T_{Op}}}\sum_{i,t}\varepsilon_{i,t}^2\right) + \frac{\varepsilon_{i,t}^2}{\frac{1}{N_{T_{Op}}}\sum_{i,t}\varepsilon_{i,t}^2}\right] \tag{2.10}$$

Using both likelihoods in equations 3.9 and 4.16, the joint estimation of the parameters can be obtained by maximizing the joint log-likelihood function:

$$\hat{\theta}^* = \arg\underset{\theta^*}{\mathrm{Max}} \quad \frac{T + N_{T_{Op}}}{2}\frac{\log L_R}{T} + \frac{T + N_{T_{Op}}}{2}\frac{\log L_{Op}}{N_{T_{Op}}} \tag{2.11}$$

where T is the number of days in the returns sample, and $N_{T_{Op}}$ is the total number of option contracts.[18]

2.2.2 Joint MLE Estimation using asset returns and VIX index

This subsection introduces a joint MLE estimation using both returns and the VIX index. In a recent paper, [Hao & Zhang (2013)] proposed a joint likelihood estimation method that incorporates VIX information to capture, in GARCH estimation, the Variance Risk Premium. Their study is based on closed-form formulas for the VIX approximations associated with several Gaussian GARCH pricing models. These formulas, similar to the one obtained in the

[18] We have $\theta^* = \{\nu, \omega, b, c, a, \eta\}$ in the case of the exponential-affine stochastic discount factor and $\theta^* = \{\nu, \omega, b, c, a, \eta, \pi\}$ in the case of the exponential U-shaped one.

present chapter for the affine IG-GARCH model, permit to compute efficiently the related log-likelihood from risk-neutral parameters. Using a similar approach [Kanniainen *et al.* (2014)] have implemented a joint maximum likelihood estimation using returns and VIX with auto-regressive disturbances to enhance the estimation performances of the GARCH option pricing model at a reasonable computational cost. More precisely, in this latter study, the likelihood function on VIX is obtained considering the following model which introduced an error process with autoregressive disturbances:

$$\left\{ \begin{array}{ll} u_t & = \varrho u_{t-1} + e_t \\ u_t & = \text{VIX}_t^{Market} - \text{VIX}_t^{Model}(h_{t+1}^*; \theta^*) \end{array} \right. \tag{2.12}$$

where $(e_t)_t$ are independent and identically distributed centered Gaussian random variables with variance Σ and where $\text{VIX}_t^{Model}(h_{t+1}^*; \theta^*)$ is obtained from equation 3.6. Therefore,

$$\log L_{\text{VIX}} = -\frac{T}{2}\left(\log(2\pi) + \log(\Sigma(1-\varrho^2))\right) + \frac{1}{2}\left(\log(1-\varrho^2)\right) - \frac{1}{2\Sigma}\left(u_1^2 + \sum_{t=2}^{T} \frac{(u_t - \varrho u_{t-1})^2}{1-\varrho^2}\right). \tag{2.13}$$

We combine this log-likelihood with the one associated with the log-returns in equation 3.9 to solve the joint likelihood optimization problem on returns and VIX as follows:

$$\bar{\theta}^* = \arg \underset{(\theta^*, \varrho)}{\text{Max}} \left(\log L_R + \log L_{\text{VIX}}\right) \tag{2.14}$$

where $\bar{\theta}^* = (\theta^*, \varrho^*)$ and ϱ^* is the estimated value of the autoregressive parameter introduced above.

2.3 Empirical results

Based on the preceding theoretical results, this section examines the empirical pricing performances of the IG-GARCH models using the two different stochastic discount factors.

2.3.1 Data properties

To implement the previous joint maximum likelihood estimation strategies using VIX or options information we use in this chapter several time series data. The first one is made of daily log-returns of the S&P500 index and the associated CBOE VIX ranging from January 07, 1999 to December 31, 2009. The series of returns is computed from closing prices. Both the returns and VIX series have 2718 daily observations available for our study. In Table 2.1, we provide the descriptive statistics of the S&P500 log-returns and VIX time series.

The second dataset is made of Wednesday's European call options written on the S&P500 from the CBOE. It contains call option prices for a large range of moneynesses and maturities. The sample period extends from January 01, 2009 to December 31, 2010. Our sample consists of option contracts on 104 Wednesdays and we apply, as most of the empirical studies in the literature (see [Heston & Nandi (2000)], [Christoffersen *et al.* (2006)] or [Kanniainen *et al.* (2014)]), the same filters as [Bakshi *et al.*(1997)]. To empirically study the real option pricing performances of our models, we split up our option dataset into an in-sample and an out-of-sample one. The models will be estimated with the Returns-Option strategy only using the in-sample data. The in-sample option data ranges from January 01, 2009 to December 31, 2009 and the out-of-sample data from January 01, 2010 to December 31, 2010. Table 2.2 (resp. Table 2.3) reports the in-sample (resp. out-of-sample) summary statistics for option data: average price, average implied volatility and the number of contracts for each moneyness/maturity[19] category. The in-sample contains 1332 contracts and the out-of-sample one 1533. Finally, for the risk-free rate that is essential to implement pricing formulas, we use the daily 3 month U.S. Treasury bills (secondary market), obtained from the U.S. Federal Reserve website.

Table 2.4, contains the estimated parameters, as well as their standard errors, for the IG-GARCH model combined with the two different stochastic discount factors using the option-returns and the VIX-Returns methodologies. All the parameters are statistically significant at conventional 5% significance levels. Instead of focusing on the individual parameter values of the models we can analyze the main financial properties. For both estimation methodologies and pricing kernels all volatility models are highly persistent under historical and risk-neutral probabilities, the leverage coefficients are negative and the levels of annualized volatility are in the same range as similar empirical studies. When we analyze the implied variance risk premium, we observe that the IG-GARCH models combined with both pricing kernels captures values in line with classical empirical studies (see [Papantonis (2016)] and references therein). This is a major difference we respect to Gaussian GARCH models.

2.3.2 Testing the validity of the stochastic discount factors

Before testing the pricing performances of the IG-GARCH model more precisely, we propose, following [Guégan *et al.* (2013)], questioning the consistency of the exponential-affine and exponential U-shaped forms of the stochastic discount factor. In this way, we perform a Generalized Method of Moments (GMM) test based on the classical martingale conditions for the risky asset and the associated derivatives. In fact, when (M_t) is a one-period stochastic discount factor we need to have

$$\begin{cases} \mathbb{E}_{\mathbb{P}}\left\{e^{Y_{t+1}}M_{t+1} \mid \mathcal{F}_t\right\} = 1 \\ \mathbb{E}_{\mathbb{P}}\left\{\frac{P_{t+1}(K,T)}{P_t(K,T)}M_{t+1} \mid \mathcal{F}_t\right\} = 1 \end{cases} \tag{2.15}$$

where $P_t(K,T)$ is the price at time t of a call option of strike K and maturity T. Therefore, we test the null hypothesis $\mathbb{E}_{\mathbb{P}}\left\{e^{Y_{t+1}}M_{t+1} \mid \mathcal{F}_t\right\} = 1$[20] using the statistics

$$t_S = \frac{1}{T}\sum_{t=1}^{T}\left(M_{t+1}\frac{S_{t+1}}{S_t} - 1\right). \tag{2.16}$$

[19] We divide the option data into 18 categories according to either moneynesses and times to expiration. The moneyness is defined as the ratio between the forward price of the underlying asset and the option's strike price.

[20] We perform a similar analysis to test the moment condition for the returns on the options for different moneynesses and different time to maturities. The results are presented in Table 2.6, Table 2.7, Table 2.8 and Table 2.9 with similar conclusions.

Under the null hypothesis, $t_S/\hat{\sigma}_T\sqrt{T}$ is asymptotically standard normal where $\hat{\sigma}_T$ is the Newey-West long-run sample variance estimate for $M_{t+1}\frac{S_{t+1}}{S_t}-1$. The results are presented in Table 2.5: for each collection of estimated parameters (see Table 2.4), the statistics proposed in equation 2.16 is computed and compared to the 5% level critical values for standard normal distribution. We find that the null hypothesis is accepted for each stochastic discount factor and estimation methodology. More precisely, the values of the GMM test statistics obtained in Table 2.5 are between −1.96 and 1.96 and the null hypothesis that the moment condition is equal to zero is not rejected at a 5% risk level. This preliminary analysis is not sufficient to discriminate both stochastic discount factors and estimation methodologies that are all compatible with the martingale restriction. In the next subsection, we investigate the related pricing performances in detail.

2.3.3 Pricing performances

Observing the general pricing performances reported at the bottom of Table 2.4, one might reach, without ambiguities, to the conclusion that, independently of the estimation method, the IG-GARCH model combined with a U-shaped pricing kernel provides a much better fit (in-sample and out-of-sample) than the classical exponential-affine approach.

In fact, the in-sample implied volatility roots mean square error IVRMSE for the period 2009 with 1322 contracts is 0.04641 for the exponential-affine SDF model using the joint MLE estimation with option-returns data, while the U-shaped SDF performs slightly better with an IVRMSE of 0.04022, which represents a 13.35% improvement as observed in Table 2.11. Analogous in-sample results are observed when estimating the models using the joint MLE with VIX-Returns data, the IVRMSE is smaller when the U-shaped SDF is used: the IVRMSE for the exponential-affine SDF is now 0.04755 versus 0.03988 for the U-shaped SDF, which represents a 16.134% improvement. We can also observe from Table 2.12 to Table 2.15 the values of the in-sample IVMRSE for different moneynesses and maturities. Therefore, the in-sample analysis strongly favors the U-shaped specification. Concerning the choice of the estimation methodology, even if the results are quite similar, in terms of computational time, we can observe from Table 2.10 that the results associated with the VIX approach are clearly faster to obtain than results from option prices.[21]

The preceding conclusion is not really surprising because an extra parameter is introduced in our approach allowing for more flexibility in calibration exercises. Therefore, it is now interesting to focus on the true test for a pricing model, the out-of-sample pricing performances for the period 2010 when the models are evaluated using the parameter estimates from the 2009 sample period. As observed in Table 2.4, when the model is estimated using option-returns information, the IVRMSE drops from 0.06113 to 0.05133 with the U-shaped pricing kernel which represents a 16.033% improvement. The same holds when VIX-Returns observations are used to estimate the model with a 16.442% improvement. Furthermore, we

[21] This conclusion was conjectured in [Papantonis (2016)]:"'This technique is expected to produce equivalent results to those obtained by using the whole cross-section of options, while at the same time being straightforward and computationally more efficient"'.

can observe from Table 2.12 to Table 2.15 that this result is homogenous regarding moneynesses and time to maturities. It is now clear that the out-of-sample results largely confirm the in-sample ones, the IG-GARCH model provides better pricing performances when the U-shaped SDF is used to obtain risk-neutral dynamics.

2.4 Conclusion

In an important paper, [Christoffersen *et al.* (2006)] proposed an option pricing model based on an IG-GARCH process and the conditional Esscher transform to underline the importance of modelling conditional skewness. One of the main features of this approach is to provide, as in [Heston & Nandi (2000)], semi-closed form formulas for call options but for non-Gaussian innovations. Recently, the monotonicity of the stochastic discount factor (often supposed to be exponential-affine of the log-returns) was discussed in the literature (see for example [Christoffersen *et al.* (2006)] and [Monfort & Pégoraro (2012)]) to favor U shapes. In this chapter, we have explored an extension of [Christoffersen *et al.* (2006)] using a U-shaped pricing kernel that increases the flexibility of the link between the historical and the risk-neutral distributions while preserving the tractability of the model. Our empirical results are clear, the in and out-of-sample pricing performances of the IG-GARCH are improved by the choice of this new pricing kernel. Furthermore, we show in this framework that an estimation strategy based on Returns-VIX information provides very competitive pricing errors at a low computational cost because expensive calibration on options can be bypassed.

Appendix: Proofs

Proposition 1.1.3

Let us first suppose that the pricing equations

$$\begin{cases} \mathbb{E}_{\mathbb{P}}\left\{e^{r} M_{t+1}^{Ushp} \mid \mathcal{F}_t\right\} = 1 \\ \mathbb{E}_{\mathbb{P}}\left\{e^{Y_{t+1}} M_{t+1}^{Ushp} \mid \mathcal{F}_t\right\} = 1 \\ \pi = \dfrac{h_{t+1}^*}{h_{t+1}} \end{cases} \tag{2.17}$$

have a unique solution denoted by $(\theta_{t+1}^*, \varepsilon_{t+1}^*, \rho_{t+1}^*)$. The preceding system can be expressed using the conditional moment generating of the pair (Y_{t+1}, y_{t+1}^{-1}) under $\mathbb{P}$:

$$\begin{cases} \mathbb{G}^{\mathbb{P}}_{(Y_{t+1}, y_{t+1}^{-1})|\mathcal{F}_t}(\theta_{t+1}^*, \rho_{t+1}^*) = e^{-r-\varepsilon_{t+1}^*} \\ \mathbb{G}^{\mathbb{P}}_{(Y_{t+1}, y_{t+1}^{-1})|\mathcal{F}_t}(\theta_{t+1}^* + 1, \rho_{t+1}^*) = e^{-\varepsilon_{t+1}^*} \\ \pi = \dfrac{h_{t+1}^*}{h_{t+1}}. \end{cases} \tag{2.18}$$

To obtain the dynamics under $\mathbb{Q}^{Ushp}$, we compute the risk-neutral conditional moment generating function of Y_{t+1}:

$$\mathbb{G}^{\mathbb{Q}^{Ushp}}_{Y_{t+1}|\mathcal{F}_t}(u) = \mathbb{E}_{\mathbb{Q}^{Ushp}}\left[e^{uY_{t+1}} \mid \mathcal{F}_t\right] = \mathbb{E}_{\mathbb{P}}\left[e^{uY_{t+1}} e^{r} M_{t+1}^{Ushp} \mid \mathcal{F}_t\right] = e^{r+\varepsilon_{t+1}^*} \mathbb{G}^{\mathbb{P}}_{(Y_{t+1}, y_{t+1}^{-1})|\mathcal{F}_t}(\theta_{t+1}^* + u, \rho_{t+1}^*).$$

Using the first equation in (5.2), we can express the risk-neutral moment generating function simply using the historical one:

$$\mathbb{G}^{\mathbb{Q}^{Ushp}}_{Y_{t+1}|\mathcal{F}_t}(u) = \frac{\mathbb{G}^{\mathbb{P}}_{(Y_{t+1}, y_{t+1}^{-1})|\mathcal{F}_t}(\theta_{t+1}^* + u, \rho_{t+1}^*)}{\mathbb{G}^{\mathbb{P}}_{(Y_{t+1}, y_{t+1}^{-1})|\mathcal{F}_t}(\theta_{t+1}^*, \rho_{t+1}^*)}.$$

Given $\mathcal{F}_t$, we know that y_{t+1} follows, under the historical probability $\mathbb{P}$, an IG distribution with degree of freedom $\delta_{t+1} = \frac{h_{t+1}}{\eta^2}$. Therefore, using (2.2), we obtain

$$\mathbb{G}^{\mathbb{Q}^{Ushp}}_{Y_{t+1}|\mathcal{F}_t}(u) = \frac{\mathbb{G}^{\mathbb{P}}_{(Y_{t+1}, y_{t+1}^{-1})|\mathcal{F}_t}(\theta_{t+1}^* + u, \rho_{t+1}^*)}{\mathbb{G}^{\mathbb{P}}_{(Y_{t+1}, y_{t+1}^{-1})|\mathcal{F}_t}(\theta_{t+1}^*, \rho_{t+1}^*)} = e^{u(r+\nu h_{t+1})} \frac{e^{\left[\delta_{t+1} - \sqrt{(\delta_{t+1}^2 - 2\rho_{t+1}^*)(1 - 2(\theta_{t+1}^* + u)\eta)}\right]}}{e^{\left[\delta_{t+1}^* - \sqrt{(\delta_{t+1}^2 - 2\rho_{t+1}^*)(1 - 2\theta_{t+1}^*\eta)}\right]}}$$

and

$$\mathbb{G}^{\mathbb{Q}^{Ushp}}_{Y_{t+1}|\mathcal{F}_t}(u) = e^{[u(r+\nu h_{t+1})] + \delta_{t+1}^*\left[1 - \sqrt{1 - 2(u)\eta^*}\right]}$$

where $\eta^* = \dfrac{\eta}{1-2\theta^*_{t+1}\eta}$ [22] and $\delta^*_{t+1} = \sqrt{(\delta^2_{t+1} - 2\rho^*_{t+1})(1-2\theta^*_{t+1}\eta)}$. Therefore, we can write

$$Y_{t+1} = r + \nu h_{t+1} + \eta^* y^*_{t+1}$$

where, given $\mathcal{F}_t$, y^*_{t+1} follows an IG distribution with degree of freedom δ^*_{t+1}. In particular risk-neutral volatility at time $t+1$ fulfills $h^*_{t+1} = \eta^*\delta^*_{t+1}$ and we deduce from

$$Y_{t+1} = r + \nu h_{t+1} + \eta^* y^*_{t+1} = r + \nu h_{t+1} + \eta y_{t+1}$$

that $y_{t+1} = \frac{\eta^* y^*_{t+1}}{\eta^*}$. Therefore, using that $\pi = \dfrac{h^*_{t+1}}{h_{t+1}}$, (2.1) gives

$$h^*_{t+1} = w^* + bh^*_t + c^* y^*_t + a^* \frac{(h^*_t)^2}{y^*_t}$$

where

$$w^* = w\pi, \quad c^* = \frac{c\pi\eta^*}{\eta}, \quad a^* = \frac{a\eta}{\pi\eta^*}.$$

To conclude the proof it only remains to express η^* using the historical parameters of the model and π. We start from

$$\delta^*_{t+1} = \frac{h^*_{t+1}}{(\eta^*)^2} = \sqrt{(\delta^2_{t+1} - 2\rho^*_{t+1})(1-2\theta^*_{t+1}\eta)}.$$

The martingale condition for the risky asset implies $\mathbb{G}^{\mathbb{Q}^{Ushp}}_{Y_{t+1}|\mathcal{F}_t}(1) = e^r$ from which we can extract ρ^*_{t+1} as a function of θ^*_{t+1}:

$$\rho^*_{t+1} = \frac{\delta^2_{t+1}}{2}\left[1 - \frac{\nu^2\eta^4}{(1-2\theta^*_{t+1}\eta)\left[1-\left(\sqrt{1-2\eta^*}\right)\right]^2}\right].$$

Therefore,

$$\frac{h^*_{t+1}}{(\eta^*)^2} = \frac{-\nu h_{t+1}}{1-\sqrt{1-2\eta^*}}$$

and

$$\pi = \frac{-\nu}{\left[1-\left(\sqrt{1-2\eta^*}\right)\right]}\left[\eta^*\right]^2.$$

Then, the parameter η^* is obtained as the solution of the following cubic equation:

$$(\eta^*)^3 + \frac{2\pi}{\nu}\eta^* + 2\frac{\pi^2}{\nu^2} = 0.$$

It is well known that this equation has a unique real solution if and only if[23]

[22] A priori, the parameter η^* depends on time through θ^*_{t+1} but as we are going to see below, θ^*_{t+1} is time independent.

[23] From the empirical values of the parameters obtained in Table 4, this condition is always fulfilled in our framework.

$$4\left(\frac{2\pi}{\nu}\right)^3 + 27\left(\frac{\sqrt{2}\pi}{\nu}\right)^4 > 0 \Leftrightarrow 27\pi > -8\nu.$$

More precisely, we get

$$\eta^* = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}$$

where $p = \frac{2\pi}{\nu}$ and $q = 2\frac{\pi^2}{\nu^2}$ and we can simplify this expression to obtain

$$\eta^* = \sqrt[3]{\frac{\pi^2}{\nu^2}\left(-1 + \sqrt{1 + \frac{8\nu}{27\pi}}\right)} + \sqrt[3]{\frac{\pi^2}{\nu^2}\left(-1 - \sqrt{1 + \frac{8\nu}{27\pi}}\right)}.$$

Finally, we may deduce from the preceding equality that

$$\theta^*_{t+1} = \frac{1}{2\eta} - \frac{1}{2\left[\sqrt[3]{\frac{\pi^2}{\nu^2}\left(-1 + \sqrt{1 + \frac{8\nu}{27\pi}}\right)} + \sqrt[3]{\frac{\pi^2}{\nu^2}\left(-1 - \sqrt{1 + \frac{8\nu}{27\pi}}\right)}\right]}$$

and that the pricing system (5.1) has a unique solution depending on the historical parameters and π. ■

VIX **as a function of the spot volatility (Section 1.1.4)**

Under both specifications of the pricing kernel, the risk-neutral dynamics of the IG-GARCH model may be written as

$$\begin{cases} X_{t+1} &= r + \nu^* h^*_{t+1} + \eta^* y^*_{t+1} \\ h^*_{t+1} &= w^* + b h^*_t + c^* y^*_t + a^* \frac{(h^*_t)^2}{y^*_t} \end{cases}$$

where, given $\mathcal{F}_t$, y^*_{t+1} follows an IG distribution with parameter $\frac{h^*_{t+1}}{\eta^*}$ under the risk-neutral probability $\mathbb{Q}$. Therefore,[24]

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}}\left[h_{t+j} \mid \mathcal{F}_{t+j-2}\right] &= \mathbb{E}_{\mathbb{Q}}\left[\frac{h^*_{t+j}}{\pi} \mid \mathcal{F}_{t+j-2}\right] \\ &= \frac{1}{\pi}\left[w^* + b h^*_{t+j-1} + \frac{c^*}{(\eta^*)^2} h^*_{t+j-1} + a^* \mathbb{E}_{\mathbb{Q}}\left[\frac{(h^*_{t+j-1})^2}{y^*_{t+j-1}} \mid \mathcal{F}_{t+j-2}\right]\right] \\ &= \frac{1}{\pi}\left[w^* + \left[b + \frac{c^*}{(\eta^*)^2} + a^*(\eta^*)^2\right] h^*_{t+j-1} + a^*(\eta^*)^4\right] \\ &= \frac{1}{\pi}\left[h^*_{t+j-1}\psi^* + h^*_0\left[1 - \psi^*\right]\right] = h_{t+j-1}\psi^* + \tilde{h}_0\left[1 - \psi^*\right] \end{aligned}$$

[24]Using the fact that an IG random variable Z with degree of freedom δ fulfills $E[\frac{1}{Z}] = \frac{1}{\delta} + \frac{1}{\delta^2}$.

where $\psi^* = b + \dfrac{c^*}{(\eta^*)^2} + a^* (\eta^*)^2$ is the variance persistence, $\tilde{h}_0 = \dfrac{h_0^*}{\pi}$ and $h_0^* = \dfrac{w^* + a^* (\eta^*)^4}{1 - \psi^*}$ is the unconditional volatility, under the risk-neutral probability. Now, using the tower property of the conditional expectation operator, the $j-$step ahead prediction of the risk-neutral volatility under the risk-neutral measure is given by

$$\mathbb{E}_{\mathbb{Q}} [h_{t+j} \mid \mathcal{F}_t] = h_{t+1} [\psi^*]^{j-1} + \tilde{h}_0 \left[1 - (\psi^*)^{j-1}\right]$$

and (3.6) follows easily from (4.13). ■

2.5 Tables and figures

Table 2.1: Estimated parameters for the IG model and the two stochastic discount factors.

Joint-Estimation	Returns-Option		Returns-VIX	
Model	M_t^{ess}	M_t^{Ushp}	M_t^{ess}	M_t^{Ushp}
Parameters:				
w	$5.1194E-06$	$7.0905E-07$	$1.0113E-05$	$6.3602E-06$
Stand.Dev	$(1.8754E-05)$	$(3.1078E-06)$	$(8.1758E-08)$	$(2.3894E-08)$
b	$2.6310E-03$	$6.9794E-01$	$5.3380E-03$	$5.4004E-01$
Stand.Dev	$(2.0975E-05)$	$(4.0179E-04)$	$(2.1561E-04)$	$(5.8376E-03)$
c	$4.7759E-05$	$7.3834E-06$	$4.4379E-05$	$1.0943E-05$
Stand.Dev	$(4.8790E-07)$	$(4.318E-04)$	$(5.4214E-09)$	$(7.012E-04)$
a	$3.3174E+03$	$6.8821E+02$	$3.3717E+03$	$6.8048E+02$
Stand.Dev	$(1.0072E-04)$	$(1.853E-01)$	$(3.2231E-01)$	$(2.7751E-03)$
η	$-7.9731E-03$	$-5.1196E-03$	$-7.4167E-03$	$-5.0788E-03$
Stand.Dev	$(9.6039E-04)$	$(9.5374E-06)$	$(5.3127E-04)$	$(4.392E-07)$
ν	$1.2583E+02$	$1.9368E+02$	$1.2590E+02$	$1.9355E+02$
Stand.Dev	$(4.7397E-05)$	$(6.9355E-03)$	$(7.509E-02)$	$(2.3781E-03)$
π	–	1.1688	–	1.2452
Stand.Dev	–	$(5.0449E-02)$	–	$(8.3454E-03)$
ϱ	–	–	$9.9256E-01$	$9.9084E-01$
Stand.Dev	–	–	$(3.7871E-04)$	$(2.0047E-05)$
Historical Model Properties:				
Persistence	0.9688	0.9962	0.9591	0.9818
Annualized volatility	0.2140	0.1931	0.2908	0.2516
Leverage coefficient	−0.0043	−0.0013	−0.0036	−0.0020
Risk-neutral Model Properties:				
Persistence	0.9711	0.9976	0.9975	0.9762
Annualized volatility	0.2223	0.24513	0.3076	0.2780
Leverage coefficient	−0.0044	−0.0013	−0.0046	−0.0025
Average variance risk premium (%):	−3.0364	−3.1732	−3.2617	−3.3828
Pricing performances:				
IVRMSE in-sample (2009)	0.0464	0.0402	0.0475	0.0398
IVRMSE out-of-sample (2010)	0.0611	0.0513	0.0630	0.0527

For the Returns-option strategy, the model is estimated using the log-returns dataset obtained from the closing prices of the S&P500 between January 07, 1999 and December 31, 2009 and the in-sample (2009) option contracts minimizing (2.11). For the Returns-VIX one, the model parameters are obtained minimizing (2.14) using the log-returns and VIX data from January 07, 1999 to December 31, 2009.

Table 2.2: GMM tests for the estimated models to test the moment condition on returns

Estimation\ SDF	M_t^{ess}	M_t^{Ushp}
Returns-option	-0.0276	0.0011
Returns-VIX	-0.0267	0.0082

We compute the statistics t_S for the IG model both combined with the Esscher and the $U-$shaped stochastic discount factors. In each case, the model parameters are estimated using the Returns-option and the Returns-VIX strategies.

Table 2.3: GMM tests, desegregated by moneynesses and times to maturities, to test the moment condition on options for the IG model combined with M_t^{ess} and estimated using the Returns-Option strategy.

	$T < 60$	$60 \leqslant T \leqslant 180$	$T > 180$	**All**
$0 < S/K < 0.975$	-0.0012	-0.0020	-0.0523	-0.0055
$0.975 < S/K < 1.00$	0.0134	0.03141	-0.0924	0.0086
$1.00 < S/K < 1.025$	-0.1548	0.00205	-0.0845	-0.0260
$1.025 < S/K < 1.05$	0.0593	0.02654	-0.0304	0.0159
$1.05 < S/K < 1.075$	-0.0951	0.01434	-0.0575	-0.0108
$1.075 < S/K$	-0.0109	-0.0041	-0.0364	-0.0076
All	-0.0027	-0.0003	-0.0248	-0.0032

Table 2.4: GMM tests, desegregated by moneynesses and times to maturities, to test the moment condition on options for the IG model combined with M_t^{Ushp} and estimated using the Returns-option strategy.

	$T < 60$	$60 \leqslant T \leqslant 180$	$T > 180$	**All**
$0 < S/K < 0.975$	-0.0175	-0.0138	-0.0757	-0.0147
$0.975 < S/K < 1.00$	-0.0397	0.0129	-0.0964	-0.0104
$1.00 < S/K < 1.025$	-0.1300	-0.0465	-0.1590	-0.0556
$1.025 < S/K < 1.05$	0.0028	0.0001	-0.1445	-0.0110
$1.05 < S/K < 1.075$	-0.1089	-0.0247	-0.1916	-0.0456
$1.075 < S/K$	-0.0415	-0.0215	-0.0614	-0.0205
All	-0.0150	-0.0088	-0.0413	-0.0097

Table 2.5: GMM tests, desegregated by moneynesses and times to maturities, to test the moment condition on options for the IG model combined with M_t^{Ushp} and estimated using the Returns-VIX strategy.

	$T < 60$	$60 \leqslant T \leqslant 180$	$T > 180$	**All**
$0 < S/K < 0.975$	−0.0209	−0.0184	−0.0805	−0.0173
$0.975 < S/K < 1.00$	−0.0476	−0.0026	−0.1213	−0.0188
$1.00 < S/K < 1.025$	−0.1695	−0.0877	−0.1825	−0.0734
$1.025 < S/K < 1.05$	−0.0099	−0.0157	−0.1533	−0.0201
$1.05 < S/K < 1.075$	−0.1300	−0.0438	−0.1903	−0.0544
$1.075 < S/K$	−0.0487	−0.0282	−0.0680	−0.0246
All	−0.0180	−0.0123	−0.0448	−0.0117

Table 2.6: Computation times (in hours) to estimate the IG model with the different estimation and risk-neutralization strategies

Estimation\ SDF	Returns-Option	Returns-VIX	
M_t^{ess}	8.0147	0.0151	(54.7 *sec*)
M_t^{Ushp}	9.1583	0.0243	(87.6 *sec*)

Table 2.7: Comparison, based on the IVRMSE, of empirical pricing performances of the IG-GARCH model using M_t^{ess} or M_t^{Ushp}

Model	Returns-Option	Returns-VIX
IVRMSE (2009)	13.351%	16.134%
IVRMSE (2010)	16.033%	16.442%

For example, the value 13.362% represents the improvement (in percentage) of the pricing error for the IG-GARCH model estimated using the Returns-Option strategy when we use the U-shaped pricing kernel instead of the exponential-affine one.

Table 2.8: In-sample IVRMSE, desegregated by moneynesses and time to maturities, using the Returns-Option estimates and M_t^{ess}.

	$T < 60$	$60 \leqslant T \leqslant 180$	$T > 180$	**All**
$0 < S/K < 0.975$	0.4756	0.0552	0.0192	0.0921
$0.975 < S/K < 1.00$	0.0298	0.0127	0.0086	0.0131
$1.00 < S/K < 1.025$	0.0234	0.0127	0.0092	0.0123
$1.025 < S/K < 1.05$	0.0242	0.0111	0.0087	0.0127
$1.05 < S/K < 1.075$	0.0314	0.0120	0.0085	0.0152
$1.075 < S/K$	0.1053	0.0321	0.0132	0.0463
All	0.1103	0.0315	0.0132	0.0464

Table 2.9: In-sample IVRMSE, desegregated by moneynesses and time to maturities, using the Returns-VIX estimates and M_t^{ess}.

	$T < 60$	$60 \leqslant T \leqslant 180$	$T > 180$	**All**
$0 < S/K < 0.975$	0.8456	0.1628	0.0352	0.1411
$0.975 < S/K < 1.00$	0.0482	0.0151	0.0092	0.0174
$1.00 < S/K < 1.025$	0.0287	0.0129	0.0081	0.0126
$1.025 < S/K < 1.05$	0.0294	0.0118	0.0077	0.0171
$1.05 < S/K < 1.075$	0.0328	0.0116	0.0082	0.0148
$1.075 < S/K$	0.0889	0.0316	0.0453	0.0429
All	0.1009	0.0366	0.0412	0.0475

Table 2.10: In-sample IVRMSE, desegregated by moneynesses and time to maturities, using the Returns-Option estimates and M_t^{Ushp}.

	$T < 60$	$60 \leqslant T \leqslant 180$	$T > 180$	**All**
$0 < S/K < 0.975$	0.6107	0.1538	0.0337	0.1270
$0.975 < S/K < 1.00$	0.0272	0.0150	0.0092	0.0153
$1.00 < S/K < 1.025$	0.0268	0.0121	0.0080	0.0119
$1.025 < S/K < 1.05$	0.0228	0.0117	0.0078	0.0130
$1.05 < S/K < 1.075$	0.0323	0.0106	0.0075	0.0121
$1.075 < S/K$	0.0642	0.0301	0.0527	0.0411
All	0.0670	0.0307	0.0406	0.0402

Table 2.11: In-sample IVRMSE, desegregated by moneynesses and time to maturities, using the Returns-VIX estimates and M_t^{Ushp}.

	$T < 60$	$60 \leqslant T \leqslant 180$	$T > 180$	**All**
$0 < S/K < 0.975$	0.3772	0.0273	0.0095	0.0686
$0.975 < S/K < 1.00$	0.0191	0.0066	0.0040	0.0072
$1.00 < S/K < 1.025$	0.0160	0.0070	0.0047	0.0071
$1.025 < S/K < 1.05$	0.0170	0.0062	0.0043	0.0077
$1.05 < S/K < 1.075$	0.0254	0.0072	0.0048	0.0109
$1.075 < S/K$	0.0913	0.0240	0.0086	0.0385
All	0.0941	0.0227	0.0082	0.0398

Table 2.12: Out-of-sample IVRMSE, desegregated by moneynesses and time to maturities, using the Returns-Option estimates and M_t^{ess}.

	$T < 60$	$60 \leqslant T \leqslant 180$	$T > 180$	**All**
$0 < S/K < 0.975$	0.0768	0.0408	0.0149	0.0685
$0.975 < S/K < 1.00$	0.0128	0.0086	0.0076	0.0094
$1.00 < S/K < 1.025$	0.0123	0.0094	0.0079	0.0096
$1.025 < S/K < 1.05$	0.0182	0.0105	0.0086	0.0120
$1.05 < S/K < 1.075$	0.0265	0.0125	0.0105	0.0160
$1.075 < S/K$	0.1215	0.0575	0.0908	0.0773
All	0.0766	0.0427	0.0540	0.0611

Table 2.13: Out-of-sample IVRMSE, desegregated by moneynesses and time to maturities, using the Returns-VIX estimates and M_t^{ess}.

	$T < 60$	$60 \leqslant T \leqslant 180$	$T > 180$	**All**
$0 < S/K < 0.975$	1.0532	0.2465	0.0531	0.1976
$0.975 < S/K < 1.00$	0.0622	0.0223	0.0145	0.0220
$1.00 < S/K < 1.025$	0.0365	0.0176	0.0125	0.0173
$1.025 < S/K < 1.05$	0.0377	0.0167	0.0119	0.0131
$1.05 < S/K < 1.075$	0.0423	0.0153	0.0115	0.0194
$1.075 < S/K$	0.1155	0.0408	0.0733	0.0606
All	0.1190	0.0503	0.0605	0.0630

Table 2.14: Out-of-sample IVRMSE, desegregated by moneynesses and time to maturities, using the Returns-Option estimates and M_t^{Ushp}.

	$T < 60$	$60 \leqslant T \leqslant 180$	$T > 180$	**All**
$0 < S/K < 0.975$	0.0638	0.0180	0.0101	0.0327
$0.975 < S/K < 1.00$	0.0097	0.0082	0.0077	0.0085
$1.00 < S/K < 1.025$	0.0107	0.0092	0.0090	0.0094
$1.025 < S/K < 1.05$	0.0140	0.0098	0.0091	0.0108
$1.05 < S/K < 1.075$	0.0235	0.0118	0.0106	0.0152
$1.075 < S/K$	0.1498	0.0537	0.0239	0.0770
All	0.0959	0.0346	0.0169	0.0513

Table 2.15: Out-of-sample IVRMSE, desegregated by moneynesses and time to maturities, using the Returns-VIX estimates and M_t^{Ushp}.

	$T < 60$	$60 \leqslant T \leqslant 180$	$T > 180$	**All**
$0 < S/K < 0.975$	0.0659	0.0190	0.0106	0.0338
$0.975 < S/K < 1.00$	0.0099	0.0086	0.0082	0.0088
$1.00 < S/K < 1.025$	0.0109	0.0095	0.0094	0.0098
$1.025 < S/K < 1.05$	0.0143	0.0102	0.0096	0.0112
$1.05 < S/K < 1.075$	0.0239	0.0122	0.0111	0.0156
$1.075 < S/K$	0.1513	0.0548	0.0247	0.0780
All	0.0972	0.0354	0.0176	0.0527

Chapter 3

A new two-step estimation strategy for non-Gaussian GARCH models

Contents

In a seminal paper, [Heston (1993)] proposed a stochastic volatility model providing a closed-form solution for the price of a European call option. More precisely, in this model, the characteristic function of the log-returns under the risk-neutral distribution has a closed-form expression and options are priced efficiently from the powerful Fast Fourier Transform (FFT) as explained in [Carr & Madan (1999)]. In particular, the parameters of the model may be calibrated from option datasets at a reasonable computational cost because Monte Carlo simulations can be bypassed. In the GARCH setting, [Heston & Nandi (2000)] offered a discrete-time framework replicating this key feature. Unfortunately, this remarkable property comes at an hefty price: in [Heston & Nandi (2000)], GARCH residuals are supposed to be Gaussian and the recursive volatility structure cannot be chosen arbitrary.[1] Nevertheless, in a recent paper, [Hao & Zhang (2013)] computed model-implied estimates of the VIX

[1]Explicit backward-recursive equations to compute efficiently the moment generating function of the log-returns are obtained because, in this Gaussian setting, both conditional expectation and variance of the volatility process are affine functions of the volatility at the preceding trading date. In [Christoffersen *et al.* (2006)], the authors proposed a non-Gaussian alternative to the Heston-Nandi model preserving the tractability but, once again, the volatility structure is selected to obtain a characteristic function that is an exponential affine function of the state variables.

index for a large family of asymmetric GARCH updating rules under Gaussian innovations. One of the main advantages of this approach is to be compatible with classical GARCH specifications and to provide an alternative approach to estimating the parameters using both log-returns and VIX information when available. In [Kanniainen *et al.* (2014)] and [Papantonis (2016)], the authors proved that these joint maximum likelihood estimation strategies improve considerably the pricing performances of affine and non-affine GARCH models when using the volatility index. However, these studies are restricted once again to a Gaussian setting.

It is now a well-known fact in the literature (see for example [Chorro *et al.* (2015)] section 2.4 and references therein) that non-Gaussian innovations provide a better fit to daily log-returns taking into account all the mass in the tails and the asymmetry. This superiority is confirmed as far as option pricing is concerned because Gaussian GARCH models are unable, in general, to capture the variance risk premium. Unfortunately, when working with non-Gaussian residuals and non-affine volatility prescriptions it is not possible to obtain closed-form expressions for the VIX index or option prices and joint maximum likelihood estimation methodologies cannot be implemented efficiently.[2]

The objective of this chapter is to propose and study a new estimation methodology for non-Gaussian GARCH models, called the two-step estimation strategy, that incorporates, with low computational cost, VIX or options information in the estimation process. More precisely, this strategy estimates separately the volatility and the distribution parameters: in the first step, we assume Gaussian innovations to estimate volatility parameters maximizing joint Option-Returns or joint VIX-Returns likelihood when closed-form formulas are available for the VIX index or plain vanilla options. In a second step, a non-Gaussian distribution is fitted from the residuals obtained in the previous stage. Thereby we both take benefit of the remarkable features of GARCH models in Gaussian environment and work with more realistic distributions. In a deep empirical study, we question the efficiency of this new flexible estimation methodology in terms of option pricing errors. We use two volatility structures (the affine [Heston & Nandi (2000)] and the non-affine[3] [Glosten *et al.* (1993)] GARCH models) combined with the Normal Inverse Gaussian distribution and risk-neutralized using the conditional Esscher transform to analyze the ability of the two-step approach to reproduce empirically observed stylized facts of cross section of options. Our message is clear, the use of non-Gaussian innovations estimated with the two-step approach improves the pricing performances of GARCH models without increasing the computational burden.

The rest of the chapter is organized as follows : in the first section, we will review both the affine-model GARCH-HN and the non-affine-model GARCH-GJR. We will also present

[2] A notable exceptions is the NIG NGARCH model associated with the extended Girsanov principle of [Badescu *et al.*(2018)] where the authors obtained an analytic solution for the VIX. Their approach will be compared to our two-step estimation strategy in the next chapter.

[3] Other non-affine GARCH structures, as the NGARCH model of [Engle & Ng (1993)], have also been tested with similar results.

also the Normal Inverse Gaussian distribution and some of its properties. Section two will describe the estimation of parameters. Moreover, the two steps modified-QML estimation methodology topic will be discussed and developed in that part. Section three will discuss the data, the results of the analysis and the forecasting of the financial time series. Finally, we will finish with a brief conclusion of the work.

3.1 Framework and building blocks

We define the section frame by the following process: Firstly, we will consider two models of time varying volatility dynamics which characterize the conditional volatility: [Heston & Nandi (2000)]'s affine GARCH-HN and the non-affine GARCH-GJR model introduced by [Glosten *et al.* (1993)]. We will then inspect the risk neutral dynamics of each model under Normal Inverse Gaussian innovation with the exponential affine stochastic discount factors M_t^{ess} i.e. the Esscher Transform

3.1.1 GARCH-HN

We next utilized the affine GARCH, introduced by [Heston & Nandi (2000)], with the following[4] dynamics:

$$\begin{cases} Y_t &= r + \lambda_0 h_t + \sqrt{h_t} z_t \\ h_t &= a_0 + a_1 \left(z_{t-1} - \gamma \sqrt{h_{t-1}}\right)^2 + b_1 h_{t-1} \end{cases} \tag{3.1}$$

with $a_0 > 0$, $a_1 \geqslant 0$, $b_1 \geqslant 0$ where r is the risk free rate, z_t are held to be as i.i.d random variables with $\mathbb{E}[z_t] = 0$ and $Var[z_t] = 1$. We represent the leverage effect by the parameter γ, which shows the negative trade-of between volatility and returns. We are able to determine that the variance persistence is $\Psi = b_1 + a_1\gamma^2$ and the average volatility level is $\mathbb{E}[h_t] = \dfrac{a_0 + a_1}{1 - \Psi}$. Considering Gaussian innovations, due to this models's affine structure, it has all the benefits of a closed-form solution for the cost of a European call option. During the estimation process, this enables us to use option data directly.

The locally risk-neutral dynamics[5] for the GARCH-HN model in the equation (3.1), under Gaussian conditional distribution, are discovered by the moment generating function of Y_t given $\mathcal{F}_{t-1}$. We can write the dynamics thus:

$$\begin{cases} Y_t &= r - \dfrac{1}{2} h_t + \sqrt{h_t} z_t^* \\ h_t &= a_0 + a_1 \left(z_{t-1}^* - \left(\gamma + \lambda_0 + \dfrac{1}{2}\right)\sqrt{h_{t-1}}\right)^2 + b_1 h_{t-1}. \end{cases} \tag{3.2}$$

where z_{t+1} are i.i.d $N(0,1)$, the persistence is fixed as $\Psi^* = b_1 + a_1\left(\gamma + \lambda_0 + \frac{1}{2}\right)^2$ and the variance's first value h_0 is fixed as equal to the average volatility level.

[4] A unique second order stationary solution exists if and only if $a_1\gamma^2 + b_1 < 1$.

[5] The parameters of M_t^{ess} can be obtained from the pricing relations and the moment generation function (for more details see [Chorro *et al.* (2015)] in page 88).

3.1.2 GARCH-GJR

We next utilised the non-affine GARCH-GJR, developed by [Glosten *et al.* (1993)], created for the analysis of the asymmetric effects of positive and negative asset returns. We define the model in the following equation[6]:

$$\begin{cases} Y_t &= r + \lambda_0 \sqrt{h_t} - \frac{h_t}{2} + \sqrt{h_t} z_t \\ h_t &= a_0 + h_{t-1}\left[b_1 + a_1 (z_{t-1})^2 + \gamma \max(0, -z_{t-1})^2\right] \end{cases} \tag{3.3}$$

with $a_0 > 0$, a_1, b_1, $\gamma \geqslant 0$ where z_t are i.i.d random variables with $\mathbb{E}[z_t] = 0$, $Var[z_t] = 1$ and the asymmetry[7] is captured by γ. When $\gamma > 0$, the model will account for leverage effect, meaning future volatility is risen more highly by bad news than good news ($z_t \geqslant 0$). We can express the persistence as $\Psi = b_1 + a_1 + \frac{\gamma}{2}$ and the unconditional variance $h_0 = \frac{a_0}{1-\Psi}$.

Under the risk-neutral measure $\mathbb{Q}^{ess}$ with Gaussian innovations, [Duan (1995)] suggests that the risk-neutral volatility dynamics may be defined thus:

$$\begin{cases} Y_t &= r - \frac{h_t}{2} + \sqrt{h_t} z_t^* \\ h_t &= a_0 + h_{t-1}\left[b_1 + a_1 \left(z_{t-1}^* - \lambda_0\right)^2 + \gamma \max\left(0, -\left(z_{t-1}^* - \lambda_0\right)\right)^2\right]. \end{cases} \tag{3.4}$$

where z_t^* are i.i.d $N(0,1)$, the persistence $\Psi^* = b_1 + [a_1 + \gamma N(\lambda_0)]\left(1 + \lambda_0^2\right) + \gamma \lambda_0 n(\lambda_0)$ where $N(.)$, $n(.)$ describes the standard normal cumulative in addition to the functions of density function. We must indicate that, unlike the affine model, the non-affine GARCH-GJR model gives us no closed-form formula for option prices. The capacity of it to exploit the option data is thus decreased, and these must instead be computed by Monte Carlo methods.

Nevertheless, in a recent work, [Kanniainen *et al.* (2014)] determined the implied VIX formulas for the GARCH-HN and GARCH-GJR models under Gaussian innovations. They employed a joint likelihood estimation methodology, including returns and VIX data, to give better pricing performances. Following [Hao & Zhang (2013)], the VIX index can be viewed as the fair-value strike for a variance swap of 21-business days:

$$\frac{1}{\tau}\left(\frac{\text{VIX}_t}{100}\right)^2 = \frac{1}{T_c}\sum_{j=1}^{T_c} \mathbb{E}_{\mathbb{Q}^{ess}}\left[h_{t+j} \mid \mathcal{F}_t\right] \tag{3.5}$$

where $\tau = 250$, $T_c = 21$. Through the use of the tower property of the conditional expectation under the risk neutral dynamics in equation 3.2, 3.3, the forecast conditional variance $\mathbb{E}_{\mathbb{Q}^{ess}}[h_{t+j} \mid \mathcal{F}_t]$ can be represented thus:

$$\mathbb{E}_{\mathbb{Q}^{ess}}\left[h_{t+j} \mid \mathcal{F}_t\right] = h_{t+1}\left[\Psi^*\right]^{j-1} + \tilde{h}_0\left[1 - (\Psi^*)^{j-1}\right] \qquad \text{with} \qquad \tilde{h}_0 = \frac{a_0 + (a_1 \cdot \mathbb{1}_{HN})}{1 - \Psi^*}$$

where the indicator[8] function of HN is denoted as $\mathbb{1}_{HN}$. Furthermore, we find $\forall t \in$

[6]The variance is similarly weak stationary under the physical probability if $\Psi = b_1 + a_1 + \frac{\gamma}{2} < 1$.

[7]Asymmetry and leverage: ([Glosten *et al.* (1993)], [Chorro *et al.* (2015)]) Asymmetry explains the fact that positive and negative shocks of equal magnitude do not have the same effects on volatility; the term leverage refers to the possibility that negative shocks increase volatility.

[8]i.e $\mathbb{1}_{HN} = 1$ when HN is used and zero otherwise.

$\{1,\cdots,T\}$, the expression of the n of HN. Moreover, we obtain $\forall t \in \{1,\cdots,T\}$ the expression of the VIX_t thus:

$$\text{VIX}_t = 100\left(\sqrt{\tau\left[h_{t+1}\frac{1-(\Psi^*)^{T_c}}{(1-\Psi^*)\,T_c} + \tilde{h}_0\left(1-\frac{1-(\Psi^*)^{T_c}}{(1-\Psi^*)\,T_c}\right)\right]}\right). \tag{3.6}$$

where $\tilde{h}_0$ is the unconditional variance under the risk-neutral measure, and Ψ^* the persistence under the risk-neutral probability.

3.1.3 NIG-GARCH models

We will rapidly cover the Normal Inverse Gaussian distribution[9] (NIG) proposed by [Barndorff-Nielsen (1998)]. We will consider $Y \sim NIG\,(\alpha,\beta,\delta,\mu)$, a random variable following a Normal Inverse Gaussian distribution along with parameters $\theta^D = (\alpha,\beta,\delta,\mu)$. Its probability density function is formed thus:

$$d_{NIG}\,(z,\alpha,\beta,\delta,\mu) = \frac{\alpha}{\pi}\, e^{\delta\left(\sqrt{\alpha^2-\beta^2}+\beta\left(\frac{z-\mu}{\delta}\right)\right)}\;\frac{K_1\left(\alpha\delta\sqrt{1+\left(\frac{z-\mu}{\delta}\right)^2}\right)}{\sqrt{1+\left(\frac{z-\mu}{\delta}\right)^2}} \tag{3.7}$$

where $\delta > 0$, $\alpha >\mid \beta \mid> 0$ where μ is the location, β the skewness, α the tail-heaviness and δ the scale. K_1 is the modified Bessel function of the third kind with index one, with mean $m = \mu + \frac{\delta\beta}{\sqrt{\alpha^2-\beta^2}}$ and variance $\sigma^2 = \frac{\delta\alpha^2}{\sqrt{\alpha^2-\beta^2}^3}$. Hence, every affinely transformed, and in particular every linearly combination, NIG random variable, is an NIG random variable. Next we will take a look at the transformed NIG random variable:

$$X = \frac{1}{\sigma}(Y-m) \quad \text{with} \quad X \sim \text{NIG}\left(\tilde{\alpha},\tilde{\beta},\tilde{\delta},\tilde{\mu}\right) \tag{3.8}$$

which is a centered version with unit variance where $\tilde{\alpha} = \sigma\alpha$, $\tilde{\beta} = \sigma\beta$, $\tilde{\delta} = \frac{\delta}{\sigma}$ and $\tilde{\mu} = \frac{-m}{\sigma} + \frac{\mu}{\sigma}$.

The stochastic discount factors M_t^{ess}, employed by us in this chapter are characterized by an exponential affine form the Esscher Transform developed by [Gerber & Shiu (1994)]. M_t^{ess} is denoted as $\forall t \in \{0,\cdots,T\}$ as $M_t^{ess} = e^{\theta_t Y_t + \varepsilon_t}$ where Y_t is the logarithm of the stock price process' returns at time t, θ_t and ε_t and $\mathbb{Q}^{ess}$ are predictable coefficients. Under $-$, the dynamic of the NIG-GARCH-HN and NIG-GARCH-GJR are altered in a non-linear fashion as detailed in [Chorro *et al.* (2015)] (chapter 3 proposition 3.4.7.):

$$z_t^* \hookrightarrow NIG\left(\tilde{\alpha},\tilde{\beta}+\sqrt{h_t}\theta_t^q,\tilde{\delta},\tilde{\mu}\right)$$

[9] This distribution is generally well suited to financial time series returns (see [Barndorff-Nielsen (1998)], [Badescu *et al.*(2015)] and [Badescu *et al.*(2017)]). We could regard the NIG distribution as a subclass of the generalized hyperbolic distributions.

where

$$\theta_t = -\frac{1}{2} - \frac{\tilde{\alpha}\tilde{\beta}\sqrt{\tilde{\delta}}}{\sqrt{h_t}\varrho^{\frac{3}{2}}} - \frac{1}{2}\sqrt{\frac{\left(\tilde{\alpha}m_t + \sqrt{\tilde{\delta}h_t}\tilde{\beta}\varrho\right)^2}{h_t\tilde{\delta}\varrho^3}\left(\frac{4\tilde{\alpha}^4\tilde{\delta}^2}{h_t\tilde{\delta}\varrho^3 + \left(\tilde{\alpha}m_t + \sqrt{\tilde{\delta}h_t}\tilde{\beta}\varrho\right)^2} - 1\right)}$$

$\varrho = \sqrt{\tilde{\alpha}^2 - \tilde{\beta}^2}$ and $m_t = \lambda_0 h_t$ for HN and $m_t = \lambda_0\sqrt{h_t} - \frac{h_t}{2}$ for GJR. However, we should point out that both models can be estimated and investigated empirically. Indeed, in this chapter we aim to suggest an alternative to facilitate the estimation of NIG-GARCH model.

3.2 Estimation of parameters

For the two-step modified Quasi-Maximum Likelihood estimation of NIG-GARCH processes, instead of immediately employing the NIG distribution, we will begin by giving our attention to the Gaussian-GARCH estimable using several different methods. We employed the joint estimation procedure of [Kanniainen *et al.* (2014)] which gives us a joint MLE estimation analysis of affine and non-affine Gaussian-GARCH models, combining information from the underlying asset returns and options data or VIX index.

We are able to approximate the Gaussian-GARCH-HN volatility model's parameters θ^V through direct calibration with option prices. Nevertheless, it must be pointed out that this approach depends having closed-form expressions of the option price, which is true for the Gaussian GARCH-HN. However, once started, this approach is extremely intensive on computing time. The amount of pricing simulation paths has an effect on computing time. To reduce the time spent, and to provide information from VIX indexes, we employ [Christoffersen *et al.* (2012)]'s joint MLE estimation option-returns in addition to the joint MLE estimation VIX-returns explained in [Kanniainen *et al.* (2014)].

3.2.1 Joint MLE estimation

Let us define the cardinal of the set of option market prices be as N. [Christoffersen *et al.* (2012)] maintains that it is possible to obtain the model parameters $\theta = \left(\lambda_0, \theta^V\right)$ by employing the joint MLE estimation techniques with option and returns information:

$$\hat{\theta} = \arg\underset{\theta}{\mathrm{Max}} \ \frac{T+N}{2}\frac{\log L_R}{T} + \frac{T+N}{2}\frac{\log L_{Op}}{N} \tag{3.9}$$

where $\log L_R$ denotes the conditional log-likelihood function associated with the returns and $\log L_{Op}$ is the conditional log-likelihood function associated with the option data. $c_i\left(h_t^R;\theta^*\right)$ denotes the model prices and $\hat{c}_i$ the option market prices. Then we may describe $\log L_{Op}$ as follows:

$$\log L_{Op} = -\frac{1}{2}\sum_{t=1}^{N}\left[\log\left(\frac{1}{N}\sum_{i=1}^{N}\varepsilon_i^2\right) + \frac{\varepsilon_i^2}{\frac{1}{N}\sum_{i=1}^{N}\varepsilon_i^2}\right] \text{ with } \varepsilon_i = \frac{c_i\left(h_t^R;\theta^*\right) - \hat{c}_i}{\hat{V}_i},$$

where $\hat{V}_i$ is the Black and Scholes Vega associated with $\hat{c}_i$, and the implied volatility errors (ε_i) are supposed to be i.i.d centered Gaussian variables with variance $\frac{1}{N}\sum_{i=1}^{N}\varepsilon_i^2$. In both Gaussian-GARCH models, $\log L_R$ is obtained thus:

$$\log L_R = -\frac{1}{2}\sum_{t=1}^{T}\left(\log(2\pi) + \log(h_t) + \frac{(Y_t - r - m_t)^2}{h_t}\right).$$

where $m_t = \lambda_0 h_t$ for HN and $m_t = \lambda_0\sqrt{h_t} - \frac{h_t}{2}$ for GJR.

According to [Kanniainen *et al.* (2014)]), it is possible to obtain the estimated parameters $\theta = \left(\lambda_0, \theta^V\right)$ associated with the joint-MLE estimation of VIX-Returns by solving the optimization:

$$\left(\hat{\theta}, \hat{\varrho}\right) = \arg\underset{(\theta,\varrho)}{\mathrm{Max}}\left(\log L_R + \log L_{\mathrm{VIX}}\right) \tag{3.10}$$

where the log-likelihood $\log L_{\mathrm{VIX}}$ on $(\mathrm{VIX}_t)_t$ can be defined as:

$$\log L_{\mathrm{VIX}} = -\frac{T}{2}\left(\log(2\pi) + \log(\Sigma(1-\varrho^2))\right) + \frac{1}{2}\left(\log(1-\varrho^2)\right) - \frac{1}{2\Sigma}\left(u_1^2 + \sum_{t=2}^{T}\frac{(u_t - \varrho u_{t-1})^2}{1-\varrho^2}\right) \tag{3.11}$$

under the assumption that the errors process $(u_t) = \mathrm{VIX}_t^{Market} - \mathrm{VIX}_t^{Model}(h_{t-1};\theta)$ is follow autoregressive disturbances $u_t = \varrho u_{t-1} + e_t$ and $(e_t)_t$ are i.i.d Gaussian random variables with mean zero and variance Σ^2.

The respective benefits of both procedures are that former studies the information from option data and from the returns proportionally, whereas the latter simultaneously amalgamates the VIX and the returns data. However, due to its use of options, the former needs closed-form expressions of the prices. Like in the case of NLS, it may only be employed in the case of Gaussian-GARCH-HN. Moreover, both of these alternatives are dependent on the Gaussian innovation hypothesis. Nevertheless, we require each of these joint estimation strategies in our novel method to estimate non-Gaussian GARCH models.

3.2.2 Two-step estimation strategy

The two-step estimation strategy, like the QML procedure, separately estimates the vector of volatility parameters $\theta = \left(\lambda_0, \theta^V\right)$ and the distribution parameters θ^D in two separate phases. This is performed successively. When fixing the volatility dynamics in equation 3.1 and 3.3, in the primary phase, the QML estimation takes it for granted that the innovations z_t follow a Gaussian distribution to approximate the volatility parameters $\left(\lambda_0, \theta^V\right)$. The secondary phase involves the estimation of the distribution parameter θ^D by maximizing:

$$\sum_{t=1}^{T} -\frac{\log(h_t)}{2} + \log\left[d_{\theta^D}\left(z_t\left(\hat{\lambda}_0, \hat{\theta}^V\right)\right)\right]$$

employing the standardized[10] residuals where d_{θ^D} is the density function of the NIG distribution.

In a similar manner, the distribution and volatility parameters are looked at separately, particularly though the reduction of the dimensions of optimization problems. Rather than looking at a normal distribution in the second step, our focus is on the density function of centered NIG distribution. The iterative algorithm requires the following steps:

- Step 1: We suppose that the $(z_t)_t$ are i.i.d $N(0;1)$ under $\mathbb{P}$. In this manner, we can estimate the vector of volatility parameters $\left(\lambda_0, \theta^V\right)$ as follows :
  - When there is a closed-form formula for option prices, we can obtain $\left(\lambda_0, \theta^V\right)$ for the maximization problem in equation (3.9).
  - When there is no closed-form formula for option prices but we do have a closed-form formula for the VIX index then, we can obtain $\left(\lambda_0, \theta^V\right)$ through the maximisation the joint VIX-Returns likelihood (3.10).
- Step 2: From the i.i.d residuals $\left(z_1\left(\hat{\lambda}_0, \hat{\theta}^V\right), \cdots, z_T\left(\hat{\lambda}_0, \hat{\theta}^V\right)\right)$ that may be extracted by the step before, We obtain the distribution vector of parameters θ^D by maximizing:

$$\sum_{t=1}^{T} -\frac{\log(h_t)}{2} + \log\left[d_{\theta^D}\left(z_t\left(\hat{\lambda}_0, \hat{\theta}^V\right)\right)\right]$$

  where d_{θ^D} is the density function of a centered NIG random variable with unit variance in equation 3.7.

In comparison to the aforementioned estimation strategies, the two-step approach enables us to take advantage both of the Gaussian hypothesis and the impact of the NIG distribution. To obtain an explicit expression of the risk neutral dynamics, we have made the assumption that the innovations are i.i.d Gaussian random variables as described [Kanniainen *et al.* (2014)]. This makes the optimization procedure smaller and enables us to estimate volatility parameters. In return, though the NIG distribution gives us no price estimations, it enables the use of specific methods for price calculation. Chiefly, as described in [Chorro *et al.* (2015)], the use of the Monte Carlos simulation with the Empirical Martingale Simulation Method (EMS) can effectively approximate the option prices. Moreover, as our results clearly illustrated, the prices that we obtained from this process performed better than those with closed-form solutions from the Gaussian hypothesis.

3.3 Empirical results

3.3.1 Data properties

In addition a dataset of options was used and written on the S&P500. We limited ourselves to contracts from Wednesday due to the quantity of option pricing models to analyze in

[10] $\left(z_1\left(\hat{\lambda}_0, \hat{\theta}^V\right), \cdots, z_T\left(\hat{\lambda}_0, \hat{\theta}^V\right)\right) = \left(\frac{Y_1 - r - \hat{\lambda}_0 h_1}{\sqrt{h_1}\left(\hat{\theta}^V\right)}, \cdots, \frac{Y_T - r - \hat{\lambda}_0 h_T}{\sqrt{h_T}\left(\hat{\theta}^V\right)}\right)$ for the GARCH-HN.

this section. It concerned 4563 options contracts quoting the prices of the period beginning in January 2nd, 2009 and ending in April 15, 2012. We split the option data set into two subcategories: the first where the model parameters are valued (to use for the affine models the joint likelihood estimation centred on returns and options) and the second category which was used to equate the pricing performances of the models. The initial category for the in-sample approximation contrast is named Dataset A from January 2nd, 2009 to December 22, 2010 and comprises 2714 and contracts. On the other hand, the subsequent category for the out-of-sample contrast is named Dataset B and comprises 1849 contracts with 67-Wednesdays from January 03, 2011 to April 15, 2012.

3.3.2 Pricing performance using Montecarlo simulation

Using Monte Carlo techniques we can implement the estimated parameters $\left(\hat{\lambda}_0, \hat{\theta}^V, \hat{\theta}^D\right)$ to decide the price options. The Empirical Martingale was used to increase the numerical efficiency of the Monte Carlo simulation as a tool to decrease the variance.[11] Simulation Method (EMS). To demonstrate this [Chorro *et al.* (2015)] conducted a detailed study of the empirical pricing performance of the (EMS). To calculate and contrast the Implied Volatility Root Mean Square (IVRMSE) and the Volatility Risk Premium (VRP) we used the estimated parameters and the simulated option prices. It is important to note that [Papantonis (2016)] defines the Volatility Risk Premium as the difference between the conditional volatility estimated with the physical and risk-neutral expectations:

$$VRP_t = \mathbb{E}_t^{\mathbb{P}}\left[Var\left(Y_{t+1}\right)\right]^{\frac{1}{2}} - \mathbb{E}_t^{\mathbb{Q}}\left[Var\left(Y_{t+1}\right)\right]^{\frac{1}{2}} = \sqrt{h_t} - \sqrt{h_t^*}.$$

where $\mathbb{E}_t^{\mathbb{P}}(.)$ and $\mathbb{E}_t^{\mathbb{Q}}$ are correspondingly the conditional expectation with the physical measure − and risk-neutral measure −. When we consider Gaussian-innovation, it is evident that the risk-neutral conditional variance h_t^* is equivalent to h_t (see [Chorro *et al.* (2015)]). If we accept this, the VRP in GARCH-HN and GARCH-GJR cases could be zero, which is not the situation of the VRP in NIG-GARCH model.

For every evaluation procedure, Table 3.1 and Table 3.2 to total the parameters estimated from the two-steps Modified-QLME estimation for NIG-GARCH model. The table 3.3, 3.4, 3.5, 3.6, 3.7 and 3.8 present to us, in a clear format, the pricing performances founded on IVRMSE, VRP and time to calculate each model. From the results of these tables we can suggest various empirical deductions. When it comes to the NIG-GARCH model, the results from these were very promising summaries for the time series method to option pricing. Normally, we get an added reduction in the IVRMSE when contrasted with the GARCH-HN model in terms of joining together NIG distribution and GARCH-GJR model.

The values of IVRMSE from the table 3.3, 3.5, the pricing performances of the NIG-GARCH-GJR is superior, compared to the NIG-GARCH-HN model (which itself stands as a good model when contrasted to Gaussian). The option pricing performance is better when

[11] The empirical martingale simulation is an fascinating technique used to decrease the variance that [Duan & Simonato (1998)] has presented. It was used extensively to heighten the numerical efficiency of the Monte Carlo estimators in GARCH option pricing models (for example [Chorro *et al.* (2015)])

using the two step approximation method, as shown in table 3.7. Furthermore, when assessing option returns, the NIG-GARCH-GJR models convey a performance which is similar to that of the class NIG-GARCH-HN. However, it is less efficient than the NIG-GARCH model based on the VIX returns. Nonetheless, we it is important to that the second method is undoubtedly due to the use of the VIX in its pricing approach, consequently accounting for the true volatility premium. When it comes to option pricing, both in-sample and out-of-sample studies favor the NIG-GARCH-GJR models with Returns-VIX co-dependant approximation (see table 3.4, 3.6 and 3.7 as well as 3.8). When considering the final result, we can confirm the calculation time (see table 3.3 and 3.5) which offers a clear indication of NIG-GARCH-GJR as an suitable method for option valuation modelling.

3.4 Conclusion

We introduce in this chapter a new estimation strategy for non-Gaussian GARCH option pricing models. This two-step inference methodology incorporates, with low computational cost, VIX or options information in the estimation process. It finds its origin in a very simple observation : under Gaussian hypotheses, some GARCH-type models have outstanding properties (closed-form expressions for the VIX and/or option prices) that fail when more realistic innovations are involved. More precisely, this strategy estimates separately the volatility and the distribution parameters assuming in the first step Gaussian innovations to estimate volatility parameters while in a second step, a non-Gaussian distribution is fitted from the residuals previously obtained . We provide a deep empirical study to illustrate the importance of combining a non-Gaussian distribution and joint likelihood estimation methodologies using VIX or options information and prove that the two-step approach improves the pricing performances of GARCH models without increasing the computational burden.

3.5 Tables and figures

Figure 3.1: Plot of h, h^* and VRP for GARCH-HN estimated with Return-VIX

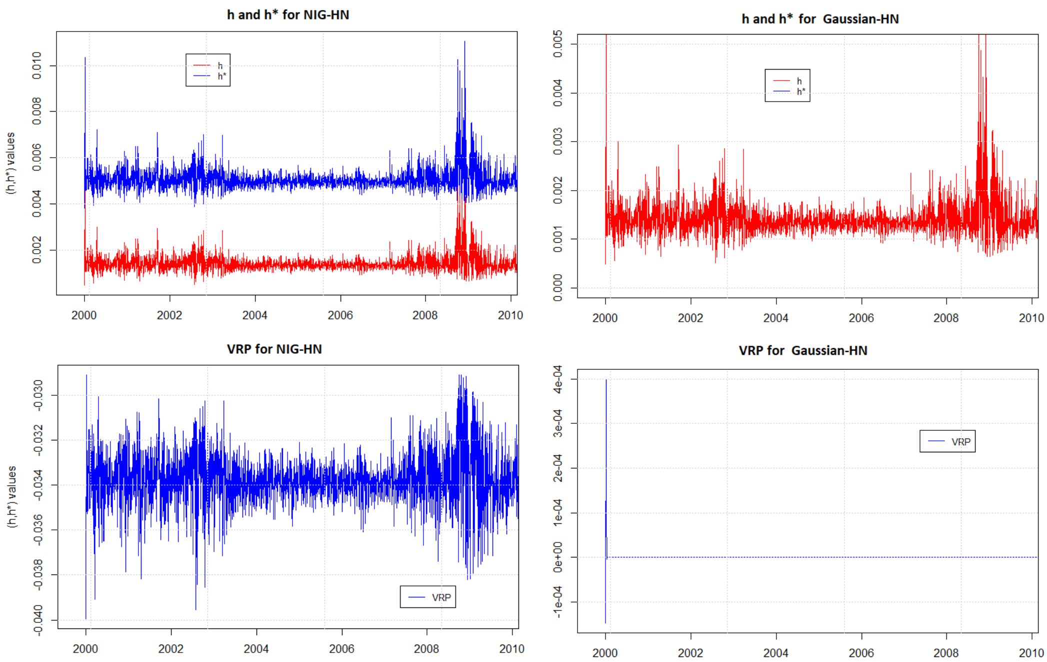


Figure 3.2: Plot h, h^* and VRP for GARCH-GJR estimated with Return-VIX

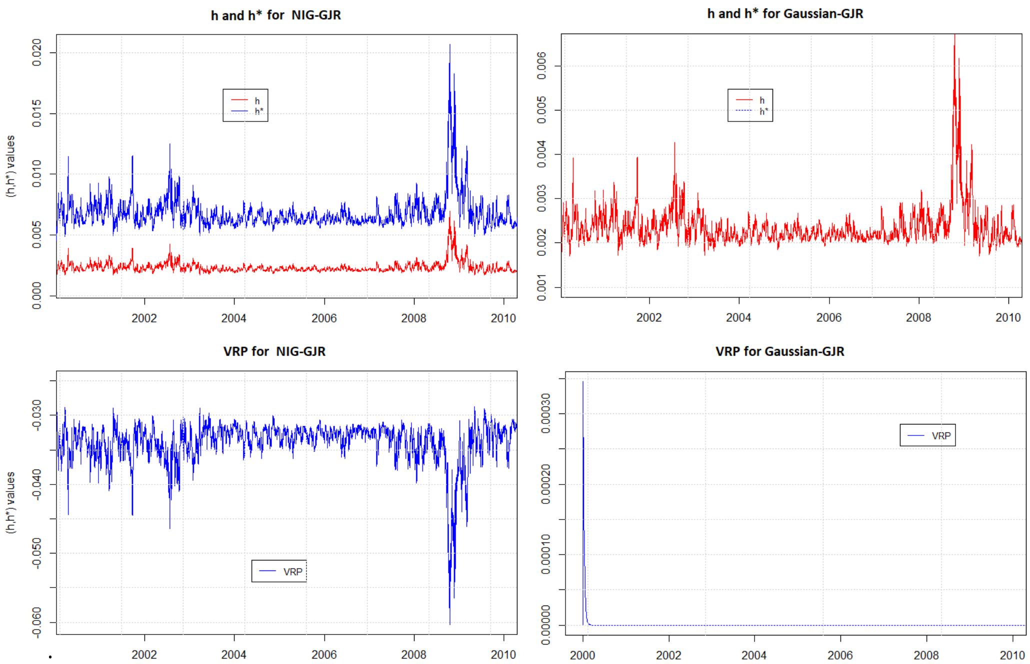


Table 3.1: Summary fits of the first step GARCH-Gaussian-Ess: Estimates and standard errors using returns dataset and the in-sample dataset (2009-2010) with Gaussian innovation and M^{ess}. Returns means MLE-estimation procedure using only returns information, Opt-VIX means Joint MLE estimation using returns and option information, Ret-VIX means Joint MLE estimation using returns and Vix information.

GARCH-type	HN-GARCH	GJR-GARCH	HN-GARCH	HN-GARCH	GJR-GARCH
Information	Returns	Returns	Opt-Ret	Ret-VIX	Ret-VIX
a_0	$3.854E-08$	$3.049E-06$	$1.859E-07$	$3.757E-12$	$4.966E-06$
Stand.Dev	(0.0044)	(0.00111)	(0.0009)	(0.0007)	(0.0004)
a_1	$2.254E-05$	$1.243E-01$	$1.542E-06$	$2.252E-05$	$1.240E-01$
Stand.Dev	(0.0001)	(0.0003)	(0.0003)	(0.0002)	(0.0000)
b_1	$8.272E-01$	$8.509E-01$	$6.500E-01$	$9.117E-01$	$8.504E-01$
Stand.Dev	(0.0035)	(0.00068)	(0.0030)	(0.0086)	(0.0015)
γ	$5.379E+01$	$2.208E-02$	$4.586E+02$	$1.423E+01$	$2.314E-02$
Stand.Dev	(0.0011)	(0.0025)	(0.0095)	(0.0088)	(0.0005)
λ_0	$1.020E+00$	$2.288E-01$	$8.596E+00$	$1.513E+00$	$1.989e^{-01}$
Stand.Dev	(0.0000)	(0.0055)	(0.0008)	(0.0501)	(0.0033)
ϱ	–	–	–	0.9992	0.8924
Stand.Dev	–	–	–	(0.0106)	(0.0012)

Table 3.2: Summary fits of the MLE-estimation procedure with NIG innovation of historical parameters using M^{ess}. This table shows the set of estimated NIG-distribution parameters obtain from the second step of the two steps estimation procedure using the results of the previous table 3.1 with in-sample data (2009-2010).

GARCH-type Information	HN Returns	GJR Returns	HN Opt-Ret	HN Ret-VIX	GJR Ret-VIX
$\hat{\alpha}$	1.2501	1.1550	1.4630	1.4365	1.3589
Stand.Dev	(0.0004)	(0.01089)	(0.0005)	(0.0008)	(0.0001)
$\hat{\beta}$	−0.0106	−0.1432	−0.0061	−0.0538	−0.0058
Stand.Dev	(0.0008)	(0.0057)	(0.0008)	(0.0003)	(0.0023)
$\hat{\delta}$	1.4728	1.0623	1.4454	1.3920	1.5336
Stand.Dev	(0.0095)	(0.0000)	(0.0005)	(0.0008)	(0.0000)
$\hat{\mu}$	2.7086	0.1327	2.1602	11.6243	7.9908
Stand.Dev	(0.0051)	(0.0076)	(0.0000)	(0.0013)	(0.0000)

Table 3.3: Results of performance analysis on IVRMSE, VRP, and computation time of each models under Gaussian-disrtibution, using the Option-Returns estimates of VIX-Returns estimates.

GARCH	HN-Ret	GJR-Ret	HN-Opt-Ret	HN-VIX-Ret	GJR-VIX-Ret
Times (h)	0.010	0.018	9.014	0.008	0.021
$-VRP$ (in %)	$3.27E-10$	$2.86E-16$	$8.88E-11$	$9.67E-09$	$7.122E-13$
in-$IVRMSE$	0.05991	0.05747	0.05574	0.05801	0.05483
out-$IVRMSE$	0.07770	0.07648	0.07339	0.07351	0.06500

Table 3.4: Comparison, based on the IVRMSE, of empirical pricing performances of the Gaussian-GARCH (GARCH-1 /GARCH-2, as example: $-4.072\% = 100 \cdot (0.05574 - 0.05801)/0.05574$), using Oprtion-Returns or VIX-Returns information.

GARCH-1	HN-Opt-Ret	HN-Opt-Ret	HN-VIX-Ret
GARCH-2	HN-VIX-Ret	GJR-VIX-Ret	GJR-VIX-Ret
in-sample	−4.072	1.6325	5.4818
out-sample	−0.163	11.4320	11.5766

Table 3.5: Results of performance analysis on IVRMSE, VRP, and computation time of each models under NIG-disrtibution, using the returns-option estimates of returns-VIX estimates.

NIG-GARCH	HN-Ret	GJR-Ret	HN-Opt-Ret	HN-Ret-VIX	GJR-Ret-VIX
Times (h)	0.016	0.024	9.071	0.017	0.036
$-VRP$ (in %)	2.906	2.867	3.011	3.213	3.006
in-$IVRMSE$	0.05739	0.05502	0.05199	0.05217	0.05124
out-$IVRMSE$	0.07004	0.06894	0.06397	0.06488	0.05956

Table 3.6: Comparison, based on the IVRMSE, of empirical pricing performances of the NIG-GARCH (GARCH-1 /GARCH-2, as example: $-0.346\% = 100 \cdot (0.05199 - 0.05217)/0.05199$), using Oprtion-Returns or VIX-Returns information.

GARCH-1	HN-Opt-Ret	HN-Opt-Ret	HN-VIX-Ret
GARCH-2	HN-VIX-Ret	GJR-VIX-Ret	GJR-VIX-Ret
in-sample	−0.346	1.4425	1.7826
out-sample	−1.422	6.8938	8.1997

Table 3.7: Comparison, based on the in-sample IVRMSE, of empirical pricing performances of the HN-GARCH and GJR-GARCH using Gaussian distribution or NIG distribution and Oprtion-Returns or VIX-Returns information. (as example $6.7276\% = 100 \cdot (0.05574 - 0.05199)/0.05574$ where 0.05574 represent the in-sample IVRMSE of the Gaussian-HN-GARCH-Opt-Ret and 0.05199 is the IVRMS of NIG-HN-GARCH-Opt-Ret).

GARCH	G-HN-Opt-Ret	G-HN-VIX-Ret	G-GJR-VIX-Ret
NIG-HN-Opt-Ret	6.7276	6.4047	8.0731
NIG-HN-VIX-Ret	10.377	10.067	11.670
NIG-GJR-Opt-Ret	5.1796	4.8513	6.5475

Table 3.8: Comparison, based on the out of sample IVRMSE, of empirical pricing performances of the HN-GARCH and GJR-GARCH using Gaussian distribution or NIG distribution and Oprtion-Returns or VIX-Returns information. As example $12.835\% = 100 \cdot (0.07339 - 0.06397)/0.07339$ where 0.07339 represent the out of sample IVRMSE of the Gaussian-HN-GARCH-Opt-Ret and 0.06397 is the out of sample IVRMS of NIG-HN-GARCH-Opt-Ret.

GARCH	G-HN-Opt-Ret	G-HN-VIX-Ret	G-GJR-VIX-Ret
NIG-HN-Opt-Ret	12.8355	12.9778	1.5846
NIG-HN-VIX-Ret	11.5955	11.7399	0.1846
NIG-GJR-Opt-Ret	18.8445	18.9770	8.3692

Chapter 4

Discriminating between GARCH models for option pricing by their ability to compute accurate VIX measures

Contents

An increasing body of literature on time series analysis has been developed for the challenges of modeling volatility using the GARCH framework. The literature goes back to [Engle (1982)] with the class of autoregressive conditional heteroscedastic (ARCH) models. The latter represented the first theoretical attempt in order to model volatility as an endogenous time-varying process which incorporated conditional variance clustering to bring the model closer to reality. Nevertheless, except for when the maximum lag of the ARCH model is large, this approach fails to reproduce the decay rate of the unconditional autocorrelation function of squared log-returns observed in financial time series. [Bollerslev (1986)] proposed the parsimonious Generalized ARCH (GARCH) model[1] allowing for a much more flexible lag structure through a moving average component. Over time, original parametric

[1] Alternative approaches were proposed by [Taylor (1986)] and [Heston (1993)] using stochastic volatility models where information from the volatility structure is needed to estimate parameters of the model.

specifications of the conditional variance have been generalized and extended in various directions to increase the practical flexibility of the model and incorporate in particular asymmetric effects and non-Gaussian innovations (see for example [Chorro *et al.* (2015)] Chapter 2).

Concerning the pricing of derivatives, [Duan (1995)] was the first paper to propose a coherent theoretical framework, namely the locally risk-neutral valuation relationship (LRNVR), to obtain risk-neutral dynamics of Gaussian GARCH models. This methodology was popularized by Heston and Nandi in [Heston & Nandi (2000)] where a discrete time affine GARCH-type model with Gaussian innovations was able to replicate one of the key features observed in continuous time literature ([Heston (1993)]): the fact that the no-arbitrage price of classical European options had semi-closed-form expression.[2]

In order to improve the numerical performances of Duan's option pricing model, three complementary areas may be explored:

1. Use more realistic GARCH processes coping with asymmetric volatility responses and non-Gaussian conditional distributions,
2. Use different risk-neutralization processes compatible with the preceding point,
3. Use, when it is possible, more information than just that of the log-returns to estimate the model.

The first point is now a classic topic and many extensions have been proposed to cope with these well-documented stylized facts. The asymmetric effects of positive and negative shocks of equal magnitude on conditional volatility, the so-called leverage effect, may be captured using a large family of extended GARCH models the most popular being probably the exponential EGARCH of [Nelson (1991)], the NGARCH model of [Engle & Ng (1993)], the GJR-GARCH of [Glosten *et al.* (1993)], the threshold GARCH of [Zakoian (1994)], and the affine HN-GARCH by [Heston & Nandi (2000)]. However, the leverage parameter of preceding specifications is not sufficient to capture all the skewness and kurtosis levels in standardized residuals. Therefore, Gaussian hypothesis for the conditional distribution of log-returns has to be relaxed and a myriad of possible choices may be used to take into account all the mass in the tails and the asymmetry ([Chorro *et al.* (2015)] Chapter 2). Among them, the Generalized Hyperbolic ([Chorro *et al.* (2012)], [Badescu *et al.*(2011)]) family and its Normal Inverse Gaussian (NIG) subclass ([Stentoft (2008)], [Badescu *et al.*(2015)]), the Inverse Gaussian (IG) distribution ([Christoffersen *et al.* (2006)]), or the mixture of Gaussian ([Badescu *et al.*(2008)]) clearly improve forecasting performances of related GARCH models.

Once a competing model has been chosen, the choice of the so-called stochastic discount factor (SDF) to obtain risk-neutral dynamics is fundamental. For this second point, two

[2] In the Duan's framework, the coefficients of the GARCH risk-neutral dynamics are just functions of the historical ones, and so may be directly estimated from the log-returns. Nevertheless, the closed-form expression permits to efficiently use available option information to calibrate the model.

constraining factors apply: this SDF has to be sufficiently flexible to provide explicit risk-neutral dynamics for a large variety of GARCH structures and innovation distributions and rich enough to produce good pricing performances. Since the seminal paper of Duan, several tools have been developed to select an equivalent martingale measure (see for example [Chorro *et al.* (2015)] Chapter 3).[3]

One of the main advantages of GARCH models, with respect for example to stochastic volatility ones, is that they may be efficiently estimated using a conditional version of the maximum likelihood estimation and a dataset of log-returns. In particular, since, in the case of exponential-affine or extended Girsanov principle SDF, the associated risk-neutral dynamics are explicit transforms of the historical ones, only log-returns information is needed to compute or approximate European option prices.[4] Even so, when an extra piece of financial information (price of plain vanilla options, the VIX index for the S&P500,...) is available it can be of interest to integrate it, in an efficient way, to the estimation process to reduce pricing errors. Therefore, following [Christoffersen *et al.* (2012)] it is now classically possible to build for some affine GARCH models (at the very least for the HN-GARCH [Heston & Nandi (2000)] and the IG-GARCH [Christoffersen *et al.* (2006)] where semi-closed form expressions for option prices are obtained) a joint maximum likelihood based on log-returns and option prices. In this setting, the affine structure of the model is mandatory: if prices are evaluated using Monte-Carlo methods, computing the likelihood function may be cumbersome. In a recent study, [Hao & Zhang (2013)] have computed VIX index formulas implied by various non-affine asymmetric Gaussian GARCH models. They presented closed-form formulas for the VIX index associated with five classical non-affine Gaussian GARCH models when [Duan (1995)] LRNVR is used. Based on this result, [Kanniainen *et al.* (2014)] proposed a fair comparison between affine and non-affine Gaussian GARCH specifications using log-returns and VIX information in the estimation.[5] For two affine GARCH models [Chorro & Fanirisoa (2016)] and [Papantonis (2016)] proved that incorporating both the physical return dynamics of the index and risk-neutral dynamics

[3]The exponential-affine SDF, M^{ess}, developed by [Bühlmann *et al.* (1996)] and [Siu *et al.* (2004)], which is based on a conditional extension of the pioneering work of [Esscher (1932)], and the SDF given by the extended Girsanov principle of [Elliott & Madan (1998)] are probably the two best known. In particular, they coincide with Duan LRNVR in the Gaussian setting. Let us also remark that extended and non-monotonic versions of the exponential-affine SDF are available for particular choices of distributions as the exponential-quadratic SDF M^{Qua} of [Monfort & Pégoraro (2012)] (see also [Christoffersen *et al.* (2013)]) for Gaussian innovations and the exponential U-shaped stochastic discount factor M^{Ush} proposed by [Chorro & Fanirisoa (2016)] in the second chapter of the present dissertation for the Inverse-Gaussian GARCH model.

[4]This is not true for M^{Qua} or M^{Ush} because, in this case, a risk-neutral parameter (the constant proportional wedge between historical and risk-neutral volatilities) has to be evaluated.

[5]Recently, a large number of studies have further investigated the ability of the VIX index as an input variable for volatility to forecast option prices. Considered as an expected volatility series, the VIX was proposed by [Whaley (1993)] and introduced by the CBOE in 1993 to serve as a market volatility indicator. The VIX captures how much the investor is willing to pay to deal with investment risks. In previous empirical papers on the importance of the VIX index, the attention focus has primarily been on the impact and the correlation of the VIX index with the stock market and returns volatility. [Giot *et al.* (2005)] and [Sarwar (2012)] have established empirical results that suggest an asymmetric relationship between stock market returns and VIX. [Cochrane *et al.* (2012)] observed the adequacy of the VIX index as an important factor in the determination of stock market returns and also of volatility.

of the VIX to estimate the parameters of GARCH option pricing models provides competitive pricing errors at a very low computational cost.[6]

This chapter attempts to fill several gaps in the GARCH option pricing literature, in particular, from an empirical point of view.

Firstly, in the spirit of [Christoffersen *et al.* (2004)] the aim of our study is to provide an intensive comparison analysis of empirical performances, in VIX index or options valuation, between different GARCH-type models using Gaussian or non-Gaussian distributions under different classes of risk-neutral measures. Furthermore, particular attention is granted on the choice of the information set (VIX, options, returns) in the estimation process. To keep the empirical analysis manageable, we only focus our attention on four classical parsimonious GARCH(1,1) structures: HN-GARCH by [Heston & Nandi (2000)], GJR-GARCH by [Glosten *et al.* (1993)], NGARCH by [Engle & Ng (1993)], and IG-GARCH by [Christoffersen *et al.* (2013)].[7] One advantage of this choice is to question the difference between affine and non-affine models. As a natural non-Gaussian alternative we favor the so-called NIG distribution not only because it is known to fit statistical properties of asset returns remarkably but also because, combined with Esscher and EGP SDF, pricing equations may be solved explicitly.[8] Furthermore, monotonic and non-monotonic pricing kernels ([Monfort & Pégoraro (2012)] and [Chorro & Fanirisoa (2016)]) are considered for Gaussian and IG distributions. To our knowledge, in the existing literature, empirical studies questioned, in general, the impact of the distribution ([Christoffersen *et al.* (2006)], [Chorro *et al.* (2012)]), the choice of the SDF ([Badescu *et al.*(2011)], [Christoffersen *et al.* (2013)], [Chorro & Fanirisoa (2016)]) or the estimation strategy ([Hao & Zhang (2013)], [Kanniainen *et al.* (2014)], [Papantonis (2016)], [Lalancette & Simonato (2017)]) on pricing performances, but few of them consider all these factors at the same time.[9] Our study is a means of making a contribution to understand the combined impact of these complementary aspects (24 combinations of GARCH-distribution-SDF-estimation are tested).

Secondly, inspired by the work of [Hao & Zhang (2013)] that explained poor pricing performances of Gaussian GARCH models by their inefficiency to capture the variance risk premium, we also explore in this chapter if it is possible to partly classify GARCH

[6]When closed-form expressions are not available, two recent studies proposed interesting alternatives. In [Lalancette & Simonato (2017)] the authors proposed, for the NGARCH model with Johnson S_U distributed driving noise, numerical approximations to make possible the computation of the implied VIX index using Monte-Carlo simulations. In [Chorro & Fanirisoa (2017)] a new estimation strategy for some non-Gaussian GARCH models is presented to include options or VIX information in the joint estimation at a low computational cost.

[7]An equivalent study could be by performed in a companion paper for Markov-switching [Elliott *et al.* (2006)], multi-component [Christoffersen *et al.* (2008)] and multiple-shock [Christoffersen *et al.* (2012)] GARCH models.

[8]Such a property is not fulfilled if we use, for example, a mixture of Gaussian distributions.

[9]For example, in [Kanniainen *et al.* (2014)] the authors study different GARCH structures with different estimation strategies, but restrict themselves to the Gaussian setting.

option pricing models by their ability to simply reproduce the VIX index. In fact, up to our knowledge, the correlation between the option pricing performances of a model and its ability to compute accurate VIX measures is not clearly established in the literature. A challenging aspect of the present study is to make VIX analysis a first-stage filter to discard the worst GARCH option pricing models. From purely numerical aspects, such a conclusion would be very interesting to back-test these models in an efficient way, using only VIX information, when available, instead of complex option datasets.

This chapter is structured along the following lines. In section 1 we first provide a partial presentation of all competing GARCH frameworks used in the empirical part. More precisely, weconsider four GARCH structures for modeling volatility as a time-varying process: HN-GARCH, GJR, NGARCH, and IG-GARCH. Then, in section 2, we recap the main risk-neutralized frameworks adopted in this chapter. Next, in section 3, we derive the related VIX index formulas. Section 4 deals with the estimation challenge, presenting methodologies based on different information sets and the related numerical results in terms of VIX approximation and option pricing. We conclude in section 5.

4.1 Competing GARCH models

We consider a financial asset with a market price at time t given by S_t and we denote by $Y_t = \log\left(\frac{S_t}{S_{t-1}}\right)$ the associated log-returns defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where $\mathbb{P}$ represents the historical probability measure. Information filtration $\{\mathcal{F}_t\}_{0 \leqslant t \leqslant T}$ is generated by log-returns supposing that $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_T = \mathcal{F}$. In what follows, we consider a general dynamics for the stock price process:

$$\begin{aligned} Y_t &= r + m_t + \sqrt{h_t} z_t \\ h_t &= F\left(z_{t-1}, h_{t-1}, \theta^V\right) \end{aligned} \tag{4.1}$$

where the z_t are i.i.d centered and reduced random variables depending on a vector of parameters θ^D, m_t is the predictable time-varying excess of returns, r is the risk-free rate and F is a mapping, compatible with realistic GARCH(1,1) volatility models that depends on a vector of parameters θ^V.[10] For our empirical horse-race we favor four particular GARCH specifications often used in the literature to cope with volatility clustering and leverage effect. Moreover, these four GARCH-type models belong to two important families: affine and non-affine frameworks. While affine GARCH models are often used because they yield a semi-closed form solution for prices of European equity options, it is now well-documented that non-affine ones provide a better fit to financial data. One important aspect of our empirical study will be to question once again this duality. Following [Kanniainen *et al.* (2014)] we choose the widely recognized NGARCH [Engle & Ng (1993)], GJR-GARCH [Glosten *et al.* (1993)], and affine HN-GARCH [Heston & Nandi (2000)] models and we add the IG-GARCH of [Christoffersen *et al.* (2006)] (see also [Chorro & Fanirisoa (2016)]) that is a notable example

[10] From now on h_0 is supposed to be constant and fixed at its unconditional level depending on the persistence of the model Ψ.

of an affine model within a non-Gaussian setting. In the next sections we briefly recall the definitions and the main properties of these specifications.

4.1.1 Affine competitors

Since the seminal work of [Heston (1993)], affine models, that led to semi-closed form expressions for option prices, are the keystone of almost all numerical studies. In the discrete time literature, the HN-GARCH [Heston & Nandi (2000)] and the IG-GARCH of [Christoffersen *et al.* (2006)] are two important contributions. More precisely, for the HN-GARCH model the historical dynamics is given by

$$\begin{cases} Y_t &= r + \lambda_0 h_t + \sqrt{h_t} z_t \\ h_t &= a_0 + a_1 \left(z_{t-1} - \gamma \sqrt{h_{t-1}} \right)^2 + b_1 h_{t-1} \end{cases} \tag{4.2}$$

with $a_0 > 0$, $a_1 \geq 0$, $b_1 \geq 0$ and for the IG-GARCH specification by

$$\begin{cases} Y_t &= r + \nu h_t + \eta z_t \\ h_t &= w + b h_{t-1} + c z_{t-1} + a \frac{h_{t-1}^2}{z_t} \end{cases} \tag{4.3}$$

with $w > 0$, $b \geq 0$, $c \geq 0$, and $a \geq 0$.

In the HN-GARCH model the z_t are supposed to be Gaussian while in the IG-GARCH they follow an Inverse Gaussian distribution with degree of freedom $\delta_t = \frac{h_t}{\eta^2}$.[11] The persistence (that will be an important quantity to express associated VIX index formula) of the HN-GARCH (resp. IG-GARCH) is given by $\Psi = b_1 + a_1\gamma^2$ (resp. $\Psi = b + \dfrac{c}{\eta^2} + a\eta^2$). Under these two hypotheses on the distributions of innovations, it is easy to prove for both models that the conditional moment generating function $\mathbb{G}^{\mathbb{P}}_{\log(S_T)|\mathcal{F}_t}(u) = E_{\mathbb{P}}[S_T^u \mid \mathcal{F}_t]$ of the log asset price under the physical measure can be written in the following log-linear form $\mathbb{G}^{\mathbb{P}}_{\log(S_T)|\mathcal{F}_t}(u) = S_t^u e^{A_t + B_t h_{t+1}}$ where the coefficients [12] A_t and B_t can be obtained by working backward from the maturity date of the option and using terminal conditions $A_T = B_T = 0$. Moreover, one important empirical consequence for the pricing of European call options is that the very particular form of the conditional moment generating function of $log(S_T)$ leads to the existence of semi-closed form expressions for prices which allow us to use Fast Fourier Transform (FFT) methodology and option information in the estimation procedure as explained in [Chorro *et al.* (2015)] Chap 4.

[11] The density function of the IG distribution is given by the one parameter family: $\mathbf{1}_{\{y>0\}} \dfrac{\delta}{\sqrt{2\pi y^3}} e^{-\left(\sqrt{y} - \delta/\sqrt{y}\right)^2/2}$ where $\delta \in \mathbb{R}_+^*$.

[12] $A_t = ru + A_{t+1} + a_0 B_{t+1} - \frac{1}{2} log(1 - 2a_1 B_{t+1})$ and $B_t = -\frac{1}{2}u + b_1 B_{t+1} + \left(\frac{u^2}{2} - 2a_1\gamma B_{t+1} u + a_1 B_{t+1}\gamma^2 \right)(1 - 2a_1 B_{t+1})^{-1}$ for the HN-GARCH model and $A(t) = A_{t+1} + ur + wB_{t+1} - \frac{1}{2} log(1 - 2a(\eta)^4 B_{t+1})$ and $B(t) = bB_{t+1} + u\nu + (\eta)^{-2} - (\eta)^{-2}\sqrt{(1 - 2a(\eta)^4 B_{t+1})(1 - 2cB_{t+1} - 2u\eta)}$ for the IG-GARCH.

4.1.2 Non-affine competitors

In order to propose asymmetric extensions of the original GARCH(1,1) model, one possibility is to modify the so-called news impact curve (NIC) introduced in [Engle & Ng (1993)]. For this purpose, we may shift a symmetric NIC to the right or consider curves centered at 0 allowing for slopes of different magnitudes on either side of the origin. These two approaches were used by [Engle & Ng (1993)] and [Glosten *et al.* (1993)] in order to introduce respectively the popular NGARCH and GJR models. In both cases, a single leverage parameter constrains the response of the conditional variance to depend on the sign of a shock. In the NGARCH model the dynamics [13] of the risky asset under historical probability is given by

$$\left\{\begin{array}{lcl} Y_t & = & r+\lambda_0\sqrt{h_t}-log(E_{\mathbb{P}}[e^{\sqrt{h_t}z_t}])+\sqrt{h_t}z_t \\ h_t & = & a_0+b_1h_{t-1}+a_1h_{t-1}\left(z_{t-1}-\gamma\right)^2 \end{array}\right. \tag{4.4}$$

with $a_0>0$, $b_1\geq 0$, $a_1\geq 0$ and for the GJR model by

$$\left\{\begin{array}{lcl} Y_t & = & r+\lambda_0\sqrt{h_t}-\dfrac{h_t}{2}+\sqrt{h_t}z_t \\ h_t & = & a_0+h_{t-1}\left[b_1+a_1\left(z_{t-1}\right)^2+\gamma\max\left(0,-\left(z_{t-1}\right)\right)^2\right] \end{array}\right. \tag{4.5}$$

with $a_0>0$, $b_1\geq 0$, $a_1\geq 0$, and $\gamma\geq 0$. The persistence of the NGARCH (resp. GJR) is given by $\Psi=b_1+a_1\left(1+\gamma^2\right)$ (resp. $\psi=b_1+a_1+\frac{\gamma}{2}$). Contrary to models presented in the preceding section, here, conditional moment generating function is not an exponential-affine function of the one step ahead volatility. To compute option prices we use in general Monte Carlo approximations. Nevertheless, as remarked in [Hao & Zhang (2013)] VIX implied formulas are available in this non-affine setting at the very least for Gaussian innovations (and other very particular cases as [Badescu *et al.*(2018)]).

4.1.3 A flexible alternative to Gaussian distribution

It is now a well-known fact that forecasting performances of GARCH-type models are improved when using non-Gaussian innovations. Historically, several interesting distributions were proposed to better account for the deviation from normality. In the present chapter we have decided to mainly focus our attention on the Normal Inverse Gaussian (NIG) distribution. This four-parameter family of distributions has been extensively used during the last decade in discrete time literature, especially for pricing issues ([Stentoft (2008)],

[13] For the NGARCH model, we take $m_t=\lambda_0\sqrt{h_t}-log(E_{\mathbb{P}}[e^{\sqrt{h_t}z_t}])$ as proposed in [Badescu *et al.*(2018)]. When innovations are Gaussian the cumulant moment generating function at the point z is equal to $\frac{z^2}{2}$ and we recover the same excess returns as in the GJR model. Nevertheless we will see that this very particular choice leads to a closed form expression for the VIX index associated with the NGARCH model with NIG innovations when extended Girsanov risk-neutralization process is used. This property is remarkable because up to our knowledge this is the unique example in the literature of an explicit VIX index formula within a non-Gaussian setting.

[Badescu *et al.*(2011)], [Guégan *et al.* (2013)], [Badescu *et al.*(2018)]): for $(\alpha, \beta, \delta, \mu)$ fulfilling $0 < |\beta| < \alpha$ and $\delta > 0$, the density of the NIG $(\alpha, \beta, \delta, \mu)$ is given by[14]

$$d_{NIG}(z, \alpha, \beta, \delta, \mu) = \frac{\alpha}{\pi} e^{\delta\left(\sqrt{\alpha^2-\beta^2}+\beta\left(\frac{z-\mu}{\delta}\right)\right)} \frac{K_1\left(\alpha\delta\sqrt{1+\left(\frac{z-\mu}{\delta}\right)^2}\right)}{\sqrt{1+\left(\frac{z-\mu}{\delta}\right)^2}}$$

where K_1 is the modified Bessel function of the third kind with index one. The mean and the variance of this distribution are respectively given by

$$m = \mu + \frac{\delta\beta}{\sqrt{\alpha^2-\beta^2}}, \qquad \sigma^2 = \frac{\delta\alpha^2}{\sqrt{\alpha^2-\beta^2}^3}. \tag{4.6}$$

Therefore, from the stability of the NIG family under affine transforms, it is possible to obtain a centered version with unit variance considering

$$\text{NIG}\left(\tilde{\alpha}, \tilde{\beta}, \tilde{\delta}, \tilde{\mu}\right) = \frac{\text{NIG}(\alpha, \beta, \delta, \mu) - m}{\sigma} \tag{4.7}$$

where $\tilde{\alpha} = \sigma\alpha$, $\tilde{\beta} = \sigma\beta$, $\tilde{\delta} = \frac{\delta}{\sigma}$ and $\tilde{\mu} = \frac{-m}{\sigma} + \frac{\mu}{\sigma}$.

4.2 Stochastic discount factors and risk-neutral dynamics

From the beginning of the 80's (see [Chorro *et al.* (2015)] Chap 3 and references therein), general methods providing arbitrage-free price processes via the notion of equivalent martingale measure (EMM) have been investigated both in discrete or continuous time frameworks. Furthermore, the choice of such an EMM is known to be equivalent to the specification of the so-called one-period stochastic discount factor (SDF). Since markets described by GARCH models are incomplete, there is a priori an infinite number of SDF available for pricing derivatives and a great challenge is to select tractable candidates for their strong economic foundations and/or empirical performances. In this section, we present the main paths to risk-neutralization that will be implemented in the numerical part to obtain arbitrage-free price approximations in Gaussian or non-Gaussian settings. More specifically, starting from the [Duan (1995)] approach particularly well-adapted to Gaussian residuals, we briefly recall the main lines of the recent advances in modeling SDF dynamics to cope with non-Gaussian innovations ([Elliott & Madan (1998)] extended Girsanov principle (EGP) and [Siu *et al.* (2004)] conditional Esscher transform) and/or have better representations of volatility risk ([Monfort & Pégoraro (2012)], [Chorro & Fanirisoa (2016)]).[15]

[14] Equivalently this distribution may be characterized by its very simple log-moment generation function given by $\kappa_{NIG}(z) = \mu z + \delta\sqrt{\alpha^2-\beta^2} - \delta\sqrt{\alpha^2-(\beta+z)^2}$. This simple expression will be used in (4.4) and also to obtain in a very simple way risk-neutral dynamics in GARCH-type models with NIG innovations.

[15] The purpose of this section is not to provide a self-contained presentation of these classical tools ([Chorro *et al.* (2015)] Chap 3) but to recall the main intuitions behind Gaussian (see Table 4.3) and non-Gaussian (see Table 4.4) risk-neutral dynamics that will be compared in the empirical part.

As in the preceding section, we consider a GARCH-type specification for the log-returns

$$\begin{aligned} Y_t &= r + m_t + \sqrt{h_t} z_t \\ h_t &= F\left(z_{t-1}, h_{t-1}, \theta^V\right) \end{aligned} \tag{4.8}$$

where the z_t are i.i.d centered random variables with unit variance.

Supposing that the z_t are i.i.d $\mathcal{N}(0,1)$, [Duan (1995)] was the first to provide a coherent theoretical CCAPM framework to obtain risk-neutral dynamics in a GARCH environment independently of the underlying GARCH structure. More precisely, if $\mathbb{Q}$ is an EMM fulfilling LRNVR[16] then

$$\begin{aligned} Y_t &= r - \tfrac{h_t}{2} + \sqrt{h_t} z_t^* \\ h_t &= F\left(z_{t-1}^* - \tfrac{m_{t-1}}{\sqrt{h_{t-1}}} - \tfrac{\sqrt{h_{t-1}}}{2}, h_{t-1}, \theta^V\right) \end{aligned} \tag{4.9}$$

where the z_t^* are i.i.d $\mathcal{N}(0,1)$ under $\mathbb{Q}$. For Gaussian models presented in the preceding section, risk-neutral dynamics deduced from the Duan's argument are given in Table 4.3. In the non-affine GJR and NGARCH setting, prices may be obtained from (4.9) using Monte Carlo approximations while in the affine HN case semi-closed form formulas are available. Nevertheless, Duan's framework relies on Gaussian hypotheses and cannot be adapted with simplicity to more general distributions. Based on this observation, [Elliott & Madan (1998)] proposed a very simple way to select a SDF based on a Girsanov-type transformation that preserves returns distribution after the change of measure by only shifting the conditional mean to fulfill the martingale restriction:[17] under the EMM $\mathbb{Q}^{EGP}$ we have

$$\begin{aligned} Y_t &= r + m_t - \nu_t + \sqrt{h_t} z_t^* \\ h_t &= F(z_{t-1}^* - \tfrac{\nu_{t-1}}{\sqrt{h_{t-1}}}, h_{t-1}, \theta^V) \end{aligned} \tag{4.10}$$

where z_t^* follows the same law as z_t under $\mathbb{P}$ and where ν_t fulfills $e^{\nu_t} = e^{-r} E_{\mathbb{P}}\left[e^{Y_t} \mid \mathcal{F}_{t-1}\right]$. When the z_t are assumed to be Gaussian, we recover the same dynamics as in (4.9). Moreover, following [Badescu *et al.*(2018)], for NIG innovations this is a tractable framework, especially when combined with the NGARCH model to obtain a closed-form formula for the associated VIX index.[18] Nevertheless, one of the major drawback of this approach, that may explain partly poor pricing performances of this method for long maturity options (see [Badescu *et al.*(2008)] and [Badescu *et al.*(2011)]), is the fact that from $\mathbb{P}$ to $\mathbb{Q}^{EGP}$ only the conditional mean is affected while the conditional variance, skewness and, kurtosis are the same. The conditional Esscher transform introduced in the GARCH setting by [Siu *et al.* (2004)] and [Gouriéroux & Monfort (2007)] is probably one of the best-known tool to select efficiently EMM. The associated SDF M^{ess} is exponential-affine of log-returns and the predictable associated coefficients of affinity are uniquely determined by the pricing equations related to the bond and the risky asset. In contrast to Duan's approach a wide variety

[16] A set of assumptions made on the utility function and the aggregated consumption growth that preserves both Gaussianity and volatility.

[17] Such a pricing kernel has also been justified from its consistency with risk-adjusted cost minimizing hedging strategies.

[18] In fact, the restriction imposed on the conditional mean in (4.4) provides explicit computations.

of return innovations may be chosen at the very least within the class of mixture or infinitely divisible distributions (see [Chorro *et al.* (2015)] Chap 3.4). Even if this tool coincides with the LRNVR in the Gaussian case, it allows for strongly non-linear relations between historical and risk-neutral volatility in the non-Gaussian setting. Furthermore, explicit[19] risk-neutral dynamics (see Table 4.4) may be obtained for the IG-GARCH model (4.3) and GARCH-type models with NIG innovations. In particular, if we suppose in (4.8) a NIG $\left(\tilde{\alpha}, \tilde{\beta}, \tilde{\delta}, \tilde{\mu}\right)$ for the z_t, we obtain ([Badescu *et al.*(2011)]) the following dynamics under the Esscher EMM:

$$\begin{aligned} Y_t &= r + m_t + \sqrt{h_t} z_t^* \\ h_t &= F\left(z_{t-1}^*, h_{t-1}, \theta^V\right) \end{aligned} \tag{4.11}$$

where z_t^* follows, under $\mathbb{Q}^{Ess}$, a NIG$(\tilde{\alpha}, \tilde{\beta} + \sqrt{h_t}\theta_t^q, \tilde{\delta}, \tilde{\mu})$ with a predictable parameter θ_t^q having an explicit form. As remarked in [Monfort & Pégoraro (2012)], the exponential-affine hypothesis concerning the SDF only allows for an equity risk premium and it may be interesting to partly solve empirical puzzles of option prices taking into account a second-order variance risk premium. To achieve this, the authors introduced an exponential-quadratic SDF M^{Qua} that extends M^{ess} adding a second moment-based source of risk information. Moreover, under Gaussian hypothesis, this new change of measure preserves the tractability of the model with a risk-neutral dynamics given by

$$\begin{aligned} Y_t &= r - \tfrac{h_t^*}{2} + \sqrt{h_t^*} z_t^* \\ h_t^* &= \pi F\left(\sqrt{\pi}\left(z_{t-1}^* - \tfrac{m_{t-1}}{\sqrt{h_{t-1}^*}} - \tfrac{\sqrt{h_{t-1}^*}}{2}\right), \tfrac{h_{t-1}^*}{\pi}, \theta^V\right) \end{aligned} \tag{4.12}$$

where the z_t^* are i.i.d $\mathcal{N}(0,1)$ under $\mathbb{Q}^{Qua}$ and π is the proportional wedge between risk-neutral and historical volatilities assumed to be constant across time.[20] As a consequence, for the HN model (4.2), the dynamics under $\mathbb{Q}^{Qua}$ remains in the same family of affine GARCH models, preserving analytic properties of the HN specification in terms of option pricing. Inspired by this new methodology, [Chorro & Fanirisoa (2016)] proposed an exponential-hyperbolic SDF M^{Ushp} that is able to cope with the same remarkable features in the case of the IG-GARCH model (4.3).

To conclude this section, let us precisely describe all related GARCH option pricing models that will be tested in the empirical part: in the affine family, the classical [Heston & Nandi (2000)] and the IG-GARCH model ([Christoffersen *et al.* (2006)] and [Chorro & Fanirisoa (2016)]) will be combined with exponential-affine and U-shaped SDF risk-neutralization processes. In these cases, Monte-Carlo methods won't be used to approximate the price of plain vanilla options. To relax the constraints on variance dynamics and

[19] In general, to obtain risk-neutral dynamics, pricing equations have to be solved numerically ([Chorro *et al.* (2012)]) at any time. However for some interesting choices (Gaussian, IG, NIG among others) solutions are analytic.

[20] The exponential-quadratic stochastic discount factor can be expressed as $M_t^{Qua} = e^{\theta_{2,t} Y_t^2 + \theta_{1,t} Y_t + \varepsilon_t}$ where $(\varepsilon_t, \theta_{1,t}, \theta_{2,t})$ are predictable coefficients. Obviously, when $\theta_{2,t} = 0$ we recover M^{ess}. The pricing equations for the bond and the risky asset impose some restrictions on these predictable coefficients that are not uniquely defined. If we want to obtain a unique solution to the pricing system an extra condition is needed. A natural candidate is to impose a constant proportional wedge $\pi = \frac{h_t^*}{h_t}$ between risk-neutral and historical conditional variances (see [Chorro *et al.* (2015)] Chap 3.5). This new risk-neutral parameter (that cannot be estimated only using any information from the log-returns) can help producing richer dynamics.

conditional distributions related to affine specifications, we will also study two classical non-affine structures namely the GJR and NGARCH models with Gaussian or NIG innovations. In the Gaussian case the dynamics will be risk-neutralized using the LRNVR or the quadratic SDF while under NIG hypotheses, exponential-affine and EGP assumptions will be favored. This great variety of models and SDF will allow us to question several key aspects of GARCH option pricing modeling. Finally, for sake of concision and simplicity all the risk-neutral dynamics used in this study are gathered in Table 4.3 for Gaussian innovations and in Table 4.4 otherwise.

4.3 Model implied CBOE VIX

Considered as the investor's expectation of volatility (see [Carr & Wu (2006)]), the CBOE VIX index can be characterized as a forecast of the 30-day risk-neutral volatility (or 22 working days) of the S&P500 index. In this chapter, we denote by Vix_t a daily-based proxy for VIX_t which is the daily-adjusted expression of the expected arithmetic average of variance (see [Hao & Zhang (2013)]):

$$\text{Vix}_t = \frac{1}{\tau}\left(\frac{\text{VIX}_t}{100}\right)^2 = \mathbb{E}_{\mathbb{Q}}\left[\frac{1}{T_c}\int_t^{t+T_c} h_u du \mid \mathcal{F}_t\right] \approx \frac{1}{T_c}\sum_{j=1}^{T_c}\mathbb{E}_{\mathbb{Q}}\left[h_{t+j} \mid \mathcal{F}_t\right] \tag{4.13}$$

where $\tau = 250$, $T_c = 22$ represents the maturity in days and $\mathbb{Q}$ is an EMM. Depending on the choice of the risk-neutral dynamics and using iterative properties of conditional expectation, the term $\mathbb{E}_{\mathbb{Q}}[h_{t+j} \mid \mathcal{F}_t]$ can be explicitly computed for a large class of Gaussian ([Hao & Zhang (2013)]) and non-Gaussian ([Chorro & Fanirisoa (2016)], [Badescu *et al.*(2018)]) GARCH models. In general, $\mathbb{E}_{\mathbb{Q}}[h_{t+j} \mid \mathcal{F}_t]$ can be expressed as a linear function of historical volatility at time $t+1$, unconditional variance, and variance persistence[21] under the selected EMM. If we can obtain analytic expressions, we have the following general form for $\mathbb{E}_{\mathbb{Q}}[h_{t+j} \mid \mathcal{F}_t]$ and Vix_t:

$$\left\{\begin{array}{l} \mathbb{E}_{\mathbb{Q}}\left[h_{t+j} \mid \mathcal{F}_t\right] = h_{t+1}\left[\Psi^*\right]^{j-1} + \tilde{h}_0\left[1-(\Psi^*)^{j-1}\right] \\ \text{Vix}_t = h_{t+1}\dfrac{1-(\Psi^*)^{T_c}}{(1-\Psi^*)\,T_c} + \tilde{h}_0\left(1-\dfrac{1-(\Psi^*)^{T_c}}{(1-\Psi^*)\,T_c}\right) \end{array}\right. \tag{4.14}$$

where expressions of $\tilde{h}_0$ and Ψ^* for particular models and SDF are reported in Table 4.5. In fact, for Gaussian models under the LRNVR and for affine models with exponential-affine or U-shaped SDF we have closed form expressions. For example, in the case of the HN model, we obtain $\tilde{h}_0 = \dfrac{a_0+a_1}{1-\Psi^*}$ and $\Psi^* = b_1 + a_1(\gamma+\lambda_0+\frac{1}{2})^2$ when an exponential-affine SDF is used while we obtain $\tilde{h}_0 = \dfrac{a_0+\pi a_1}{1-\Psi^*}$ and $\Psi^* = b_1+\pi^2 a_1\left(\dfrac{\gamma}{\pi}+\dfrac{\lambda_0}{\pi}+\dfrac{1}{2}\right)^2$ under the quadratic SDF.

Unfortunately, in the case of NIG innovations (a notable exception is the NIG NGARCH model associated with the EGP of [Badescu *et al.*(2018)]) or when an exponential-quadratic SDF is used with the Gaussian NGARCH and GJR structures we do not have closed-form

[21] Where the variance persistence is the coefficient in front of h_t in $\mathbb{E}_{\mathbb{Q}}[h_{t+1} \mid \mathcal{F}_t]$.

formulas for the implied Vix_t. However, as explained in [Lalancette & Simonato (2017)] we can still use Monte Carlo simulations to approximate conditional expectation $\mathbb{E}_{\mathbb{Q}}\left[h_{t+j} \mid \mathcal{F}_t\right]$ and Vix_t.

4.4 Methodology and empirical results

In this section, we present the main points emerging from this analysis. First, we carry out numerical experiments to analyze pricing performances of all competing GARCH models, focusing on affine/non-affine structures, the risk-neutralization process and the estimation methodology. A pool of 24 possible combinations (Model/SDF/Estimation) will thereby be tested to try to understand the impact of underlying factors. Furthermore, a second experiment aims to question the possibility of partly ranking GARCH option pricing models by their ability to simply reproduce VIX dynamics, instead of using a heavy set of option data. More specifically, after a brief description of the data, we present the main lines of classical joint likelihood estimation methodologies based on Option-Returns or VIX-Returns data (see for example [Kanniainen *et al.* (2014)] and reference therein) and that of the two-step estimation strategy recently introduced in [Chorro & Fanirisoa (2017)] for NIG-GARCH processes. Then, when closed-form expressions for option prices are not available, we recall how Monte Carlo approximations may be implemented efficiently in the GARCH framework using the powerful and simple adjustment proposed by [Duan & Simonato (1998)]. Finally, this section ends with a presentation of the results based on our empirical findings.

4.4.1 Data description

The present study used S&P500 daily returns and VIX data from January 07, 1999 to December 22, 2010, which are composed of 2718 observations covering about 12 years. We plotted in Figure 4.1 the S&P500 and CBOE VIX indexes with their log-returns series while Table 4.1 displayed associated summary statistics.[22] This information set was used to implement both classical conditional maximum likelihood strategies and joint estimation strategies based on returns and VIX information.

We also used a dataset of options written on the S&P500 obtained from Bloomberg. Due to the number of option pricing models to test in this chapter, we restricted ourselves to Wednesday's contracts.[23] Therefore, it concerned 4563 options contracts whose prices were quoted during the period spanning from January 2nd, 2009 to April 15, 2012. We divided the option data set into two subsets: one in which model parameters are estimated (to implement for the affine models the joint likelihood estimation based on returns and options) and another subset used to compare pricing performances of models. The first subset, used for the in-sample estimation and comparison, is called Dataset A from January 2nd, 2009 until December 22, 2010 and contains 2714 contracts. However, the second subset for the out-of-sample comparison is called Dataset B and contains 1849 contracts with

[22] Let us remark that VIX data from January 03, 2011 until April 15, 2012 are also used in the empirical part to test the ability of GARCH option pricing models to forecast VIX dynamics.

[23] We apply to our dataset the same filters as described in [Christoffersen *et al.* (2012)].

67-Wednesdays from January 03, 2011 until April 15, 2012. This will be used to test the out-of-sample ability to capture the behavior of the index option smile. Summary statistics for option data are reported in Table 4.2 for both Dataset A and B: this table shows the number of contracts, the average price, and the average implied volatility across moneynesses and times to maturity. The patterns in the Dataset B are clearly similar to those in the in-sample Dataset A.

Depending on the chosen estimation strategy, the in-sample dataset of returns is combined with in-sample VIX data or Dataset A to estimate the model as explained in the next section. Futhermore, usual in and out-of-sample option pricing performances are studied: we use in-sample estimated parameters to compute approximate prices (from FFT or Monte-Carlo approximations depending on the structure of the model) for the contracts in Dataset A and B to analyze associated errors. In the out-of sample exercise presented above, we assumed that model's parameters are constant over the whole sample period (Dataset B). Obviously, this may appear as unrealistic and unfair for the simulation and relaxing this assumption will highlight the robustness of our conclusions. Therefore, in a complementary numerical experiment, we allowed model parameters to change over time through a rolling window estimation strategy for the 67 Wednesdays in the Dataset B.[24] For each Wednesday in dataset B, we estimated each model and use corresponding parameters to price options next Wednesday.[25]

4.4.2 Estimation methodologies

In this section we denote by ϑ the set of risk-neutral parameters associated with historical dynamics (4.1). When conditional Esscher transform or extended Girsanov principle are used to obtain risk-neutral dynamics we simply have $\vartheta = (\theta^D, \theta^V)$ while $\vartheta = (\theta^D, \theta^V, \pi)$ in the case of U-shaped pricing kernels.[26] Moreover, we denote by T (resp. N) the number of VIX and log-returns daily observations (resp. N the cardinal of the set of option market prices) involved in the estimation process. One of the main advantages of the GARCH machinery is that historical model parameters (θ^D, θ^V) may be easily obtained, from a simple log-returns dataset, using a conditional version of the classical maximum likelihood estimator maximizing

$$\log L_R(\theta^D, \theta^V) = \sum_{t=1}^{T} \log \left(\frac{1}{\sqrt{h_t}} f_{\theta^D} \left(\frac{Y_t - (r + m_t)}{\sqrt{h_t}} \right) \right)$$

where f_{θ^D} is the probability density function of the model innovations. However, the proportional wedge between historical and risk-neutral volatility π cannot be estimated only

[24] We assume constant windows of 12 years (resp. 2 years) for log-returns and VIX data (resp. for options).

[25] We particularly use estimated in-sample parameters as initial values for the optimization performed the first Wednesday while we initialize parameters of the following Wednesday estimation process by using parameters obtained the previous week.

[26] As defined in the preceding sections, θ^D is the vector of innovation parameters, θ^V represents the volatility parameters, and π is the proportional wedge between risk-neutral and historical volatilities supposed to be constant.

using returns data. Moreover, during the last decade, several empirical studies underlined the real interest to incorporate in the estimation process VIX or option information, when available, to improve related pricing performances. Therefore, we present below two joint likelihood estimation strategies used in the empirical part:

Joint estimation strategy using Option-Returns information: We consider a set of option market prices $(\hat{c}_1, ..., \hat{c}_N)$ and define associated weighted Vega errors $\varepsilon_i = \frac{c_i - \hat{c}_i}{\hat{V}_i}$ where c_i and $\hat{V}_i$ are the model prices and the Black and Scholes Vega associated with $\hat{c}_i$. Following [Trolle & Schwartz (2009)], we suppose that the (ε_i) are i.i.d centered Gaussian variables with variance $\frac{1}{N}\sum_{i=1}^{N}\varepsilon_i^2$. Therefore, the associated option log-likelihood is given by

$$\log L_{Op}(\vartheta) = -\frac{1}{2}\sum_{i=1}^{N}\left[\log\left(\frac{1}{N}\sum_{i=1}^{N}\varepsilon_i^2\right) + \frac{\varepsilon_i^2}{\frac{1}{N}\sum_{i=1}^{N}\varepsilon_i^2}\right]$$

and we obtain the joint Option-Returns likelihood (see [Christoffersen *et al.* (2013)]):

$$\frac{T+N}{2}\frac{\log L_R((\theta^D, \theta^V))}{T} + \frac{T+N}{2}\frac{\log L_{Op}(\vartheta)}{N}. \tag{4.15}$$

One of the major drawbacks of this approach is the requirement to evaluate several times the objective function (4.15) in the maximization process. In the case of affine GARCH models presented above, independently of the choice of the exponential-affine or exponential U-shaped SDF, closed-form expressions for option prices are available and make this process computationally acceptable. As noticed in Section 4.3, for most of Gaussian GARCH specifications and for the NIG NGARCH model combined with the EGP it is possible to obtain closed-form expressions for the implied VIX. Therefore, as provided by [Kanniainen *et al.* (2014)], a similar strategy based on VIX information and not on options one may be implemented.

Joint estimation strategy using VIX-Returns information: To build the VIX log-likelihood we suppose with [Kanniainen *et al.* (2014)] (see also [Chorro & Fanirisoa (2016)] or [Badescu *et al.*(2018)]) that VIX pricing errors $u_t = \text{VIX}_t^{Market} - \text{VIX}_t^{Model}$ follow autoregressive disturbances $u_t = \varrho u_{t-1} + e_t$ where $(e_t)_t$ are i.i.d Gaussian random variables with mean zero and variance Σ^2 are where $\mid \varrho \mid < 1$ to ensure stationarity. Consequently the VIX log-likelihood is given by

$$\begin{aligned}\log L_{\text{VIX}}(\vartheta, \varrho) \;&= -\frac{T}{2}\left(\log(2\pi) + \log(\Sigma(1-\varrho^2))\right) + \frac{1}{2}\left(\log(1-\varrho^2)\right)\\ &-\frac{1}{2\Sigma}\left(u_1^2 + \sum_{t=2}^{T}\frac{(u_t - \varrho u_{t-1})^2}{1-\varrho^2}\right)\end{aligned} \tag{4.16}$$

and we obtain the joint VIX-returns likelihood $\left(\log L_R(\theta^D, \theta^V) + \log L_{\text{VIX}}(\vartheta, \varrho)\right)$.

Finally, a last estimation strategy will be used in the empirical part for particular GARCH models with NIG innovations. This strategy, first introduced in [Chorro & Fanirisoa (2017)], derives from a very simple finding: under Gaussian hypotheses, some GARCH-type models have outstanding properties (closed-form expressions for VIX and/or option prices) that fail when NIG innovations are involved. Therefore, inspired by the so-called quasi-maximum likelihood (QML) estimator, a two-step approach is possible to take benefit of these remarkable features in a Gaussian environment:

Two-step estimation strategies using VIX-Returns or Option-Returns information: As in the QML approach, this two-step strategy estimates separately volatility and distribution parameters assuming Gaussian innovations in the first step. We start from a GARCH-type model with NIG innovations

- Step 1: We assume that the $(z_t)_t$ are i.i.d $\mathcal{N}(0,1)$ under $\mathbb{P}$. Subsequently, we can estimate the vector of volatility parameters θ^V as follows :
  - In the cases where we have a closed-form formula for option prices, $\hat{\theta}^V$ may be obtained by maximizing the joint Option-Returns likelihood (4.15).
  - In the cases where we do not have a closed-form formula for option prices but we have a closed-form formula for the VIX index then $\hat{\theta}^V$ may be obtained maximizing the joint VIX-Returns likelihood.
- Step 2: From the i.i.d residuals $\left(z_1\left(\hat{\theta}^V\right),\cdots,z_T\left(\hat{\theta}^V\right)\right)$ that may be extracted from the previous step, the distribution vector of parameters θ^D is obtained maximizing

$$\sum_{t=1}^{T} -\frac{\log(h_t)}{2} + \log\left[f_{\theta^D}\left(\frac{Y_t-(r+m_t)}{\sqrt{h_t}}\right)\right]$$

  where f_{θ^D} is the density function of a centered NIG random variable with unit variance as introduced in Section 4.1.3.

To summarize, in our empirical study, the HN model with Gaussian innovations and the IG-GARCH model (risk-neutralized using Esscher or U-shaped SDF) will be estimated using the returns, the joint VIX-Returns and the joint Option-Returns likelihoods. The GJR and NGARCH models with Gaussian innovations (risk-neutralized using Esscher SDF) will be estimated using the returns and the joint VIX-Returns likelihood. The HN, GJR, NGARCH with NIG innovations (risk-neutralized using Esscher SDF) will be estimated using the returns and the two-step estimation strategy. The GJR and NGARCH models with Gaussian innovations (risk-neutralized using the quadratic SDF) will be estimated using the joint VIX-Returns likelihood.[27]

The estimated parameter values and their respective standard errors, obtained from using the different sets of information, are reported in Table 4.6 (resp. Table 4.14) for Gaussian

[27] In this case, and only in this case, the methodology of [Lalancette & Simonato (2017)] will be used to approximate the VIX using Monte Carlo methods.

GARCH models combined with the exponential-affine (resp. the quadratic) SDF. For NIG parameters, the results of the two-step estimation exercises are presented in Table 4.10, while Table 4.20 shows estimates for the IG-GARCH model under both M^{Ess} and M^{Ushp}. Finally, for the NIG-NGARCH model risk-neutralized using the EGP, the joint VIX-Returns likelihood estimates are illustrated in Table 4.18. In all cases, results are roughly in the same range as those obtained in many other previous empirical studies.

We notice for the IG-GARCH model that parameter estimates are remarkably stable across the different approaches. Concerning the other GARCH specifications, instead of focusing on the individual values of each parameter, we remark that global features of each model (implied persistence, leverage effect parameter) differ only a little from one strategy to another.[28] We classically obtain high historical persistences and all models and estimation approaches clearly indicate the leverage effect. Moreover, in the case of the two U-shaped pricing kernels, the proportional wedge between the risk-neutral and the historical volatilities is significantly estimated to be greater than 1, with values ranging between 1.24 and 1.72 (see Tables 4.14 and 4.20) for the Gaussian HN and the IG-GARCH models, as observed in empirical studies. Last but not least, as remarked in [Kanniainen *et al.* (2014)], for the joint VIX-Returns estimation strategy, the autocorrelation coefficient ϱ is uniformly close to 1 with a minimum value of 0.81 for the Gaussian HN model combined with the quadratic SDF.

Concerning parameters of the NIG distribution, we can see from Tables 4.10 and 4.18 that the observed (negative) values of skewness vary from -0.01 to -0.34 and that observed kurtosis vary from 1.42 to 2.62. These values provide evidence by their departure from normality and they are in the same range as those obtained in previous studies (see for example [Badescu *et al.*(2011)]).

4.4.3 Criteria for the option and VIX pricing analysis

Once a particular GARCH model has been properly estimated using a well-chosen set of historical financial information, we obtain explicitly from Tables 4.3 and 4.4 the related risk-neutral dynamics depending on the choice of the underlying SDF. For the HN-GARCH model with Gaussian innovations ([Heston & Nandi (2000)] and [Monfort & Pégoraro (2012)]) and the IG-GARCH model ([Christoffersen *et al.* (2006)] and [Chorro & Fanirisoa (2016)]), under both exponential-affine and U-shaped SDF, we have quasi-closed-form solutions for pricing vanilla European options efficiently from FFT methodology (see for example [Chorro *et al.* (2015)] Chap 4.2) that massively decrease the required time to price a full option book. For other non-affine specifications, prices are approximated using Monte Carlo simulation using 15000 trajectories.[29]To test the quality of these price approximations we

[28]For example, we can deduce from Tables 4.6 and 4.14 that in the case of the GJR GARCH specification we obtain historical (resp. risk-neutral) persistences around 0.986 (resp. around 0.996) and a leverage parameter γ between 0.022 and 0.023.

[29]An important point to emphasize here is the use in our study of the so-called empirical martingale simulation methodology (EMS) proposed by [Duan & Simonato (1998)] to reduce drastically the variance of Monte Carlo estimators. As remarked for example in [Badescu *et al.*(2015)], EMS is an essential tool to improve numerical efficiency of Monte Carlo methods especially in the GARCH setting and to use a reasonable

will use, in the empirical part, the in (Dataset A), out (Dataset B) and Wednesday (rolling window strategy) Implied Volatility Root Mean Squared Error (IVRMSE[30]) that measure the discrepancy between model and option prices:

$$IVRMSE = \sqrt{\frac{1}{N}\sum_i \left(\frac{c_i - \hat{c}_i}{\hat{V}_i}\right)^2}$$

where c_i is the option price given by the model, $\hat{c}_i$ the corresponding market price and $\hat{V}_i$ the Black and Scholes Vega associated with $\hat{c}_i$. Moreover, another interesting economic criteria will be the magnitude of the average annualized volatility risk premium (VRP) as defined in [Papantonis (2016)] in order to understand why an equity risk premium is in general not sufficient to produce realistic price levels. Finally, in order to discuss the correlation between option pricing performances and the capacity of implied VIX to fit the market VIX, we will use the measures of adequacy introduced in [Qiang *et al.* (2015)], namely, the mean percentage error (MPE_{VIX}), the mean percentage absolute error (MAE_{VIX}) and the root mean squared error ($RMSE_{\text{VIX}}$) defined below:

$$MPE_{\text{VIX}} = \frac{1}{N}\sum_{j=1}^{N}\left(\frac{\text{VIX}_j^{Model}}{\text{VIX}_j^{Market}} - 1\right),\ MAE_{\text{VIX}} = \frac{1}{N}\sum_{j=1}^{N}\left(\left|\frac{\text{VIX}_j^{Model}}{\text{VIX}_j^{Market}} - 1\right|\right)$$
$$\text{and } RMSE_{\text{VIX}} = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(\text{VIX}_j^{Model} - \text{VIX}_j^{Market}\right)^2}. \tag{4.17}$$

4.4.4 Empirical findings

Our study relies on 24 combinations of GARCH-distribution-SDF-estimation. To make the presentation much more readable, we group them into five different categories: the Gaussian-GARCH models combined with M^{Ess}, the NIG-GARCH models combined with M^{Ess}, the Gaussian-GARCH models combined with M^{Qua}, the NIG-NGARCH model risk-neutralized using the EGP, and the IG-GARCH model. For each group, we present in a specific table (see Tables 4.7, 4.11, 4.15, 4.19 and 4.21) option and VIX fitting performances based on the criteria introduced in the preceding section. Furthermore, we report for each model the related estimation time and the variance risk premium as defined in [Papantonis (2016)]. For a selected subclass containing more than one element, we provide internal pairwise comparisons in terms of computational time of estimation and in-sample pricing performances (see Tables 4.8, 4.12, 4.16 and 4.22) and in terms of out-of-sample and weekly out-of-sample option valuation errors (see Tables 4.9, 4.13, 4.17 and 4.23). Finally, general results are provided to allow for broader conclusions: in Table 4.24, out-of-sample

number of simulations to compute option prices. Nevertheless, in spite of their efficiency when combined with EMS, Monte Carlo approximations for non-affine models make impossible to use option information in the estimation process at a realistic computational cost.

[30]In the bulk of recent studies ([Christoffersen *et al.* (2012)], [Kanniainen *et al.* (2014)], [Chorro & Fanirisoa (2016)], [Badescu *et al.*(2017)]) this indicator was used to measure pricing performances because Vega-weighted errors do not vary too much across maturities and moneyness contrary to price errors.

performances of the best models in each category are compared while we can find in Table 4.25 (resp. Table 4.26) a summary of VIX and option performance measures of the 24 competitors (resp. the corresponding ranking). Regarding results presented in Table 4.26 we can easily notice that ranks related to option (resp. to VIX) valuation are mostly independent of the choice of the underlying criteria selected from in sample, out-of-sample or weekly out-of sample IVRMSE (resp. from RMSE, MPE or MAE). Thus, in the following, numerical comparisons will rest on out-of-sample IVRMSE and VIX RMSE. We start our analysis at a group level.

We deduce from Table 4.7 that, when they are estimated only using returns, pricing performances of Gaussian GARCH models seem to be independent of the choice of the GARCH structure with IVRMSE ranging from 0.07648 to 0.07770 under Duan's LRNVR. When an extra piece of financial information is introduced into the estimation process, we obtain the smallest IVRMSE of 0.065 for the non-affine specifications especially the GJR model. This is in line with the existing literature that favors non-affine Gaussian stochastic volatility models (see [Kanniainen *et al.* (2014)] and references therein). Table 4.11 leads to similar conclusions in the NIG environment while Table 4.15 confirms the slight superiority of non-affine Gaussian specifications when using an exponential-quadratic SDF. Nevertheless, option valuation errors under Gaussian distribution and exponential-affine SDF are the worst of all competitors. A plausible explanation comes from the fact that these models generate very small variance risk premia (see Table 4.7) which are not in line with empirical observations. In fact, as reported in Tables 4.11, 4.15, 4.19 and 4.21, when we use non-Gaussian alternatives and/or U-shaped pricing kernels we recover VRP between -2.867% (for the NIG-GJR model estimated using returns only) and -3.75411% (for the IG-GARCH model estimated using Option-Returns information) that are in line with a bulk of empirical studies ([Papantonis (2016)]). For Gaussian distribution and exponential-affine SDF, the variance risk is neglected and an equity risk premium is not sufficient to produce realistic price levels.

In Table 4.11, the overall IVRMSE is between 0.05929 and 0.07004 for NIG-GARCH models risk-neutralized with the Esscher SDF with values that are all smaller than corresponding values for Gaussian innovations. The minimal IVRMSE of 0.05929 is obtained in the case of the NIG-NGARCH model by estimating with the two-step estimation strategy using VIX-Returns information as introduced in [Chorro & Fanirisoa (2017)]. Not surprisingly, a finer modeling approach of conditional skewness improves considerably the quality of price approximations. The two-step estimation strategy using VIX-Returns information helps to substantially improve performances at a parsimonious computational cost. The improvement (of around 14% for non-affine specifications) from using VIX information is also fundamental in this framework because returns based estimation strategy only leads to IVRMSE ranging from 0.06894 to 0.07004. This is confirmed in Table 4.19 for the NIG-NGARCH model associated with the EGP with an IVRMSE of 0.05935.

Working with non-Gaussian residuals is not the only way to generate more realistic VRP than Gaussian-GARCH ones. We present in table 4.15 the IVRMSE of different

Gaussian-GARCH models when an exponential-quadratic SDF is used to price options. It is worth noting that in this approach, it is not possible to directly estimate models from returns market quotes because the extra parameter π is involved in the risk-neutral dynamics. We obtain good IVRMSE between 0.06006 and 0.06331 that consistently outperform the Gaussian counterpart with exponential-affine SDF. Even if they are slightly worse than corresponding values for NIG-GARCH models for out-of-sample IVRMSE, the hierarchy is reversed when considering the next week pricing errors build on the rolling window estimation strategy. Both a modeling approach based on realistic conditional skewness and a modeling approach incorporating a variance premium in the pricing kernel seem to capture valuable empirical features. Therefore, a natural question is how is it possible that these two aspects are more complementary rather than competitive? The IG-GARCH model appears as an interesting candidate to tackle this issue.

For the IG-GARCH model, we obtain (see Table 4.21) out-of-sample IVRMSE between 0.067427 (in the case of the Esscher SDF estimated using returns only) and 0.056641 (for the U-shaped SDF and Returns-Options estimation strategy). Once again, a dataset of returns is not sufficient to produce competitive results. Furthermore, when the joint VIX-Returns estimation process is performed we obtain an IVRMSE of 0.057568 that is much closer to the best value at a considerably shorter computation time. The U-shaped pricing kernel of [Chorro & Fanirisoa (2016)] outperforms by around 7% the Esscher SDF in a conditionally Inverse-Gaussian environment and produces the best performances observed in this chapter: conditional skewness is a key factor of GARCH option pricing models that becomes outstanding when associated with a non-standard SDF.

When using GARCH option pricing models, the modeler is faced with four degrees of freedom: the GARCH structure, the distribution of the innovations, the pricing kernel, and the estimation strategy. Now we conclude the analysis of option pricing errors brought together in Table 4.25 with more general considerations on the impact of each factor caeteris paribus. Let us start with marginal effects: the impact of the choice of a non-affine GARCH structure accounting for the leverage effect is small with a 2.2% improvement in favor of the GJR model. In the same way, in the case of the NIG-NGARCH model estimated using returns and VIX information, the Esscher and the extended Girsanov principle SDF give rise to almost identical results with a difference of 1.4% for the benefit of the exponential-affine parameterization (see also [Badescu *et al.*(2011)] and [Badescu *et al.*(2015)] that deliver the same conclusion). Finally, using an estimation strategy based on options and returns information only improves by around 1% the IVRMSE with respect to its VIX-Returns counterpart (however, this improvement is around 10.5% when using returns only) as already observed in [Chorro & Fanirisoa (2016)]. Nevertheless, for this latter point we have to keep in mind that this slight 1% upgrade comes at a very high computational cost as reported in Tables 4.8, 4.12, 4.16 and 4.22. More decisively, the NIG distribution reduces the valuation error of around 11, 6% in comparison to Gaussian innovations while, in the affine family, the IG-GARCH model outperforms by 10.6% the Gaussian HN model. Concerning, the choice of the SDF, we clearly observe, both in the Gaussian and in the Inverse-Gaussian case that U-shaped parameterizations yield respectively to 13% and 7% lower IVRMSE (see [Christoffersen *et al.* (2013)], [Chorro & Fanirisoa (2016)], and [Badescu *et al.*(2017)] for similar findings). In the light of these observations, it is not surprising to see from Table 4.24 that, when we compare out-of-sample pricing errors between the best models of each sub-group, the most interesting performances are delivered by a model with non-Gaussian innovations, risk-neutralized using a U-shaped SDF and estimated maximizing the joint VIX-Returns log-likelihood, namely, the IG-GARCH model. We conclude that, when it is possible, the combination of all these factors is fundamental to producing competitive valuation errors.[31] The best model is not the most richly parameterized but a parsimonious one able to cope with classical stylized facts in terms of historical dynamics and risk representation.

Even if the ultimate criterion to compare GARCH option pricing models is the value of the pricing errors associated with a large real-world dataset of option prices, its computation may lead to large numerical issues in particular when Monte Carlo approximations are needed. This is true, not only during the estimation stage, but also to compute the objective function. To conclude this section, we question the possibility of deducing option pricing performances of a GARCH model from its capacity to forecast VIX dynamics. In Table 4.26, we have reported the ranks of the 24 models considered in this article regarding VIX and options adequacy measures introduced in Section 4.4.3. For example,[32] when we measure the relationship between rankings obtained from out-of-sample pricing errors and VIX RMSE we obtain a significant Spearman's rank correlation coefficient of 0.92. Moreover, top ten models obtained using VIX RMSE criterion are mainly as highly ranked as using options based criterion. The most important conclusion is that the ranking of models is well-preserved independently of the chosen option or VIX adequacy measure: examining the performance of a model in fitting VIX time series gives a very good indication on related pricing performances at a very reasonable computational cost. VIX analysis appears in this way as a very interesting and parsimonious first-stage evaluation to discard the worst GARCH option pricing models.

4.5 Conclusion

In this chapter, we have examined pricing performances of a large collection of GARCH models by questioning the global synergy between the choice of the affine/non-affine GARCH specification, the use of competing alternatives to the Gaussian distribution, the selection of an appropriate SDF and the choice of different estimation strategies based on several sets of financial information and on standard minimization algorithms. Therefore, 24 combinations of GARCH/distribution/SDF/estimation are tested using a large option

[31] It is also important to remark that noteworthy results are obtained with non-affine GARCH structures in NIG environment when VIX information is used in the estimation process. In this case, the residual error of around 3% comes from the necessity to use classical SDF to obtain risk-neutral dynamics.

[32] In Table 4.26, we notice that the rankings related to options (or VIX) valuation are essentially independent of the choice of the adequacy measures. For example, Spearman's rank correlation coefficient between in and out-of sample pricing errors ranking methodologies is equal to 0.97 . Consequently, we focus our attention on out-of-sample pricing performances and VIX RMSE.

dataset written on the S&P500. To do this, an intensive empirical comparison is performed not only based on in and out-of-sample pricing performances, but also using a weekly rolling window strategy where the model is estimated each Wednesday to price options one week later. Uniformly for these three criteria, the IG-GARCH model risk-neutralized using a U-shaped pricing kernel provides the best results. This gives evidence for the importance of using a non-Gaussian distribution combined with a non-standard stochastic discount factor that takes account for the variance risk premium. Of course, to estimate the variance risk aversion parameter, historical returns are not sufficient and an extra financial information is required. At this point, we have found that the joint VIX-Returns likelihood estimation provides competitive pricing errors at a very interesting computational cost with respect to option based estimation processes. This latter finding holds for all models considered in this chapter. For non-affine GARCH specifications, we found that, under NIG innovations, very interesting pricing errors are obtained when, and only when, VIX information is incorporated into the estimation strategy. This is efficiently possible for the NGARCH model using the EGP risk-neutralization process or using the two-step estimation strategy developed in [Chorro & Fanirisoa (2017)].

Finally, we have questioned in this study the possibility to deduce option pricing performances of a GARCH model from its capacity to forecast VIX dynamics. When we ranked models using options or VIX criteria we obtained a highly significant Spearman's rank correlation coefficient of 0.92. Therefore, examining the performance of a model in fitting VIX time series gives a very good indication on related pricing performances at a very reasonable computational cost. VIX analysis appears in this way as a very interesting and parsimonious first-stage evaluation to discard the worst GARCH option pricing models.

4.6 Tables and figures

Figure 4.1: S&P500 and VIX closing prices (top) and daily log-returns (bottom) from January 7, 1999 to December 22, 2010.

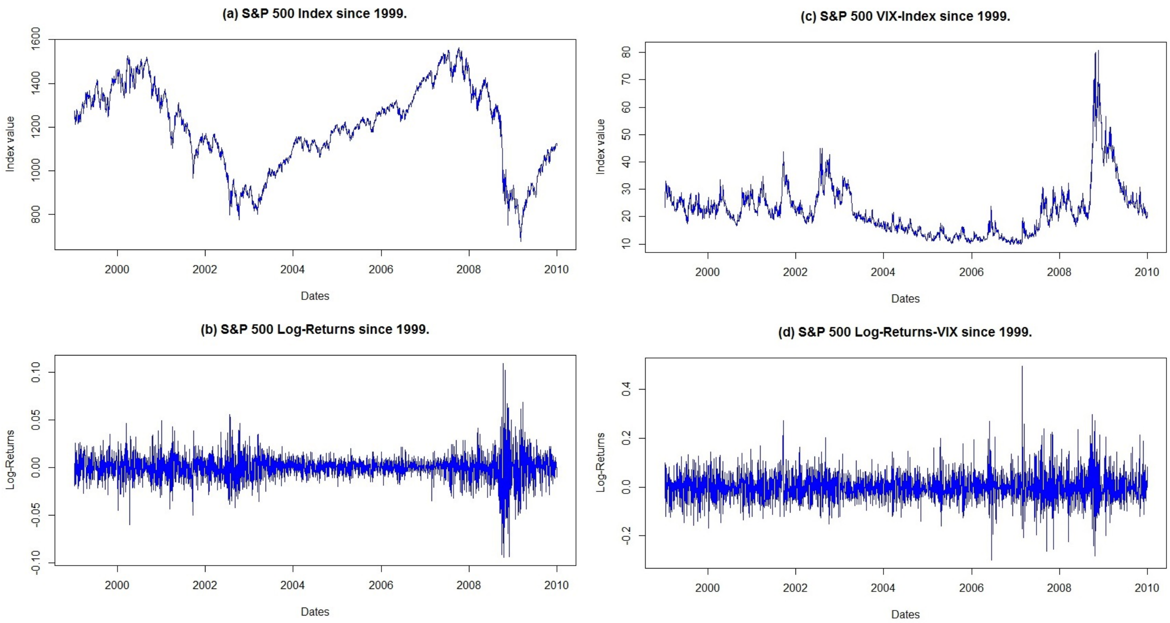


Table 4.1: Descriptive statistics of the S&P500 and VIX datasets covering the period January 7, 1999-December 22, 2010.

	Number of observations	Min	Max	Mean	Std Dev	Skewness	Kurtosis
Price index	2718	676.53	1565.15	1182.75	190.14	−0.0959	−0.6909
Log-returns	2718	−0.0947	0.1096	−0.0001	0.0139	−0.1214	7.3758
VIX index	2718	9.8900	80.8600	22.1859	9.6098	1.8853	5.6964
Log VIX	2718	−0.3506	0.4960	−0.0001	0.0613	0.5697	4.1682

Table 4.2: Properties of the in-sample (Dataset A) options data (2009-2010) and the out-of-sample (Dataset B) options data (2011-2012), the table shows the number of contracts, the average price, and the average implied volatility across moneynesses and times to maturities.

Option Dataset	Dataset A				Dataset B			
S/K \ T	< 60	$[60, 180]$	> 180	**All**	< 60	$[60, 180]$	> 180	**All**
Number of call option contracts :								
$0 < S/K < 0.975$	178	607	286	1071	107	419	214	740
$0.975 < S/K < 1.00$	40	103	44	187	36	80	46	162
$1.00 < S/K < 1.025$	36	96	54	186	30	75	41	146
$1.025 < S/K < 1.05$	35	93	37	165	31	75	37	143
$1.05 < S/K < 1.075$	37	93	40	170	28	72	29	129
$1.075 < S/K$	122	546	267	935	79	312	138	529
All	448	1538	728	2714	311	1033	505	1849
Average call price :								
$0 < S/K < 0.975$	8.558	23.392	41.658	24.536	7.436	21.804	42.351	23.863
$0.975 < S/K < 1.00$	28.133	59.176	84.700	57.336	25.047	59.893	84.423	56.454
$1.00 < S/K < 1.025$	42.764	71.741	96.643	70.383	45.442	76.560	103.004	75.002
$1.025 < S/K < 1.05$	59.721	87.681	109.272	85.558	66.109	95.260	119.433	93.600
$1.05 < S/K < 1.075$	77.534	103.012	125.367	101.971	88.434	116.030	139.506	114.656
$1.075 < S/K$	133.310	170.220	187.118	163.549	147.551	178.710	197.623	174.628
All	58.337	85.870	107.460	83.889	63.336	91.376	114.390	89.701
Average implied volatility from call options :								
$0 < S/K < 0.975$	0.212	0.209	0.210	0.210	0.161	0.174	0.182	0.172
$0.975 < S/K < 1.00$	0.223	0.231	0.233	0.229	0.177	0.198	0.205	0.194
$1.00 < S/K < 1.025$	0.228	0.230	0.235	0.231	0.202	0.207	0.211	0.207
$1.025 < S/K < 1.05$	0.239	0.240	0.233	0.237	0.202	0.210	0.213	0.208
$1.05 < S/K < 1.075$	0.259	0.245	0.235	0.246	0.226	0.222	0.211	0.220
$1.075 < S/K$	0.308	0.267	0.255	0.277	0.260	0.235	0.228	0.241
All	0.245	0.237	0.234	0.238	0.204	0.207	0.208	0.207

Table 4.3: Summary information on risk-neutral dynamics of each Gaussian GARCH model analyzed in this chapter.

Model	Risk-neutral dynamics	Risk-neutral parameters and distribution	References
HN-GARCH :			
Gaus-Ess	$Y_t = r - \frac{h_t}{2} + \sqrt{h_t} z_t^*$	$z_t^* \hookrightarrow N(0,1)$	[Heston & Nandi (2000)]
	$h_t = a_0 + a_1 \left(z_{t-1}^* - \gamma^* \sqrt{h_{t-1}}\right)^2 + b_1 h_{t-1}$	$\gamma^* = \gamma + \lambda_0 + \frac{1}{2}$	[Kanniainen *et al.* (2014)]
Gaus-Qua	$Y_t = r - \frac{h_t^*}{2} + \sqrt{h_t^*} z_t^*$	$z_t^* \hookrightarrow N(0,1)$, $\pi = \frac{h_t^*}{h_t}$	[Monfort & Pégoraro (2012)]
	$h_t^* = \pi a_0 + \pi^2 a_1 \left(z_{t-1}^* - \gamma^* \sqrt{h_{t-1}^*}\right)^2 + b_1 h_{t-1}^*$	$\gamma^* = \left(\frac{\gamma}{\pi} + \frac{\lambda_0}{\pi} + \frac{1}{2}\right)$	[Chorro *et al.* (2015)] (Chap. 3.5)
GJR-GARCH :			
Gaus-Ess	$Y_t = r - \frac{h_t}{2} + \sqrt{h_t} z_t^*$		
	$h_t = a_0 + h_{t-1} \left[b_1 + a_1 \left(z_{t-1}^* - \lambda_0\right)^2\right]$	$z_t^* \hookrightarrow N(0,1)$	[Kanniainen *et al.* (2014)]
	$+\gamma h_{t-1} \max\left(0; -\left(z_{t-1}^* - \lambda_0\right)\right)^2$		
Gaus-Qua	$Y_t = r - \frac{h_t^*}{2} + \sqrt{h_t^*} z_t^*$	$z_t^* \hookrightarrow N(0,1)$, $\pi = \frac{h_t^*}{h_t}$	[Monfort & Pégoraro (2012)]
	$h_t^* = a_0^* + h_{t-1}^* b_1 + a_1^* h_{t-1}^* \left(z_{t-1}^* - \lambda_t^*\right)^2$	$a_0^* = a_0 \pi$, $a_1^* = a_1 \pi$, $\gamma^* = \gamma \pi$	[Chorro *et al.* (2015)] (Chap. 3.5)
	$+\gamma^* h_{t-1}^* \max\left(0; -\left(z_{t-1}^* - \lambda_t^*\right)\right)^2$	$\lambda_t^* = \frac{\lambda_0}{\sqrt{\pi}} - \frac{\sqrt{h_{t-1}^*}}{2}\left(\frac{1}{\pi} - 1\right)$	
NGARCH :			
Gaus-Ess	$Y_t = r - \frac{h_t}{2} + \sqrt{h_t} z_t^*$	$z_t^* \hookrightarrow N(0,1)$	
	$h_t = a_0 + b_1 h_{t-1} + a_1 h_{t-1} \left(z_{t-1}^* - \gamma^*\right)^2$	$\gamma^* = \lambda_0 + \gamma$	[Kanniainen *et al.* (2014)]
Gaus-Qua	$Y_t = r - \frac{h_t^*}{2} + \sqrt{h_t^*} z_t^*$	$z_t^* \hookrightarrow N(0,1)$,	[Monfort & Pégoraro (2012)]
	$h_t^* = a_0^* + b_1 h_{t-1}^* + a_1^* h_{t-1}^* \left(z_{t-1}^* - \gamma_t^*\right)^2$	$\pi = \frac{h_t^*}{h_t}$, $a_0^* = a_0 \pi$, $a_1^* = a_1 \pi$	[Chorro *et al.* (2015)]
		$\gamma_t^* = \frac{\gamma}{\sqrt{\pi}} + \frac{\lambda_0}{\sqrt{\pi}} - \frac{\sqrt{h_{t-1}^*}}{2}\left(\frac{1}{\pi} - 1\right)$	

Table 4.4: Summary information on risk-neutral dynamics of each non-Gaussian GARCH model analyzed in this chapter.

Model	Risk-neutral dynamics	Risk-neutral parameters[33] and distribution	References
HN-GARCH :			
NIG-Ess	$Y_t = r + m_t + \sqrt{h_t} z_t^*$ $h_t = a_0 + a_1 \left(z_{t-1}^* - \gamma\sqrt{h_{t-1}}\right)^2 + b_1 h_{t-1}$ $m_t = \lambda_0\sqrt{h_t} - \frac{1}{2}h_t$	$z_t^* \hookrightarrow NIG\left(\tilde{\alpha}, \tilde{\beta} + \sqrt{h_t}\theta_t^q, \tilde{\delta}, \tilde{\mu}\right)$ $\theta_t^q = -\frac{1}{2} - \frac{\tilde{\alpha}\tilde{\beta}\sqrt{\tilde{\delta}}}{\sqrt{h_t}\varrho_t^{\frac{3}{2}}} - \frac{1}{2}\Theta_t,$ $\varrho = \sqrt{\tilde{\alpha}^2 - \tilde{\beta}^2}$	[Badescu *et al.*(2011)] [Chorro *et al.* (2012)]
GJR-GARCH :			
NIG-Ess	$Y_t = r + m_t + \sqrt{h_t} z_t^*$ $h_t = a_0 + h_{t-1}\left[b_1 + a_1\left(z_{t-1}^* - \lambda_0\right)^2\right]$ $+\gamma h_{t-1} \max\left(0; -\left(z_{t-1}^* - \lambda_0\right)\right)^2$ $m_t = \lambda_0\sqrt{h_t} - \frac{1}{2}h_t$	$z_t^* \hookrightarrow NIG\left(\tilde{\alpha}, \tilde{\beta}\sqrt{h_t} + \theta_t^q, \tilde{\delta}, \tilde{\mu}\right)$ $\theta_t^q = -\frac{1}{2} - \frac{\tilde{\alpha}\tilde{\beta}\sqrt{\tilde{\delta}}}{\sqrt{h_t}\varrho_t^{\frac{3}{2}}} - \frac{1}{2}\Theta_t$ $\varrho = \sqrt{\tilde{\alpha}^2 - \tilde{\beta}^2}$	[Badescu *et al.*(2011)] [Chorro *et al.* (2012)]
NGARCH :			
NIG-Ess	$Y_t = r + m_t + \sqrt{h_t} z_t^*$ $h_t = a_0 + b_1 h_{t-1} + a_1 h_{t-1}\left(z_{t-1}^* - \gamma\right)^2$ $m_t = \lambda_0\sqrt{h_t} - \frac{1}{2}h_t$	$z_t^* \hookrightarrow NIG\left(\tilde{\alpha}, \tilde{\beta}\sqrt{h_t} + \theta_t^q, \tilde{\delta}, \tilde{\mu}\right)$ $\theta_t^q = -\frac{1}{2} - \frac{\tilde{\alpha}\tilde{\beta}\sqrt{\tilde{\delta}}}{\sqrt{h_t}\varrho_t^{\frac{3}{2}}} - \frac{1}{2}\Theta_t$ $\varrho = \sqrt{\tilde{\alpha}^2 - \tilde{\beta}^2}$	[Badescu *et al.*(2011)] [Chorro *et al.* (2012)]
NIG-EGP	$Y_t = r + m_t - \nu_t + \sqrt{h_t} z_t^*,$ $h_t = a_0 + b_1 h_{t-1} + a_1 h_{t-1}\left(z_{t-1}^* - \frac{\nu_{t-1}}{\sqrt{h_{t-1}}} - \gamma\right)^2$	$z_t^* \hookrightarrow NIG\left(\tilde{\alpha}, \tilde{\beta}, \tilde{\delta}, \tilde{\mu}\right)$ $e^{\nu_t} = e^{-r}\mathbb{E}_{\mathbb{P}}\left[e^{Y_t} \mid \mathcal{F}_{t-1}\right]$ $\nu_t = \tilde{\mu}\sqrt{h_t} + \sqrt{\tilde{\alpha}^2 - \tilde{\beta}^2} - \sqrt{\tilde{\alpha}^2 - (\tilde{\beta} + \tilde{\delta}\sqrt{h_t})^2}$	[Badescu *et al.*(2015)]
IG-GARCH :			
IG-Ess	$Y_t = r + \nu^* h_{t+1}^* + \eta^* z_t^*$ $h_t^* = w^* + b^* h_{t-1}^* + c^* z_{t-1}^* + a^* \frac{(h_{t-1}^*)^2}{z_{t-1}^*}$	$z_t^* \hookrightarrow IG\left(\frac{h_t}{\eta^2}\right)$ $\eta^* = \frac{\eta}{1 - 2\theta^*\eta},\ \theta^* = \frac{1}{2}\left[\eta^{-1} - \frac{1}{\nu^2\eta^3}\left[1 + \frac{\nu^2\eta^3}{2}\right]^2\right]$ $c^* = c\left(\frac{\eta^*}{\eta}\right)^{\frac{5}{2}},\ a^* = a\left(\frac{\eta^*}{\eta}\right)^{-\frac{5}{2}},\ \nu^* = \nu\left(\frac{\eta^*}{\eta}\right)^{-\frac{3}{2}},\ w^* = w\left(\frac{\eta^*}{\eta}\right)^{\frac{3}{2}},$	[Christoffersen *et al.* (2006)]
IG-Ushp	$Y_t = r + \nu^* h_{t+1}^* + \eta^* z_t^*$ $h_t^* = w^* + b h_t^* + c^* y_t^* + a^* \frac{(h_t^*)^2}{z_t^*}$	$z_t^* \hookrightarrow IG\left(\frac{h_t}{\eta^2}\right),\ \pi = \frac{h_t^*}{h_t}$ $\eta^* = \sqrt[3]{\frac{\pi^2}{\nu^2}\left(-1 + \sqrt{1 + \frac{8\nu}{27\pi}}\right)} + \sqrt[3]{\frac{\pi^2}{\nu^2}\left(-1 - \sqrt{1 + \frac{8\nu}{27\pi}}\right)}$ $\nu^* = \frac{\nu}{\pi},\ w^* = w\pi,\ c^* = \frac{c\pi\eta^*}{\eta},\ a^* = \frac{a\eta}{\pi\eta^*}$	[Chorro & Fanirisoa (2016)]

$$\Theta_t = \sqrt{\frac{\left(\tilde{\alpha}m_t + \sqrt{\tilde{\delta}h_t}\tilde{\beta}\varrho_t\right)^2}{h_t\tilde{\delta}\varrho_t^3}\left(\frac{4\tilde{\alpha}^4\tilde{\delta}^2}{h_t\tilde{\delta}\varrho_t^3 + \left(\tilde{\alpha}m_t + \sqrt{\tilde{\delta}h_t}\tilde{\beta}\varrho_t\right)^2} - 1\right)}$$

Table 4.5: We present expressions of the parameters $\tilde{h}_0$ and Ψ^* associated with the closed-form expression of Vix_t in equation 4.14 for different GARCH structures, SDF and conditional distributions.

GARCH models	$\tilde{h}_0$	Ψ^*
HN-GARCH :		
Gaussian-Ess	$\frac{a_0 + a_1}{1 - \Psi^*}$	$b_1 + a_1(\gamma + \lambda_0 + \frac{1}{2})^2$
Gaussian-Qua	$\frac{a_0 + \pi a_1}{1 - \Psi^*}$	$b_1 + \pi^2 a_1 \left(\frac{\gamma}{\pi} + \frac{\lambda_0}{\pi} + \frac{1}{2}\right)^2$
GJR-GARCH :		
Gaussian-Ess	$\frac{a_0}{1 - \Psi^*}$	$b_1 + [a_1 + \gamma N(\lambda_0)]\left(1 + \lambda_0^2\right) + \gamma \lambda_0 n(\lambda_0)$
NGARCH :		
Gaussian-Ess	$\frac{a_0}{1 - \Psi^*}$	$b_1 + a_1(1 + (\lambda_0 + \gamma)^2)$
NIG-EGP	$\frac{a_0}{1 - \Psi^*}$	$b_1 + a_1(1 + (\lambda_0 + \gamma)^2)$
IG-GARCH :		
Ess	$\frac{w + a(\eta)^4}{1 - \Psi^*}$	$b + \frac{c^*}{(\eta^*)^2} + a^*(\eta^*)^2$
Ushp	$\frac{w + \frac{a\eta}{\pi^2}(\eta^*)^3}{(1 - \psi^*)}$	$b + \frac{c^*}{(\eta^*)^2} + a^*(\eta^*)^2$

Table 4.6: Parameter estimates and standard errors of Gaussian GARCH models combined with the Esscher SDF. Returns means MLE estimation procedure using only returns information, Ret-VIX means Joint MLE estimation using returns and options information, Ret-VIX means Joint MLE estimation using returns and VIX information. The estimation is based on log-returns and VIX datasets from January 7 1999 to December 22 2010 and on the in-sample dataset of options (2009-2010).

GARCH-type Information	HN-GARCH Returns	GJR-GARCH Returns	NGARCH Returns	HN-GARCH Opt-Ret	HN-GARCH Ret-VIX	GJR-GARCH Ret-VIX	NGARCH Ret-VIX
a_0	$3.854E-08$	$3.049E-06$	$1.677E-06$	$1.859E-07$	$3.757E-12$	$4.966E-06$	$3.557E-06$
Stand.Dev	(0.0044)	(0.0011)	(0.0000)	(0.0009)	(0.0007)	(0.0004)	(0.0009)
a_1	$2.254E-05$	$1.243E-01$	$6.174E-02$	$1.542E-06$	$2.252E-05$	$1.240E-01$	$6.172E-02$
Stand.Dev	(0.0001)	(0.0003)	(0.0014)	(0.0003)	(0.0002)	(0.0000)	(0.0007)
b_1	$8.272E-01$	$8.509E-01$	$8.446E-01$	$6.500E-01$	$9.117E-01$	$8.504E-01$	$7.956E-01$
Stand.Dev	(0.0035)	(0.0006)	(0.0011)	(0.0030)	(0.0086)	(0.0015)	(0.0009)
γ	$5.379E+01$	$2.208E-02$	$1.174E+00$	$4.586E+02$	$1.423E+01$	$2.314E-02$	$4.701E-08$
Stand.Dev	(0.0011)	(0.0025)	(0.0001)	(0.0095)	(0.0088)	(0.0005)	(0.0206)
λ_0	$1.020E+00$	$2.288E-01$	$8.911E-07$	$8.596E+00$	$1.513E+00$	$1.989E-01$	$8.452E-01$
Stand.Dev	(0.0000)	(0.0055)	(0.0010)	(0.0008)	(0.0501)	(0.0033)	(0.0050)
ϱ	–	–	–	–	0.9992	0.8924	0.9542
Stand.Dev	–	–	–	–	(0.0106)	(0.0012)	(0.0110)

Table 4.7: Option pricing performances and VIX predictability (see section 4.4.3) of Gaussian-GARCH models combined with the Esscher SDF. The results are based on estimates provided in Table 4.6. The column **Model Properties** presents computational time of estimation in hours and Variance Risk Premium.

GARCH-type Information	HN Returns	GJR Returns	NGARCH Returns	HN Opt-Ret	HN Ret-VIX	GJR Ret-VIX	NGARCH Ret-VIX
Model Properties							
Times (h)	0.010	0.018	0.025	9.014	0.008	0.021	0.014
$-VRP$ (in %)	$3.27E-10$	$2.86E-16$	$2.26E-12$	$8.88E-11$	$9.67E-09$	$7.122E-13$	$1.326E-16$
Predictability of VIX							
MPE_{VIX}	0.0113	0.0106	−0.0088	−0.0092	−0.0023	−0.0016	−0.0012
MAE_{VIX}	0.0128	0.0125	0.0106	0.0104	0.0058	0.0059	0.0057
$RMSE_{VIX}$	0.2849	0.2771	0.2475	0.1869	0.1844	0.1740	0.1748
Pricing performances							
in-$IVRMSE$	0.0599	0.0574	0.0571	0.0557	0.0580	0.0548	0.0557
out-$IVRMSE$	0.0777	0.0764	0.0766	0.0733	0.0735	0.0650	0.0729
We-$IVRMSE$	0.0662	0.0651	0.0652	0.0610	0.0614	0.0592	0.0592

Table 4.8: Model comparisons based on computational time of estimation and on in-sample IVRMSE given in Table 4.7. The upper triangular part of the matrix illustrates relative difference (in percentage) of the computational time of estimation between the i-th and the j-th models, as example: $-80\% = 100 * (0.010 - 0.018)/0.010$. The lower triangular part of the matrix illustrates relative difference (in percentage) of the in-sample IVRMSE between the j-th and the i-th models, as example: $4.072\% = 100 * (0.059916 - 0.057476)/0.059916$.

GARCH-type	HN Ret	GJR Ret	NGARCH Ret	HN Opt-Ret	HN VIX-Ret	GJR VIX-Ret	NGARCH VIX-Ret
HN-Ret	–	−80.00	−150.00	$-9.004E+04$	20.000	−110.00	−40.00
GJR-Ret	4.072	–	−38.888	$-4.997E+04$	55.555	−16.666	22.222
NGARCH-Ret	4.561	0.509	–	$-3.595E+04$	68.000	16.000	44.000
HN-Opt-Ret	6.960	3.010	2.513	–	99.911	99.767	99.844
HN-VIX-Ret	3.179	−0.930	−1.448	−4.063	–	−162.50	−75.00
GJR-VIX-Ret	8.480	4.595	4.106	1.634	5.475	–	33.333
NGARCH-VIX-Ret	7.035	3.088	2.592	0.080	3.982	−1.579	–

Table 4.9: Model comparisons based on out-of-sample IVRMSE and Wednesday-IVRMSE given in Table 4.7. The upper triangular part of the matrix illustrates relative difference (in percentage) of the out-of-sample IVRMSE between the i-th and the j-th models, as example: $1.5650\% = 100 * (0.077701 - 0.076485)/0.077701$. The lower triangular part of the matrix illustrates relative difference (in percentage) of the Wednesday-IVRMSE between the j-th and the i-th models, as example: $1.736\% = 100 * (0.06626 - 0.06511)/0.06626$.

GARCH-type	HN Ret	GJR Ret	NGARCH Ret	HN Opt-Ret	HN VIX-Ret	GJR VIX-Ret	NGARCH VIX-Ret
HN-Ret	–	1.565	1.393	5.540	5.394	16.340	6.054
GJR-Ret	1.736	–	−0.174	4.045	3.890	15.010	4.560
NGARCH-Ret	1.509	−0.230	–	4.212	4.057	15.160	4.727
HN-Opt-Ret	7.923	6.297	6.512	–	−0.162	11.430	0.536
HN-VIX-Ret	7.259	5.621	5.838	−0.7212	–	11.570	0.697
GJR-VIX-Ret	10.640	9.062	9.271	2.950	3.645	–	−12.30
NGARCH-VIX-Ret	10.530	8.954	9.163	2.836	3.531	−0.118	–

Table 4.10: Parameter estimates and standard errors of the NIG distribution for GARCH models combined with the Esscher SDF. These parameters have been obtained using the standard maximum-likelihood algorithm for the residuals extracted from Table 4.6.

GARCH-type Information	HN Returns	GJR Returns	NGARCH Returns	HN Opt-Ret	HN Ret-VIX	GJR Ret-VIX	NGARCH Ret-VIX
α	1.2501	1.1550	1.2702	1.4630	1.4365	1.3589	1.4536
Stand.Dev	(0.0004)	(0.0108)	(0.0036)	(0.0005)	(0.0008)	(0.0001)	(0.0009)
β	−0.0106	−0.1432	−0.0025	−0.0061	−0.0538	−0.0058	−0.0061
Stand.Dev	(0.0008)	(0.0057)	(0.0015)	(0.0008)	(0.0003)	(0.0023)	(0.0001)
δ	1.4728	1.0623	1.6204	1.4454	1.3920	1.5336	1.4538
Stand.Dev	(0.0095)	(0.0000)	(0.0005)	(0.0001)	(0.0008)	(0.0000)	(0.0000)
μ	2.7086	0.1327	1.9734	2.1602	11.6243	7.9908	2.0178
Stand.Dev	(0.0051)	(0.0076)	(0.0055)	(0.0000)	(0.0013)	(0.0000)	(0.0003)

Table 4.11: Option pricing performances and VIX predictability (see section 4.4.3) of NIG-GARCH models combined with the Esscher SDF. The results are based on estimates provided in Tables 4.6 and 4.10. The column **Model Properties** presents computational time of estimation in hours and Variance Risk Premium. .

GARCH-type Information	HN Returns	GJR Returns	NGARCH Returns	HN Opt-Ret	HN Ret-VIX	GJR Ret-VIX	NGARCH Ret-VIX
Model Properties							
Times (h)	0.016	0.024	0.031	9.071	0.017	0.036	0.019
$-VRP$ (in %)	2.906	2.867	2.956	3.011	3.213	3.006	3.366
Predictability of VIX							
MPE_{VIX}	−0.0054	−0.0043	−0.0029	−0.0010	−0.0006	0.0011	0.0010
MAE_{VIX}	0.0074	0.0062	0.0053	0.0054	0.0051	0.0047	0.0049
$RMSE_{VIX}$	0.1884	0.1758	0.1703	0.1509	0.1314	0.1130	0.1308
Pricing performances							
in-$IVRMSE$	0.0573	0.0550	0.0567	0.0519	0.0521	0.0512	0.0463
out-$IVRMSE$	0.0700	0.0689	0.0690	0.0639	0.0648	0.0595	0.0592
We-$IVRMSE$	0.0595	0.0592	0.0586	0.0510	0.0514	0.0504	0.0504

Table 4.12: Model comparisons based on computational time of estimation and on in-sample IVRMSE given in Table 4.11. The upper triangular part of the matrix illustrates relative difference (in percentage) of the computational time of estimation between the i-th and the j-th models, as example: $-50\% = 100 * (0.016 - 0.024)/0.016$. The lower triangular part of the matrix illustrates relative difference (in percentage) of the in-sample IVRMSE between the j-th and the i-th models, as example: $4.116\% = 100 * (0.057391 - 0.055029)/0.057391$.

GARCH-type	HN Ret	GJR Ret	NGARCH Ret	HN Opt-Ret	HN-VIX-Ret VIX-Ret	GJR VIX-Ret	NGARCH VIX-Ret
HN-Ret	–	−50.00	−93.75	$-5.659E+04$	−6.250	−125.00	−18.75
GJR-Ret	4.116	–	−29.16	$-3.769E+04$	29.166	−50.00	20.833
NGARCH-Ret	1.203	−3.038	–	$-2.916E+04$	45.161	−16.12	38.709
HN-Opt-Ret	9.404	5.515	8.301	–	99.812	99.603	99.790
HN-VIX-Ret	9.094	5.192	7.987	−0.342	–	−111.76	−11.76
GJR-VIX-Ret	10.70	6.871	9.617	1.435	1.771	–	47.222
NGARCH-VIX-Ret	19.26	15.79	18.27	10.088	11.180	9.579	–

Table 4.13: Model comparisons based on out-of-sample IVRMSE and Wednesday-IVRMSE given in Table 4.11. The upper triangular part of the matrix illustrates relative difference (in percentage) of the out-of-sample IVRMSE between the i-th and the j-th models, as example: $1.565\% = 100 * (0.070042 - 0.068946)/0.070042$. The lower triangular part of the matrix illustrates relative difference (in percentage) of the Wednesday-IVRMSE between the j-th and the i-th models, as example: $0.420\% = 100 * (0.05950 - 0.05925)/0.05950$.

GARCH-type	HN Ret	GJR Ret	NGARCH Ret	HN Ret	HN VIX-Ret	GJR VIX-Ret	NGARCH VIX-Ret
HN-Ret	–	1.565	1.483	8.669	7.360	14.960	15.34
GJR-Ret	0.420	–	−0.088	7.217	5.887	13.610	14.00
NGARCH-Ret	1.361	0.945	–	7.291	5.972	13.680	14.07
HN-Opt-Ret	14.240	13.870	13.050	–	−1.433	6.888	7.30
HN-VIX-Ret	13.560	13.200	12.370	−0.783	–	8.203	8.61
GJR-VIX-Ret	15.260	14.900	14.090	1.195	1.964	–	0.45
NGARCH-VIX-Ret	15.230	14.870	14.060	1.156	1.925	−0.039	–

Table 4.14: Parameter estimates and standard errors of Gaussian GARCH models combined with the exponential-quadratic SDF. Returns means MLE estimation procedure using only returns information, Ret-VIX means Joint MLE estimation using returns and options information, Ret-VIX means Joint MLE estimation using returns and VIX information. The estimation is based on log-returns and VIX datasets from January 7 1999 to December 22 2010 and on the in-sample dataset of options (2009-2010).

GARCH-type Information	HN-GARCH Opt-Ret	HN-GARCH Ret-VIX	GJR-GARCH Ret-VIX	NGARCH Ret-VIX
a_0	$5.7547E-14$	$1.0014E-12$	$4.966E-06$	$1.780E-06$
Stand.Dev	(0.0009)	(0.0003)	(0.0008)	(0.0000)
a_1	$1.5139E-06$	$1.5048E-06$	$1.241E-01$	$3.877E-02$
Stand.Dev	(0.0368)	(0.0002)	(0.0003)	(0.0065)
b_1	$6.500E-01$	$6.5121E-01$	$8.504E-01$	$9.329E-01$
Stand.Dev	(0.0032)	(0.0066)	(0.0004)	(0.0000)
γ	$4.5869E+02$	$4.586E+02$	$2.3142E-02$	$1.277E-07$
Stand.Dev	(0.0036)	(0.0095)	(0.0002)	(0.0078)
λ_0	$8.596E+00$	$8.672E+00$	$1.989E-01$	$4.583E-01$
Stand.Dev	(0.0006)	(0.0036)	(0.0015)	(0.0004)
π	$1.6723E+00$	$1.722E+00$	$1.2785E+00$	$1.2413E+00$
Stand.Dev	(0.0048)	(0.0022)	(0.0012)	(0.0092)
ϱ	–	0.8099	0.9546	0.9170
Stand.Dev	–	(0.0003)	(0.06235)	(0.0023)

Table 4.15: Option pricing performances and VIX predictability (see section 4.4.3) of Gaussian-GARCH models combined with the exponential-quadratic SDF. The results are based on estimates provided in Table 4.14. The column **Model Properties** presents computational time of estimation in hours and Variance Risk Premium.

GARCH-type Information	HN-GARCH Opt-Ret	HN-GARCH Ret-VIX	GJR-GARCH Ret-VIX	NGARCH Ret-VIX
Model Properties				
Times (h)	10.326	0.019	1.053	0.961
$-VRP$ (in %)	3.2301	3.2475	3.3562	3.5628
Predictibility of VIX				
MPE_{VIX}	−0.0012	−0.0004	−0.0003	−0.0004
MAE_{VIX}	0.0049	0.0047	0.0041	0.0041
$RMSE_{VIX}$	0.1445	0.1248	0.1119	0.1103
Pricing performances				
in-$IVRMSE$	0.0511	0.0513	0.0507	0.0492
out-$IVRMSE$	0.0627	0.0633	0.0628	0.0600
We-$IVRMSE$	0.0514	0.0515	0.0508	0.0493

Table 4.16: Model comparisons based on computational time of estimation and on in-sample IVRMSE given in Table 4.15. The upper triangular part of the matrix illustrates relative difference (in percentage) of the computational time of estimation between the i-th and the j-th models, as example: $99.816\% = 100 * (10.326 - 0.019)/10.326$. The lower triangular part of the matrix illustrates relative difference (in percentage) of the in-sample IVRMSE between the j-th and the i-th models, as example: $-0.528\% = 100 * (0.05110 - 0.05137)/0.05110$.

GARCH-type	HN-Opt-Ret	HN-VIX-Ret	GJR-VIX-Ret	NGARCH-VIX-Ret
HN-Opt-Ret	–	99.816	89.802	90.693
HN-VIX-Ret	−0.528	–	−5442	−4957
GJR-VIX-Ret	0.626	1.149	–	8.736
NGARCH-VIX-Ret	3.640	4.146	3.033	–

Table 4.17: Model comparisons based on out-of-sample IVRMSE and Wednesday-IVRMSE given in Table 4.15. The upper triangular part of the matrix illustrates relative difference (in percentage) of the out-of-sample IVRMSE between the i-th and the j-th models, as example: $-0.892\% = 100 * (0.06275 - 0.06331)/0.06275$. The lower triangular part of the matrix illustrates relative difference (in percentage) of the Wednesday-IVRMSE between the j-th and the i-th models, as example: $-0.097\% = 100 * (0.05147 - 0.05152)/0.05147$.

GARCH-type	HN-Opt-Ret	HN-VIX-Ret	GJR-VIX-Ret	NGARCH-VIX-Ret
HN-Opt-Ret	–	−0.892	−0.223	4.287
HN-VIX-Ret	−0.097	–	0.663	5.133
GJR-VIX-Ret	1.243	1.339	–	4.500
NGARCH-VIX-Ret	4.216	4.309	3.010	–

Table 4.18: Parameter estimates and standard errors of the NIG-NGARCH model combined with the EGP SDF. The estimation is based on Joint MLE estimation using returns and VIX information from January 7 1999 to December 22 2010.

Vol Parameters	a_0	a_1	b_1	γ	λ_0	ϱ
Values	$1.896E-06$	$3.877E-02$	$9.329E-01$	$7.110E-01$	$9.937E-02$	0.9163
Stand.Dev	(0.0035)	(0.0008)	(0.0000)	(0.0001)	(0.0000)	(0.0004)
NIG Parameters	α	β	δ	μ	–	–
Values	$2.961E+00$	$-9.441E-01$	1.5877	0.5341	–	–
Stand.Dev	(0.0002)	(0.0013)	(0.0000)	(0.0019)	–	–

Table 4.19: Option pricing performances and VIX predictability (see section 4.4.3) of Gaussian-GARCH models combined with exponential-quadratic SDF. The results are based on estimates provided in Table 4.18.

VIX Performances	$-VRP$ (*in* %)	MPE_{VIX}	MAE_{VIX}	$RMSE_{VIX}$
Values	3.3645	−0.0001	0.0039	0.1066
Pricing performances	Times (h)	in-IVRMSE	out-IVRMSE	we-IVRMSE
Values	0.0101	0.0480	0.0593	0.0488

Table 4.20: Parameter estimates and standard errors of the IG-GARCH model combined with Esscher and U-shaped SDF. Returns means MLE estimation procedure using only returns information, Ret-VIX means Joint MLE estimation using returns and options information, Ret-VIX means Joint MLE estimation using returns and VIX information. The estimation is based on log-returns and VIX datasets from January 7 1999 to December 22 2010 and on the in-sample dataset of options (2009-2010).

Joint-Estimation	Returns	Returns-Option		Returns-VIX	
Model		M_t^{ess}	M_t^{Ushp}	M_t^{ess}	M_t^{Ushp}
Parameters:					
w	$1.2061E-06$	$9.7699E-06$	$1.0185E-05$	$2.2341E-06$	$5.3156E-05$
Stand.Dev	(0.0000)	(0.0006)	(0.0002)	(0.0002)	(0.0004)
b	$2.3052E-03$	$1.0159E-03$	$1.7211E-03$	$2.3184E-03$	$1.8603E-03$
Stand.Dev	(0.0000)	(0.0000)	(0.0001)	(0.0003)	(0.0003)
c	$4.9024E-05$	$4.5379E-05$	$4.5118E-05$	$4.8949E-05$	$4.8233E-05$
Stand.Dev	(0.0000)	(0.0000)	(0.0001)	(0.0009)	(0.0005)
a	$3.3174E+03$	$3.3317E+03$	$3.3174E+02$	$3.3174E+03$	$3.3174E+03$
Stand.Dev	(0.0000)	(0.0001)	(0.0007)	(0.0002)	(0.0005)
η	$-7.972E-03$	$-7.5314E-03$	$-7.4936E-03$	$-7.9552E-03$	$-7.928E-03$
Stand.Dev	(0.0000)	(0.0000)	(0.0000)	(0.0002)	(0.0010)
ν	$1.2584E+02$	$1.2594E+02$	$1.2573E+02$	$1.2583E+02$	$1.2584E+02$
Stand.Dev	(0.0009)	(0.0001)	(0.0003)	(0.0003)	(0.0001)
π	–	–	1.4005	–	1.6325
Stand.Dev	–	–	(0.0084)	–	(0.0035)
ϱ	–	–	–	$9.9552E-01$	$9.9386E-01$
Stand.Dev	–	–	–	(0.0070)	(0.0034)

Table 4.21: Option pricing performances and VIX predictability (see section 4.4.3) of the IG-GARCH model combined with Esscher and U-shaped SDF. The results are based on estimates provided in Table 4.20. The column **Model Properties** presents the computational time of estimation in hours and the Variance Risk Premium.

Joint-Estimation	Returns	Returns-Option		Returns-VIX	
Model		M_t^{ess}	M_t^{Ushp}	M_t^{ess}	M_t^{Ushp}
Model Properties:					
Times (h)	0.036	9.2684	10.3984	0.0152	0.0348
VPR	3.1785	3.7541	3.5165	3.7042	3.4563
Predictability of VIX:					
MPE_{VIX}	−0.0011	−0.0004	−0.0003	−0.00008	−0.00019
MAE_{VIX}	0.0051	0.0043	0.0040	0.0039	0.0039
$RMSE_{VIX}$	0.1364	0.1315	0.1061	0.1010	0.0990
Pricing performances:					
in-IVRMSE	0.0543	0.0461	0.0435	0.0464	0.0438
out-IVRMSE	0.0674	0.0610	0.0566	0.0618	0.0575
we-IVRMSE	0.0514	0.0500	0.0480	0.0510	0.0480

Table 4.22: Model comparisons based on computational time of estimation and on in-sample IVRMSE given in Table 4.21. The upper triangular part of the matrix illustrates relative difference (in percentage) of the computational time of estimation between the i-th and the j-th models, as example: $-2.564\ ^{+05}\% = 100*(0.036-9.268)/0.036$. The lower triangular part of the matrix illustrates relative difference (in percentage) of the in-sample IVRMSE between the j-th and the i-th models, as example: $15.08\% = 100*(0.054358-0.046160)/0.054358$.

IG-GARCH-type	Ess-Ret	Ess-Opt-Ret	Ushp-Opt-Ret	Ess-VIX-Ret	Ushp-VIX-Ret
Ess-Ret	–	$-2.5E+05$	$-2.8E+05$	57.777	3.333
Ess-Opt-Ret	15.08	–	−12.191	99.836	99.624
Ushp-Opt-Ret	19.89	5.663	–	99.853	99.665
Ess-VIX-Ret	14.49	−0.699	−6.745	–	−128.94
Ushp-VIX-Ret	19.29	4.957	−0.748	5.617	–

Table 4.23: Model comparisons based on the out-of-sample IVRMSE and the Wednesday-IVRMSE given in Table 4.21. The upper triangular part of the matrix illustrates relative difference (in percentage) of the out-of-sample IVRMSE between the i-th and the j-th models, as example: $9.446\% = 100 * (0.067427 - 0.061058)/0.067427$. The lower triangular part of the matrix illustrates the relative difference (in percentage) of Wednesday-IVRMSE between the j-th and the i-th models, as example: $2.759\% = 100 * (0.05147 - 0.05152)/0.05147$.

IG-GARCH-type	Ess-Ret	Ess-Opt-Ret	Ushp-Opt-Ret	Ess-VIX-Ret	Ushp-VIX-Ret
Ess-Ret	–	9.446	16.000	8.295	14.620
Ess-Opt-Ret	2.759	–	7.234	−1.271	5.716
Ushp-Opt-Ret	6.722	4.076	–	−9.168	−1.637
Ess-VIX-Ret	0.891	−1.920	−6.251	–	6.899
Ushp-VIX-Ret	6.567	3.916	−0.166	5.726	–

Table 4.24: Model comparisons based on the out-of-sample IVRMSE and the Wednesday-IVRMSE for best competitors of each sub-group. Due to the weak difference between the results obtained using Opt-Ret or VIX-Ret information we favor the IVRMSE obtained from Joint MLE estimation using returns and VIX to reduce computational burden. The upper triangular part of the matrix illustrates relative difference (in percentage) of the out-of-sample IVRMSE between the i-th and the j-th models. The lower triangular part of the matrix illustrates relative difference (in percentage) of Wednesday-IVRMSE between the j-th and the i-th models.

	GJR Gaus-Ess	NGARCH NIG-Ess	NGARCH Gaus-Qua	NGARCH NIG-EGP	IG Ushp
GJR-Gaus-Ess	–	8.780	7.603	8.695	11.440
NGARCH-NIG-Ess	14.81	–	−1.290	−0.092	2.913
NGARCH-Gaus-Qua	16.74	2.260	–	1.182	4.149
NGARCH-NIG-EGP	17.50	3.152	0.912	–	3.003
IG-Ushp	18.78	4.659	2.454	1.556	–

Table 4.25: Option pricing performances and VIX predictability for the 24 competitors considered in this chapter.

GARCH	IVRMSE			VIX		
Model	*in*	*out*	*We*	*RMSE*	*MAE*	*MPE*
G.HN.Ret.Ess	0.05939	0.07770	0.06647	0.28494	0.01287	0.011305
G.GJR.Ret.Ess	0.05747	0.07733	0.06511	0.27710	0.01254	0.010604
G.NGARCH.Ret.Ess	0.05718	0.07661	0.06526	0.24753	0.01065	−0.00886
G.HN.Op.Ret.Ess	0.05574	0.07339	0.06101	0.18696	0.01044	−0.00929
G.HN.Ret.VIX.Ess	0.05801	0.07351	0.06145	0.18444	0.00588	−0.00237
G.GJR.Ret.VIX.Ess	0.05483	0.06500	0.05921	0.17404	0.00594	−0.00168
G.NGARCH.Ret.VIX.Ess	0.05570	0.07299	0.05928	0.17480	0.00574	−0.00127
NIG.HN.Ret.Ess	0.05739	0.07004	0.05950	0.18843	0.00748	−0.00546
NIG.GJR.Ret.Ess	0.05502	0.06894	0.05925	0.17585	0.00629	−0.00437
NIG.NGARCH.Ret.Ess	0.05670	0.06900	0.05869	0.17036	0.00533	−0.00296
NIG.HN.Op.Ret.Ess	0.05199	0.06397	0.05103	0.15090	0.00545	−0.00108
NIG.HN.Ret.VIX.Ess	0.05217	0.06488	0.05143	0.13141	0.00513	−0.00060
NIG.GJR.Ret.VIX.Ess	0.05124	0.05956	0.05042	0.11303	0.00471	0.001135
NIG.NGARCH.Ret.VIX.Ess	0.04633	0.05929	0.05044	0.13082	0.00499	0.001045
G.HN.Op.Ret.Qua	0.05110	0.06275	0.05147	0.14451	0.00491	−0.00125
G.HN.Ret.VIX.Qua	0.05137	0.06331	0.05152	0.12488	0.00474	−0.00044
G.GJR.Ret.VIX.Qua	0.05124	0.06289	0.05108	0.11196	0.00410	−0.00036
G.NGARCH.Ret.VIX.Qua	0.05168	0.06240	0.05005	0.11032	0.00417	−0.00040
NIG.NGARCH.Ret.VIX.EGP	0.05016	0.06012	0.04967	0.10662	0.00395	−0.00013
IG.Ret.Ess	0.05435	0.06742	0.05147	0.13641	0.00511	−0.00118
IG.Opt.Ret.Ess	0.04616	0.06105	0.05005	0.13159	0.00432	−0.00049
IG.Opt.Ret.Ushp	0.04354	0.05664	0.04801	0.10616	0.00400	−0.00031
IG.Ret.VIX.Ess	0.04648	0.06183	0.05101	0.10108	0.00398	−0.00008
IG.Ret.VIX.Ushp	0.04387	0.05756	0.04809	0.09909	0.00396	−0.00019

Table 4.26: Rankings of the 24 competitors considered in this chapter in terms of option pricing performances and VIX predictability obtained in Table 4.25.

GARCH	IVRMSE			VIX		
$Model$	in	out	We	$RMSE$	MAE	MPE
$G.HN.Ret.Ess$	24	24	24	24	24	24
$G.GJR.Ret.Ess$	22	23	22	23	23	23
$G.NGARCH.Ret.Ess$	20	22	23	22	22	21
$G.HN.Op.Ret.Ess$	18	20	20	20	21	22
$G.HN.Ret.VIX.Ess$	23	21	21	19	17	17
$G.GJR.Ret.VIX.Ess$	15	14	16	16	18	16
$G.NGARCH.Ret.VIX.Ess$	17	19	18	17	16	15
$NIG.HN.Ret.Ess$	21	18	19	21	20	20
$NIG.GJR.Ret.Ess$	16	16	17	18	19	19
$NIG.NGARCH.Ret.Ess$	19	17	15	15	14	18
$NIG.HN.Op.Ret.Ess$	12	12	9	14	15	11
$NIG.HN.Ret.VIX.Ess$	13	13	11	10	13	9
$NIG.GJR.Ret.VIX.Ess$	8	4	6	7	8	12
$NIG.NGARCH.Ret.VIX.Ess$	4	3	7	9	11	10
$G.HN.Op.Ret.Qua$	7	9	12	13	10	14
$G.HN.Ret.VIX.Qua$	10	11	14	8	9	7
$G.GJR.Ret.VIX.Qua$	9	10	10	6	5	5
$G.NGARCH.Ret.VIX.Qua$	11	8	4	5	6	6
$NIG.NGARCH.Ret.VIX.EGP$	6	5	3	4	1	2
$IG.Ret.Ess$	14	15	13	12	12	13
$IG.Opt.Ret.Ess$	3	6	5	11	7	8
$IG.Opt.Ret.Ushp$	1	1	1	3	4	4
$IG.Ret.VIX.Ess$	5	7	8	2	3	1
$IG.Ret.VIX.Ushp$	2	2	2	1	2	3

Chapter 5

Appendix

5.1 Appendix

Proposition 2.3

Let us first suppose that the pricing equations

$$\begin{cases} \mathbb{E}_{\mathbb{P}}\left\{e^{r}M_{t+1}^{Ushp} \mid \mathcal{F}_t\right\}=1 \\ \mathbb{E}_{\mathbb{P}}\left\{e^{Y_{t+1}}M_{t+1}^{Ushp} \mid \mathcal{F}_t\right\}=1 \\ \pi=\dfrac{h_{t+1}^{*}}{h_{t+1}} \end{cases} \tag{5.1}$$

have a unique solution denoted by $(\theta_{t+1}^{*},\varepsilon_{t+1}^{*},\rho_{t+1}^{*})$. The preceding system can be expressed using the conditional moment generating of the pair (Y_{t+1},y_{t+1}^{-1}) under $\mathbb{P}$:

$$\begin{cases} \mathbb{G}^{\mathbb{P}}_{(Y_{t+1},y_{t+1}^{-1})|\mathcal{F}_t}(\theta_{t+1}^{*},\rho_{t+1}^{*})=e^{-r-\varepsilon_{t+1}^{*}} \\ \mathbb{G}^{\mathbb{P}}_{(Y_{t+1},y_{t+1}^{-1})|\mathcal{F}_t}(\theta_{t+1}^{*}+1,\rho_{t+1}^{*})=e^{-\varepsilon_{t+1}^{*}} \\ \pi=\dfrac{h_{t+1}^{*}}{h_{t+1}}. \end{cases} \tag{5.2}$$

To obtain the dynamics under $\mathbb{Q}^{Ushp}$, we compute the risk-neutral conditional moment generating function of Y_{t+1}:

$$\mathbb{G}^{\mathbb{Q}^{Ushp}}_{Y_{t+1}|\mathcal{F}_t}(u)=\mathbb{E}_{\mathbb{Q}^{Ushp}}\left[e^{uY_{t+1}} \mid \mathcal{F}_t\right]=\mathbb{E}_{\mathbb{P}}\left[e^{uY_{t+1}}e^{r}M_{t+1}^{Ushp} \mid \mathcal{F}_t\right]=e^{r+\varepsilon_{t+1}^{*}}\mathbb{G}^{\mathbb{P}}_{(Y_{t+1},y_{t+1}^{-1})|\mathcal{F}_t}(\theta_{t+1}^{*}+u,\rho_{t+1}^{*}).$$

Using the first equation in (5.2), we can express the risk-neutral moment generating function simply using the historical one:

$$\mathbb{G}^{\mathbb{Q}^{Ushp}}_{Y_{t+1}|\mathcal{F}_t}(u)=\frac{\mathbb{G}^{\mathbb{P}}_{(Y_{t+1},y_{t+1}^{-1})|\mathcal{F}_t}(\theta_{t+1}^{*}+u,\rho_{t+1}^{*})}{\mathbb{G}^{\mathbb{P}}_{(Y_{t+1},y_{t+1}^{-1})|\mathcal{F}_t}(\theta_{t+1}^{*},\rho_{t+1}^{*})}.$$

Given $\mathcal{F}_t$, we know that y_{t+1} follows, under the historical probability $\mathbb{P}$, an IG distribution with degree of freedom $\delta_{t+1}=\frac{h_{t+1}}{\eta^2}$. Thus, using (2.2), we obtain

$$\mathbb{G}^{\mathbb{Q}^{Ushp}}_{Y_{t+1}|\mathcal{F}_t}(u) = \frac{\mathbb{G}^{\mathbb{P}}_{(Y_{t+1},y^{-1}_{t+1})|\mathcal{F}_t}(\theta^*_{t+1}+u,\rho^*_{t+1})}{\mathbb{G}^{\mathbb{P}}_{(Y_{t+1},y^{-1}_{t+1})|\mathcal{F}_t}(\theta^*_{t+1},\rho^*_{t+1})} = e^{u(r+\nu h_{t+1})}\frac{e^{\left[\delta_{t+1}-\sqrt{(\delta^2_{t+1}-2\rho^*_{t+1})(1-2(\theta^*_{t+1}+u)\eta)}\right]}}{e^{\left[\delta^*_{t+1}-\sqrt{(\delta^2_{t+1}-2\rho^*_{t+1})(1-2\theta^*_{t+1}\eta)}\right]}}$$

and

$$\mathbb{G}^{\mathbb{Q}^{Ushp}}_{Y_{t+1}|\mathcal{F}_t}(u) = e^{[u(r+\nu h_{t+1})]+\delta^*_{t+1}\left[1-\sqrt{1-2(u)\eta^*}\right]}$$

where $\eta^* = \dfrac{\eta}{1-2\theta^*_{t+1}\eta}$ [1] and $\delta^*_{t+1} = \sqrt{(\delta^2_{t+1}-2\rho^*_{t+1})(1-2\theta^*_{t+1}\eta)}$. Therefore, we can write

$$Y_{t+1} = r+\nu h_{t+1}+\eta^* y^*_{t+1}$$

where, given $\mathcal{F}_t$, y^*_{t+1} follows an IG distribution with degree of freedom δ^*_{t+1}. In particular the risk neutral volatility at time $t+1$ fulfills $h^*_{t+1} = \eta^*\delta^*_{t+1}$ and we deduce from

$$Y_{t+1} = r+\nu h_{t+1}+\eta^* y^*_{t+1} = r+\nu h_{t+1}+\eta y_{t+1}$$

that $y_{t+1} = \frac{\eta^* y^*_{t+1}}{\eta^*}$. Thus, using that $\pi = \dfrac{h^*_{t+1}}{h_{t+1}}$, (2.1) gives

$$h^*_{t+1} = w^*+bh^*_t+c^*y^*_t+a^*\frac{(h^*_t)^2}{y^*_t}$$

where

$$w^* = w\pi, \quad c^* = \frac{c\pi\eta^*}{\eta}, \quad a^* = \frac{a\eta}{\pi\eta^*}.$$

To conclude the proof it only remains to express η^* using the historical parameters of the model and π. We start from

$$\delta^*_{t+1} = \frac{h^*_{t+1}}{(\eta^*)^2} = \sqrt{(\delta^2_{t+1}-2\rho^*_{t+1})(1-2\theta^*_{t+1}\eta)}.$$

The martingale condition for the risky asset implies $\mathbb{G}^{\mathbb{Q}^{Ushp}}_{Y_{t+1}|\mathcal{F}_t}(1) = e^r$ from which we can extract ρ^*_{t+1} as a function of θ^*_{t+1}:

$$\rho^*_{t+1} = \frac{\delta^2_{t+1}}{2}\left[1-\frac{\nu^2\eta^4}{(1-2\theta^*_{t+1}\eta)\left[1-\left(\sqrt{1-2\eta^*}\right)\right]^2}\right].$$

Thus,

$$\frac{h^*_{t+1}}{(\eta^*)^2} = \frac{-\nu h_{t+1}}{1-\sqrt{1-2\eta^*}}$$

and

$$\pi = \frac{-\nu}{\left[1-\left(\sqrt{1-2\eta^*}\right)\right]}\left[\eta^*\right]^2.$$

[1] A priori, the parameter η^* depends on time through θ^*_{t+1} but as we are going to see below, θ^*_{t+1} is time independent.

Then, the parameter η^* is obtained as the solution of the following cubic equation:

$$(\eta^*)^3 + \frac{2\pi}{\nu}\eta^* + 2\frac{\pi^2}{\nu^2} = 0.$$

It is well known that this equation has a unique real solution if and only if[2]:

$$4\left(\frac{2\pi}{\nu}\right)^3 + 27\left(\frac{\sqrt{2}\pi}{\nu}\right)^4 > 0 \;\Leftrightarrow\; 27\pi > -8\nu.$$

More precisely, we get

$$\eta^* = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}$$

where $p = \dfrac{2\pi}{\nu}$ and $q = 2\dfrac{\pi^2}{\nu^2}$ and we can simplify this expression to obtain

$$\eta^* = \sqrt[3]{\frac{\pi^2}{\nu^2}\left(-1 + \sqrt{1 + \frac{8\nu}{27\pi}}\right)} + \sqrt[3]{\frac{\pi^2}{\nu^2}\left(-1 - \sqrt{1 + \frac{8\nu}{27\pi}}\right)}.$$

Finally, we may deduce from the preceding equality that

$$\theta^*_{t+1} = \frac{1}{2\eta} - \frac{1}{2\left[\sqrt[3]{\frac{\pi^2}{\nu^2}\left(-1 + \sqrt{1 + \frac{8\nu}{27\pi}}\right)} + \sqrt[3]{\frac{\pi^2}{\nu^2}\left(-1 - \sqrt{1 + \frac{8\nu}{27\pi}}\right)}\right]}$$

and that the pricing system (5.1) has a unique solution depending on the historical parameters and π.

■

VIX as a function of the spot volatility (Section 2.1.4)

Under both specifications of the pricing kernel, the risk-neutral dynamics of the IG-GARCH model may be written as

$$\begin{cases} Y_{t+1} = \log\left(\frac{S_{t+1}}{S_t}\right) & = \; r + \nu^* h^*_{t+1} + \eta^* y^*_{t+1} \\ h^*_{t+1} & = \; w^* + b^* h^*_t + c^* y^*_t + a^* \frac{(h^*_t)^2}{y^*_t} \end{cases}$$

where, given $\mathcal{F}_t$, y^*_{t+1} follows an IG distribution with parameter $\frac{h^*_{t+1}}{\eta^*}$ under the risk-neutral probability $\mathbb{Q}$. Thus[3],

[2] From the empirical values of the parameters obtained in Table 4, this condition is always fulfilled in our framework.

[3] Using the fact that an IG random variable Z with degree of freedom δ fulfills $E[\frac{1}{Z}] = \frac{1}{\delta} + \frac{1}{\delta^2}$.

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}\left[h^*_{t+j} \mid \mathcal{F}_{t+j-2}\right] &= w^* + bh^*_{t+j-1} + \frac{c^*}{(\eta^*)^2}h^*_{t+j-1} + a^*\mathbb{E}_{\mathbb{Q}}\left[\frac{(h^*_{t+j-1})^2}{y^*_{t+j-1}} \mid \mathcal{F}_{t+j-2}\right] \\
&= w^* + \left[b + \frac{c^*}{(\eta^*)^2} + a^*\left(\eta^*\right)^2\right]h^*_{t+j-1} + a^*\left(\eta^*\right)^4 \\
&= h^*_{t+j-1}\psi^* + h^*_0\left[1-\psi^*\right]
\end{aligned}$$

where $\psi^* = b + \frac{c^*}{(\eta^*)^2} + a^*\left(\eta^*\right)^2$ is the variance persistence, and $h^*_0 = \frac{w^* + a^*\left(\eta^*\right)^4}{1-\psi^*}$ is the unconditional volatility, under the risk-neutral probability. Now, using the tower property of the conditional expectation operator, the $j-$step ahead prediction of the risk-neutral volatility under the risk neutral measure is given by

$$\mathbb{E}_{\mathbb{Q}}\left[h^*_{t+j} \mid \mathcal{F}_t\right] = h^*_{t+1}\left[\psi^*\right]^{j-1} + h^*_0\left[1-\left(\psi^*\right)^{j-1}\right]$$

and (3.6) follows easily from (4.13).

■

Bibliography

[Aït-Sahalia & Lo (1998)] Y. Aït-Sahalia and A. W. Lo (1998), Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices, *Journal of Finance* , **53**, 499-547. (Cited on page 25.)

[Aït-Sahalia *et al.*(2013)] Y. Aït-Sahalia, J. Fan and Y. Li. (2013), The leverage effect puzzle: Disentangling sources of bias at high frequency, *Journal of Financial Economics, Elsevier* , **109**(1), 224-249. (Cited on page 6.)

[Andersen *et al.*(2003)] T.G. Andersen, T. Bollerslev, F.X. Diebold and P. Labys (2003), Modeling and Forecasting Realized Volatility, *Econometrica*, **71**, 579-625. (Cited on page 3.)

[Baillie & DeGennaro (1990)] R. T. Baillie and R. P. DeGennaro (1990), Stock returns and volatility, *Journal of Financial and Quantitative Analysis*, **25**, 203-214. (Cited on page 12.)

[Baillie & Bollerslev (1989)] R. T. Baillie and T. Bollerslev (1989) Common stochastic trends in a system of exchange rates, *Journal of Monetary Economics* , **44**, 81-167. (Cited on page 13.)

[Badescu *et al.*(2008)] A. Badescu, R. Kulperger and E. Lazar (2008), Option valuation with normal mixture GARCH, *Studies in Nonlinear Dynamics & Econometrics*, **12**(2), 1-42. (Cited on pages 42, 45, 84 and 91.)

[Badescu *et al.*(2009)] A. Badescu, R.J. Elliott and T.K. Siu (2009), Esscher transforms and consumption-based models, *Insurance: Mathematics and Economics*, **45**(3), 337-347. (Cited on pages 42 and 47.)

[Badescu *et al.*(2011)] A. Badescu, R.J. Elliott, R. Kulperger, J. Miettinen and T.K. Siu, (2011), A comparison of pricing kernels for GARCH option pricing with generalized hyperbolic distributions, *International Journal of Theoretical and Applied Finance*, **14**(5), 669-708. (Cited on pages 40, 42, 84, 86, 90, 91, 92, 98, 101 and 107.)

[Badescu *et al.*(2015)] A. Badescu, R. Elliott and J.P. Ortega (2015), Non-Gaussian GARCH Option Pricing Models and their Continuous Time Limits, *European Journal of Operational Research* , **247**(3), 820-830. (Cited on pages 14, 42, 71, 84, 98, 101 and 107.)

[Badescu *et al.*(2017)] A. Badescu, Z. Cui and J.P. Ortega (2017), Non-affine GARCH option pricing models, variance dependent kernels, and diffusion limits, *Journal of Financial Econometrics*, **15**(4), 602-648. (Cited on pages 49, 71, 99 and 102.)

[Badescu *et al.*(2018)] A. Badescu, Y. Chen, M. Couch and Z. Cui (2018), Variance swaps valuation under non-affine GARCH models and their diffusion limits. *Quantitative Finance*, (to appear 2018), Print. (Cited on pages 68, 89, 90, 91, 93 and 96.)

[Bakshi *et al.*(1997)] G. Bakshi, C. Cao, and Z. Chen (1997), Empirical performance of alternative option pricing models, *Journal of Finance* **52**, 2003-2050. (Cited on pages 25, 30, 42 and 52.)

[Bakshi *et al.*(2015)] G. Bakshi, D. Madan and G. Panayotov (2015), Heterogeneity in Beliefs and Volatility Tail Behavior, *Journal of Financial and Quantitative Analysis* , **50**(6), 1389-1414. (Cited on page 42.)

[Bates (1996)] D. S. Bates, (1996), Testing option pricing models. In: *Handbook of Statistics: Statistical Methods in Finance* , (G.S. Maddale & C. R. Rao, eds.), **14**, 567-611. Amsterdam: Elsevier. (Cited on page 42.)

[Barndorff-Nielsen (1977)] O. E. Barndorff-Nielsen (1977), Exponentially decreasing distributions for the logarithm of particle size, *Proc. R. Soc.*, **353**, 401-419. (Cited on page 13.)

[Barndorff-Nielsen (1998)] O. E. Barndorff-Nielsen (1998), Processes of Normal Inverse Gaussian Type, *Finance and Stochastics*, **2** (1), 41-68. (Cited on pages 13, 44 and 71.)

[Barone-adesi *et al.* (2008)] G. Barone-adesi , R. Engle and L. Mancini (2008), A GARCH Option Pricing Model with Filtered Historical Simulation, *Review of Financial Studies*, **21**, 1223-1258. (Cited on pages 12, 25 and 30.)

[Barone-adesi *et al.* (2012)] G. Barone-Adesi and Dall'O, H. (2012), Is the price kernel monotone? ACRN, *Journal of Finance and Risk Perspectives* , **1** , 43-68. (Cited on page 25.)

[Beine *et al.* (2002)] M. Beine, A. Bénassy-Quéré, C. Lecourt (2002), Central bank intervention and foreign exchange rates: new evidence from FIGARCH estimations, *Journal of International Money and Finance*, **21** (1), 115-144. (Cited on page 13.)

[Bera & Higgins (1993)] A. K. Bera and M. L. Higgins. (1993), ARCH models: properties, estimation and testing, *J.Economic Surveys*, **7**, 305-366. (Cited on page 8.)

[Beare & Schmidt (2016)] B. K Beare and Schmidt, L. D. W. (2016), An empirical test of pricing kernel monotonicity, *the Journal of Applied Econometrics*, **31**, 338-356. (Cited on page 25.)

[Berkes *et al.* (2003)] I. Berkes, L. Horváth, and P. Kokoszka, GARCH processes: structure and estimation, *Bernoulli*, **9**, 201-227. (Cited on page 8.)

[Black & Scholes (1973)] F. Black and M. Scholes, (1973), The pricing of Options and corporate liabilities, *Journal of political Economy*, **81**, 637-659. (Cited on pages 1, 3 and 41.)

[Black & Scholes (1976)] F. Black and M. Scholes, (1976), The pricing of commodity contracts, *Journal of Financial Economics*, **3**, 167-179. (Cited on page 6.)

[Bollerslev (1986)] T. Bollerslev (1986), Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, **31**, 307-327. (Cited on pages 3, 7, 10, 41 and 83.)

[Bollerslev *et al.* (1992)] T. Bollerslev,R. Y. Chou, and K. F. Kroner, (1992), ARCH modeling in finance: A review of the theory and empirical evidence, *Journal of Econometrics*, **52**, 5-59. (Cited on page 7.)

[Bollerslev (1987)] T. Bollerslev, (1987), A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return, *Review of Economics and Statistics*, **69**, 542-547. (Cited on pages 12 and 13.)

[Bollen & Inder (2002)] B. Bollen and B. Inder (2002), Estimating Daily Volatility in Financial Markets Utilizing Intraday Data, *Journal of Empirical Finance*, **9**, 551-562 (Cited on page 3.)

[Boothe & Glassman (1987)] P. Boothe P, D. Glassman, (1987), The statistical distribution of exchange rates, *Journal of International Economics*, **22**, 297-319. (Cited on page 13.)

[Brennan (1979)] M. Brennan (1979),The Pricing of Contingent Claims in Discrete-Time Models. *Journal of Finance*, **34**, 53-68. (Cited on pages 6, 16, 20 and 42.)

[Brailsford & Faff (1996)] T. Brailsford and R. Faff (1996) An evaluation of volatility forecasting techniques, *Journal of Banking and Finance*, **20**, 419-38. (Cited on page 12.)

[Bühlmann *et al.* (1996)] H. Bühlmann, F. Delbaen, P. Embrechts & A.N. Shiryaev (1996), No Arbitrage, Change of Measure and Conditional Esscher Transforms in a Semi-Martingale Model of Stock Processes, *CWI of Economics and Statistics*, **9**, 291-317. (Cited on pages 21, 42, 45 and 85.)

[Byun & Cho (2013)] S. J. Byun and H. Cho, (2013), Forecasting carbon futures volatility using GARCH models with energy volatilities, *Energy Economics*, **40**, 207-221. (Cited on page 12.)

[Campbell *et al.* (1997)] J. Campbell, A. W. Lo and A. C. MacKinlay (1997) *The Econometrics of Financial Markets* Princeton, New Jersey: Princeton University Press. (Cited on page 42.)

[Carr & Madan (1999)] P. Carr and D. B. Madan (1999), Option valuation using the fast Fourier transform, *Journal of Computational Finance*, **2**(4), 61-73. (Cited on pages 32, 34, 48 and 67.)

[Carr *et al.* (2002)] P. Carr, H. Geman, D. B. Madan and M. Yor (2002), The Fine Structure of Asset Returns: An Empirical Investigation, *Journal of Business*, **75**, 305-332. (Cited on page 3.)

[Carr & Wu (2006)] P. Carr and L. Wu (2006), A tale of two indices, *Journal of Derivatives*, **13**, 13-29. (Cited on page 93.)

[Chabi-Yo *et al.* (2008)] F. Chabi-Yo, R.Garcia and E. Renault (2008), State dependence can explain the risk aversion puzzle, *Review of Financial Studies*, **21**, 973-1011. (Cited on pages 25 and 42.)

[Chorro *et al.* (2012)] C. Chorro, D. Guégan and F. Ielpo (2012), Option pricing for GARCH-type models with generalized hyperbolic innovations, *Quantitative Finance* , **12**(7), 1079-1094. (Cited on pages 37, 40, 42, 46, 84, 86, 92 and 107.)

[Chorro *et al.* (2015)] C. Chorro, D. Guégan and F. Ielpo (2015), *A Time Series Approach to Option Pricing: Models, Methods and Empirical Performances,* Springer-Verlag Berlin Heidelberg (7), XVI, 188. (Cited on pages 8, 9, 10, 12, 13, 26, 33, 42, 45, 48, 68, 69, 70, 71, 74, 75, 84, 85, 88, 90, 92, 98, 106 and .)

[Chorro & Fanirisoa (2016)] C. Chorro, R. Fanirisoa (2016), *Option Valuation with IG-GARCH model and an **U**-Shaped pricing kernel.* Chapter 1 of the PhD dissertation. (Cited on pages 40, 85, 86, 87, 90, 92, 93, 96, 98, 99, 101, 102 and 107.)

[Chorro & Fanirisoa (2017)] C. Chorro, R. Fanirisoa (2017), *Two steps estimation strategies for NIG-GARCH processes using VIX or option information.* Chapter 2 of the PhD dissertation. (Cited on pages 86, 94, 97, 100 and 103.)

[Christoffersen *et al.* (2004)] P. Christoffersen and K. Jacobs (2004), Which GARCH model for option valuation? , *Management Science*, **50**, 1204-1221. (Cited on pages 8, 30, 40 and 86.)

[Christoffersen *et al.* (2006)] P. Christoffersen, S. Heston and K. Jacobs (2006), Option valuation with conditional Skewness , *Journal of Econometrics*, **131**, 253-284. (Cited on pages 13, 15, 16, 39, 40, 42, 43, 44, 45, 46, 50, 51, 52, 55, 67, 84, 85, 86, 87, 88, 92, 98, 107 and .)

[Christoffersen *et al.* (2008)] P. Christoffersen, K. Jacobs, C. Ornthanalai, and Y. Wang (2008),Option Valuation with Long-Run and Short-Run Volatility Components, *Journal of Financial Economics*, **90**, 272-297. (Cited on page 86.)

[Christoffersen *et al.* (2012)] P. Christoffersen, K. Jacobs and C. Ornthanalai (2012), Dynamic Jump Intensities and Risk Premiums: Evidence from $S\&P500$ Returns and Options, *Journal of Financial Economics*, **106**, 447-472. (Cited on pages 9, 43, 50, 51, 72, 85, 86, 94, 99 and .)

[Christoffersen *et al.* (2013)] P. Christoffersen, S. Heston and K. Jacobs (2013), Capturing Option Anomalies with a Variance-Dependent Pricing Kernel, *Review of Financial studies*, **26**(8), 1963-2006. (Cited on pages 10, 12, 13, 25, 40, 42, 49, 85, 86, 96 and 102.)

[Cochrane (2001)] J. H. Cochrane (2001), Asset Pricing. Princeton, New Jersey: Princeton University Press. (Cited on page 42.)

[Cochrane *et al.* (2012)] J. H. Cochrane, I. Mansur, and B. Odusami, (2013),Volatility persistence in metal returns: A FIGARCH approach, *Journal of Economics and Business*, **64**(4), 287-305. (Cited on page 85.)

[Ding *et al.* (1993)] Z. Ding, C. W. J. Granger and R. F. Engle(1993), A Long Memory Property of Stock Market Returns and a New Model. *Journal of Empirical Finance*, **1**, 83-106. (Cited on pages 11 and 42.)

[Duan (1995)] J. C. Duan (1995), The GARCH option pricing model, *Mathematical Finance* **5**, 13-32. (Cited on pages 9, 10, 19, 20, 21, 30, 41, 48, 70, 84, 85, 90 and 91.)

[Duan & Simonato (1998)] J.C. Duan and J.G. Simonato (1998), Empirical Martingale Simulation for Asset Prices, *Management Science*, **44**, 1218-1233. (Cited on pages 37, 75, 94 and 98.)

[Duan (1999)] J.C. Duan (1999), Conditionally fat-tailed distributions and the volatility smile in options, Working paper. (Cited on pages 21 and 42.)

[Duan *et al.* (2005)] J.C. Duan, P. Ritchken and Z. Sun (2005), Jump Starting GARCH: Pricing and Hedging Options with Jumps in Returns and Volatilities. Working Paper, University of Toronto. (Cited on page 45.)

[Duan *et al.* (2006)] J.C. Duan, P. Ritchken and Z. Sun (2006), Approximating GARCH-Jump Models, Jump Diffusion Processes, and Option Pricing, *Mathematical Finance*, **16**, 21-52. (Cited on page 45.)

[Engle (1982)] R. F. Engle (1982), Autoregressive conditional heterocedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, **50** (4), 987-1007. (Cited on pages 3, 7, 41 and 83.)

[Engle & Gonzalez-Rivera (1991)] R. F. Engle and G. Gonzalez-Rivera (1991), Semiparametric estimation of ARCH model, *Journal of Business and Economic Statistics*, **9** (4), 345-358. (Cited on page 3.)

[Engle & Ng (1993)] R. F. Engle and Victor K. Ng (1993), Measuring and testing the impact of News on Volatility, *Journal of Finance*, **48**, 1749-1777. (Cited on pages 11, 31, 68, 84, 86, 87 and 89.)

[Engle *et al.* (1987)] R. F. Engle, D. M. Lilien and R. P. Robins (1987), Estimating time varying risk premia in the term structure,*Journal of Business and Economic Statistics* , **9**, 345-359. (Cited on page 9.)

[Elliott & Madan (1998)] R.J. Elliott and D.B. Madan (1998), A discrete time equivalent martingale measure, *Mathematical Finance*, **8**, 127-152. (Cited on pages 17, 21, 23, 24, 42, 49, 85, 90 and 91.)

[Elliott *et al.* (2006)] R.J. Elliott, T.K. Siu and L. Chan (2006), Option Pricing for GARCH Models with Markov Switching, *International Journal of Theoretical and Applied Finance*, **9** (6), 825-841. (Cited on page 86.)

[Esscher (1932)] F. Esscher (1932), On the Probability Function in the Collective Theory of Risk, *Skandinavisk Aktuarietidskrift*, **15**, 175-195. (Cited on pages 21, 45 and 85.)

[Fama (1965)] E.F. Fama (1965), The Behaviour of Stock Market Prices, *Journal of Business*, **38**, 34-105. (Cited on page 3.)

[Fernández & Steel (1998)] C. Fernández and M. F. J. Steel (1998), On Bayesian modeling of fat tails and skewness, *J.Amer. Statist. Assoc*, **93**, 359-371. (Cited on page 13.)

[Follmer & Schweizer (1991)] H. Follmer and M. Schweizer (1991), Hedging of contingent claims, *Applied stochastic analysis* , **5**, p. 389. (Cited on page 17.)

[Forsberg & Bollerslev (2002)] L. Forsberg and T. Bollerslev (2002), Bridging the gap between the distribution of realized (ECU) volatility and ARCH modelling (of the euro): The GARCH-NIG model. *Journal of Applied Econometrics*, **17**, 535-548. (Cited on page 15.)

[Francq & Zakoian (2004)] C. Francq and J. M. Zakoian (2004), Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes, *Bernoulli*, **10** (4), 605-637. (Cited on page 28.)

[French *et al.* (1987)] K. R. French, G. W. Schwert and R. E. Stambaugh (1987), Expected Stock Returns and Volatility, *Journal of financial Economics* , **19**, 3-29. (Cited on page 3.)

[Gerber & Shiu (1994)] H. Gerber and E. Shiu (1994), Option Pricing by Esscher Transforms, *Transactions of the Society of Actuaries*, **XLVI**, 99-191. (Cited on pages 16, 17, 21, 22, 42, 45 and 71.)

[Giot *et al.* (2005)] P. Giot, Relationships between Implied Volatility Indexes and Stock Index Returns, *Journal of Portfolio Management*, **31**, 92-100. (Cited on page 85.)

[Giraitis *et al.* (2005)] L. Giraitis, R. Leipus, and D. Surgailis (2005), Recent advances in ARCH modelling In G. TeyssiÎijere and A. Kirman (eds.) Long-Memory in Economics, *Springer-Verlag*, Heidelberg. (Cited on page 8.)

[Glosten *et al.* (1993)] L. R. Glosten, R. Jagannathan and D. E. Runkle (1993), On the relation between the expected value and the volatility if the nominal excess return on stocks, *Journal of Finance*, **48** (5), 1779-1801. (Cited on pages 10, 11, 13, 15, 31, 42, 68, 69, 70, 84, 86, 87 and 89.)

[Gouriéroux & Monfort (2007)] C. Gouriéroux and A. Monfort (2007), Econometric Specifications of Stochastic Discount Factor Models, *Journal of Econometrics*, **136**, 509-530. (Cited on pages 42, 46 and 91.)

[Guégan *et al.* (2013)] D. Guégan, F. Ielpo and H. Lalaharison (2013), Option pricing with discrete time jump processes, *Journal of Economic Dynamics & Control*, **37**, 2417-2445. (Cited on pages 53 and 90.)

[Hardle & Hafner (2000)] W. Hardle and C.M. Hafner (2000), Discrete time option pricing with flexible volatility estimation, *Finance and Stochastics*, **4**, 189-207. (Cited on page 8.)

[Hansen & Singleton (1982)] L. P. Hansen and K. J. Singleton (1982), Generalized instrumental variables estimation of nonlinear rational expectations models, *Econometrica* , **50**, 1269-1286. (Cited on pages 16 and 42.)

[Hansen & Singleton (1983)] L. P. Hansen, K.J. Singleton (1983), Stochastic consumption, risk aversion, and the temporal behavior of asset returns, *Journal of Political Economy*, **91**,249-265. (Cited on page 42.)

[Hansen & Lunde (2005)] L. P. Hansen, A. Lunde (2005), A forecast comparison of volatility models: does anything beat a GARCH(1,1), *Journal of Applied Econometrics*, **20**(7), 873-889. (Cited on pages 8 and 9.)

[Hao & Zhang (2013)] J. Hao and and J. E. Zhang (2013), GARCH Option Pricing Models, the CBOE VIX, and Variance Risk Premium, *Journal of Financial Econometrics* **11**(3), 556-580. (Cited on pages 40, 48, 51, 67, 70, 85, 86, 89 and 93.)

[Harrison& Kreps (1979)] J. M. Harrison and D. M. Kreps (1979), Martingales and arbitrage in multiperiod securities markets, *Journal of Economic Theory*, **20**, 381-408. (Cited on page 42.)

[Heston (1993)] S.L. Heston (1993), A closed-form solution for options with stochastic volatility with applications to bond and currency options, *The Review of Financial Studies*, **6**(2), 327-343. (Cited on pages 10, 48, 67, 83, 84 and 88.)

[Heston & Nandi (2000)] S. Heston and S. Nandi, (2000), A closed-form GARCH option valuation model, *Review of Financial Studies*, **13**, 585-625. (Cited on pages 10, 30, 31, 32, 39, 41, 43, 45, 50, 52, 55, 67, 68, 69, 84, 85, 86, 87, 88, 92, 98 and 106.)

[Hull (2011)] Hull, J. C. (2011), Options, Futures, and Other Derivatives, *Boston*, **8** (Cited on page 2.)

[Huisman *et al.* (1998)] R. Huisman , K. Koedijk, C. Kool, F. Palm, (1998), The fat-tailedness of FX returns. *Mimeo*, Maastricht University. (Cited on page 13.)

[Hsieh (1989)] D. Hsieh, (1989), Testing for Nonlinearity in Daily Foreign Exchange Rate Changes, *Journal of Business*, **62**, 339-368. (Cited on page 12.)

[Jackwerth (2000)] J. Jackwerth (2000), Recovering Risk Aversion from Option Prices and Realized Returns, *Review of Financial Studies* , **13**, 433-451. (Cited on page 25.)

[Johnson *et al.* (1994)] N. L. Johnson, S. Kotz and N. Balakrishnan (1994), Continuous Univariate Distributions, *New York: Wiley*, **2 v**. (Cited on pages 15 and 44.)

[Kanniainen *et al.* (2014)] J. Kanniainen, L. Binghuan and Y. Hanxue (2014), Estimating and using GARCH models with VIX data for option valuation, *Journal of Banking & Finance*, **43**(C), 200-211. (Cited on pages 30, 31, 40, 43, 49, 50, 52, 68, 70, 72, 73, 74, 85, 86, 87, 94, 96, 98, 99, 100, 106 and .)

[Koedijk *et al.* (1992)] K. G. Koedijk, P. A. Stork , C. G. de Vries, (1992), Differences between foreign exchange rate regimes: the view from the tails, *Journal of International Money and Finance*, 462 - 473. (Cited on page 13.)

[Lambert & Laurent (2001)] P. Lambert and S. Laurent, (2001), Modelling financial time series using GARCH-type models and a skewed student density, *Mimeo*, Universite de Liege. (Cited on page 13.)

[Lalancette & Simonato (2017)] S. Lalancette and J.-G. Simonato, (2017), The role of the conditional skewness and kurtosis in VIX index valuation, *European Financial Management* **23**(2), 325-354. (Cited on pages 40, 86, 94 and 97.)

[Liu & Brorsen (1995)] S. Liu and B. W. Brorsen, (1995), Maximum Likelihood Estimation of a GARCH-Stable Model, *Journal of Applied Econometrics*, **10**, 273-285. (Cited on page 13.)

[Loudon *et al.* (2000)] G. Loudon, W. Watt and P. Yadav (2000), An empirical analysis of alternative parametric ARCH models, *Journal of Applied Econometrics*, **15**, 117-36. (Cited on page 12.)

[Markham & Sharpe (2002)] J. W. Markham, M.E. Sharpe (2002), *A Financial History of the United States: From Christopher Columbus to the Robber Barons (1492-1900)*, Volume 1, Retrieved November 26, 2014 (Cited on page 1.)

[Mandelbrot (1963)] B. Mandelbrot (1963), The Variation of Certain Speculative Prices, *Journal of Business*, **36**, 299-319 (Cited on page 3.)

[McMillan *et al.* (2000)] D. McMillan, A. Speight, and O. Apgwilym (2000), Forecasting UK stock market volatility, *Applied Financial Economics*, **10**, 435-480. (Cited on page 12.)

[Monfort & Pégoraro (2012)] A. Monfort and F. Pégoraro (2012), Asset pricing with second order Esscher transforms. *Journal of Banking and Finance*, **36**(6), 1678-1687. (Cited on pages 21, 25, 26, 39, 40, 42, 43, 45, 47, 55, 85, 86, 90, 92, 98 and 106.)

[Nelson (1991)] D. B. Nelson (1991), Conditional heteroscedasticity in asset returns, *Econometrica*, **59** ,347-370. (Cited on pages 10, 11, 13, 15, 42 and 84.)

[Pagan *et al.* (1990)] A. Pagan and G. Schwert (1990), Alternative models for conditional stock volatility, *Journal of Econometrics*, **45**, 267-90. (Cited on page 12.)

[Papantonis (2016)] I. Papantonis (2015), Volatility risk premium implications of GARCH option pricing models, *Economic Modelling*, **58**, 104-115. (Cited on pages 40, 49, 53, 54, 68, 75, 85, 86, 99 and 100.)

[Qiang *et al.* (2015)] L. Qiang, Q. Gaoxiu and S. Guo (2015), VIX forecasting and variance risk premium: A new GARCH approach , *North American Journal of Economics and Finance* **34**, 314-322. (Cited on pages 48 and 99.)

[Ross (1978)] S. A. Ross (1978), A simple approach to the valuation of risky streams, *Journal of Business*, **51**, 453-475. (Cited on page 42.)

[Rubinstein (1976)] M. Rubinstein, (1976), The Valuation of Uncertain Income Streams and the Pricing of Options. *Bell Journal of Economics*, **7**, 407-425. (Cited on pages 6, 16, 17, 20 and 42.)

[Sarwar (2012)] G. Sarwar, (2012), Is VIX an Investor Fear Gauge in BRIC Equity Markets?, *Journal of Multinational Financial Management* **22** (3), 55-65. (Cited on page 85.)

[Stentoft (2008)] L. Stentoft, (2008), American Option Pricing Using GARCH models and the Normal Inverse Gaussian Distribution, *Journal of Financial Econometrics*, **6** (4), 540-582. (Cited on pages 13, 42, 84 and 89.)

[Simonato & Stentoft (2015)] J.G Simonato and L. Stentoft, (2015), Which pricing approach for options under GARCH with non-normal innovations? Working paper. (Cited on page 42.)

[Siourounis (2002)] D. Siourounis (2002), Modeling volatility and testing for efficiency in emerging capital markets: the case of the Athens stock exchange, *Applied Financial Economics*, **12**, 47-55 (Cited on page 12.)

[Siu *et al.* (2004)] T.K. Siu, H. Tong and H. Yang (2004), On Pricing Derivatives under GARCH models: A Dynamic Gerber-Shiu Approach, *North American Actuarial Journal*, **8**, 17-31. (Cited on pages 42, 45, 85, 90 and 91.)

[Taylor (1986)] S. J. Taylor (1986), Modelling Financial Time Series, *Chichester: John Wiley.* (Cited on page 83.)

[Terasvirta (2009)] T. Terasvirta, (2009), An Introduction to Univariate GARCH Models. In: Andersen, T.G., Davis, R.A., Kreiss, J.P. and Mikosch, T. (eds.), *Handbook of Financial Time Series*, 17-42. Springer. (Cited on page 41.)

[Trolle & Schwartz (2009)] A.B. Trolle and E.S. Schwartz, (2009), Unspanned stochastic volatility and the pricing of commodity derivatives, *Review of Financial Studies*, **22**, 4423-4461. (Cited on page 96.)

[Whaley (1993)] R. E. Whaley, (1993), Derivatives on market volatility: Hedging tools long overdue, *Journal of Derivatives* **1**, 71-84. (Cited on page 85.)

[Wedderburn (1974)] R.W.M. Wedderburn (1974),Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method *Biometrika*, **61** (3), 439-447. (Cited on page 28.)

[Yu (2002)] J. Yu (2002), Forecasting volatility in the New Zealand stock market, *Applied Financial Economics*, **12**, 193-202. (Cited on page 12.)

[Zakoian (1994)] J. M. Zakoian, (1994), Threshold heteroskedastic models, *Journal of Economic Dynamics and Control*, **18** (5), 931-955. (Cited on pages 10, 11, 42 and 84.)

[Ziegler (2007)] A. Ziegler, (2007), Why does Implied Risk Aversion Smile?, *Review of Financial Studies*, **20**, 859-904. (Cited on page 42.)

Abstract

For the purposes of pricing European options under practical models, we have studied GARCH process applications, using different pricing kernels alongside a strategy of estimation that integrates returns, options, or VIX information. We conducted this analysis on a large option dataset found in the S&P500 from 1999 to 2012.

In the first chapter, our aim is to extend the IG-GARCH approach and the conditional Escher transform, which emphasise the importance of modelling conditional skewness first dealt with by [Christoffersen *et al.* (2006)]. To this end, we have constructed a related method centred on an extended and non-monotonic version of the exponential affine pricing kernel. This last version includes a U-shaped exponential function, with the aim of improving the flexibility of the link between the risk-neutral and historical distributions, as the tractability of the model is well preserved. Most importantly, in this context, an estimation strategy based on information of returns-VIX provides noteworthy errors on pricing which comes at a low computational, following [Kanniainen *et al.* (2014)]. Consequently, it is possible to combine historical returns at a rational cost with options or VIX information in the estimation process. This creates more accurate joint likelihood as explained in [Christoffersen *et al.* (2012)] and [Kanniainen *et al.* (2014)]. But other than that, the empirical study provides us with very convincing evidence. Indeed, it demonstrates the superiority of the exponential U-shaped pricing kernel for approximating the options' price from the $S\&P500$ for the relevant period. We have carefully performed a GMM test, a crucial step to check the validity of each pricing kernel regarding the martingale conditions. Concerning the IG-GARCH model associated with both exponential affine and exponential U-shaped pricing kernels in addition to estimation using options (VIX information), we have produced a comparative analysis of in-sample and out-of-sample pricing performances. We then calculated the Implied Volatility Root Mean Square (IVRMSE) allocated to each model. We did this in order to estimate and compare the price errors on both returns-option data and VIX data as performance measurements. The empirical results were clear-cut: choosing this new pricing kernel improved the in and out of sample pricing performances of the IG-GARCH.

Let us move on to chapter 2. We here attempt to provide an alternative procedure to the GARCH estimation challenges. Therefore we give greater focus to the analysis of the GARCH-process model estimation. Naturally, this comes with the Normal Inverse Gaussian Distribution innovation (NIG), based on using the two-step Modified-Quasi Maximum Likelihood (QML). The approach also creates uncertainty regarding the GARCH-HN and GARCH-GJR efficiency for option pricing through the use of the empirical martingale simulation. Furthermore, we see a noticeable improvement in the approximation method for the development of numerical efficiency of the Monte Carlo simulation about the GARCH option pricing models described by [Chorro *et al.* (2015)]. These results enable us to draw the conclusion that the GARCH-GJR model with NIG distribution should be regarded as satisfactory means of price process forecasting which provides a more accurate option price.

We also notice that the results of assessing out-of-sample system forecasts, referred to as NIG-GARCH-GJR, combined with information on the VIX index as well as what turned out to be the most effective predictive model on the basis of Inverse Root Mean Square Error. Additionally, we obtained clear results demonstrating that the improvement of the GARCH-GJR models' performance is supported by the VIX-index information. Naturally, this comes with NIG error assumption which supports the findings of previous studies where the VIX-index is held as acceptable for estimation and forecasts pricing options on GARCH type models. Due to these findings, we are able to assert that the option pricing forecast's efficiency and computation speed are also enhanced.

Now, let's move on to chapter 3 where we tried 24 combinations associated with 4 different GARCH frameworks (HN, GJR, IG-GARCH, NGARCH), 3 distributions (Gaussian, NIG, Inverse Gaussian), 4 SDF. There are several points about to discuss: notably the question of the total synergy between the choice of affine or non-affine GARCH specification, the chief uses of competing alternatives to the Gaussian distribution, selecting a correct SDF and the choice of diverse approximation strategies which relies on numerous financial information sets and standard minimization algorithms. We were able to do this by carrying out a rigorous empirical comparison based, not only on the in and out-of-sample pricing performances, but principally on the employment of a weekly rolling window strategy in which each Wednesday model estimation was performed on price options one week later. We have discovered that the risk-neutralized IG-GARCH model which uses a U-shaped pricing kernel gives the most accurate results in a consistent manner, and this, for these three criteria (in, out-of-sample and weekly rolling window). Furthermore, this finding shows that the use of a non-Gaussian distribution in conjunction with a non-standard stochastic discount factor which takes account of the variance risk premium is vital. Moreover, we have confirmed out that the combined VIX-Returns likelihood approximation gives competitive pricing errors at remarkably low computational cost regarding option based estimation procedures. We have observed that this assertion is applicable to every model investigated in this article. Regarding the non-affine GARCH specifications, we obtain noteworthy pricing errors, under NIG innovations only when we integrate VIX information into the strategy of approximation. Concerning the NGARCH type model, this is possible to a high degree of efficiency when we employ the EGP risk neutralization process or the Two-steps approximation strategy discussed in chapter 2. According to our results, the model using options or VIX criteria is significantly more appropriate for pricing $S\&P500$ options, due to the highly significant Spearman's rank correlation coefficient of 0.9 we obtained. In consequence, an extremely effective indication of the correlated pricing performances (at very low computational cost) is obtained through the examination of model performance in a suitable VIX time series. In this way, VIX analysis can be seen as a very useful and parsimonious first-stage evaluation for filtering out the worst GARCH option pricing models.

Keywords: Option valuation, Pricing kernel, VIX index, non-Gaussian GARCH, S&P500.

Resumé

Studente : **N^o étudiant:** 11029499
Nom: Fanirisoa Z. RAHANTAMIALISOA
N^o INE: 0g411k03dt

Nous avons étudié, dans cette recherche, les applications des processus GARCH dans la théorie de l'évaluation des options pour calculer les prix des options européennes, en utilisant différents "pricing kernel" ainsi qu'une stratégie d'estimation intégrant des informations sur les rentabilités, les options ou les informations issus du VIX-index. Nous avons appliqué l'analyse sur un jeu de données d'option de l'indice S&P500, de 1999 à 2012. Le noyau de cette recherche est structurée dans le but :

- Premièrement : d'approfondir l'approche IG-GARCH qui souligne l'importance de capture la caractère leptokurtique et les queues épaisses des distributions empiriques des rentabilités comme proposée par [Christoffersen *et al.* (2006)].

- Deuxièmenent : de proposer une extension non monotone de l'exponentiel affine "Pricing kernel" qui possède une fonction exponentielle en forme de U pour augmenter la flexibilité du lien entre les distributions risque au neutre et les distributions historiques, car le traitement du modèle est bien préservée.

- Troisièmement : de proposer une stratégie d' estimation basée sur les informations de return-VIX qui génère des erreurs de valorisation d'option très raisonnable avec un faible coût de calcul comme détaillé dans [Christoffersen *et al.* (2012)] et [Kanniainen *et al.* (2014)].

- Quatrièmement : de proposer une procédure alternative aux défis de l'estimation de GARCH basée sur l'utilisation du "Two steps Modified-Quasi Maximum Likelihood (QML)". Nous nous concentrons sur l'estimation des processus GARCH avec la Normal Inverse Gaussian (NIG) comme distribution conditionnelle.

Passons maintenant aux performances, qui sont analysées plus en détail au chapitre 3. l'analyse est basée sur l'utilisation d'une vaste collection de modèles GARCH. Nous nous interrogerons notamment sur la synergie entre le choix de la spécification GARCH affine ou non, l'utilisation d'alternatives à la distribution gaussienne, la sélection d'un SDF correct ainsi que le choix de diverses stratégies d'estimation. Ainsi, nous avons testé 24 combinaisons associées à 4 différentes spécifications de GARCH (HN, GJR, IG-GARCH, NGARCH), 3 distributions (Gaussien, NIG, Gaussien inverse), 4 SDF. Il convient de noter que nous avons pu y parvenir en effectuant une comparaison empirique approfondie basée non seulement sur les performances des prix "in-sample" et "out of sample", mais principalement sur l'utilisation d'une stratégie de fenÃłtre glissante hebdomadaire dans laquelle chaque estimation du modèle du mercredi a été faite pour évaluer les options une semaine plus tard. Ensuite, nous avons

calculé le carré moyen de la volatilité implicite (IVRMSE) attribué à chaque modèle. Cela a été fait pour faire une estimation et une comparaison des erreurs de prix sur les données de retour option et les données VIX en tant que mesures de performance. Nous avons obtenu des résultats empiriques très précis et très convaincants:

- En effet, il démontre la supériorité du "Pricing kernel" exponentiel en U en termes d'approximation du prix des options pour l'indice du S&P500 pour la période considérée. Nous avons soigneusement effectué un test GMM nécessaire au contrôle de la validité de chaque Pricing kernel. Le choix de ce Pricing kernel a amélioré les performances de la valorisation des options en utilisant le modèle IG-GARCH.

- En ce qui concerne les spécifications GARCH affines et non affines dans le cadre des innovations NIG, nous pouvons conclure des résultats du chapitre 2 que le modèle GARCH-GJR associé aux distributions NIG est considéré comme un candidat adéquat si nous parlons de processus de détermination du prix d'option.

- En outre, nous avons obtenu des résultats clairs qui montrent que les informations du VIX-index soutiennent l'amélioration des performances des modèles GARCH-GJR avec l'hypothèse d'innovation NIG. Grâce à ces résultats, nous pouvons affirmer que l'efficacité des prévisions de valorisation des options et leur temps de calcul sont également améliorés. Sur cette question, nous avons découvert que l'estimation combinée avec les VIX-index et les rendements donne des erreurs de valorisation concurrentielles à un coût de calcul très raisonnable comparé aux procédures d'estimation par options.

- Pour le modèle NGARCH, les résultats sont aussi attractifs dans le cas où nous utilisons le processus de neutralisation du risque EGP avec la stratégie d'estimation en deux étapes que nous venons de développer au chapitre 2.

Par conséquent, l'analyse des différents modèles en utilisant des informations issues du VIX fournit une très bonne indication sur les performances à un coût de calcul très raisonnable. De cette manière, la comparaison des performances à partir des informations issues du VIX semblé être une première étape d'évaluation très intéressante et parcimonieuse pour éliminer les pires GARCH modèles de valuation des prix d'option.

Keywords: Option valuation, Pricing kernel, VIX index, non-Gaussian GARCH, $S\&P500$.

Astratto

Studente : **Matricola:** 956107
Dottorandoo : Fanirisoa Z. textscRAHANTAMIALISOA
Ciclo:

In questa ricerca, abbiamo studiato, le applicazioni dei processi GARCH nella teoria della valutazione delle opzioni per calcolare i prezzi delle opzioni europee, utilizzando diverse "Pricing kernel", anché una strategia di stima includendo informazioni su ritorni, le opzioni o le informazioni nate dal VIX . Abbiamo applicato l'analisi su un scampolo di dati di opzione del' indice S&P500, del 1999 al 2012.

- Primo: per approfondire l'approccio di IG-GARCH che enfatizza l'importanza dell'acquisizione del carattere leptokurtic e spesse code delle distribuzioni empiriche dei rendimenti come proposto da [Christoffersen *et al.* (2006)].
- Secondo: per proporre un'estensione non-monotono del esponenziale affine "Pricing kernel", che ha una funzione esponenziale a forma di "U" per aumentare la flessibilità del legame tra il rischio di distribuzioni neutri e distribuzioni storiche.
- Terzo: per proporre una strategia per stimare sulla base delle informazioni dei rendimenti e VIX, che genera errori molto ragionevoli di valutazione delle opzioni con un basso costo di calcolo come dettagliato in [Christoffersen *et al.* (2012)] e [Kanniainen *et al.* (2014)].
- Quarto: per proporre una procedura alternativa per le sfide della stima GARCH basata sull'uso dei "Two steps Modified-Quasi Maximum Likelihood(QML)". Ci concentriamo sulla stima del processo GARCH con il NIG "Normal Inverse Gaussian" come distribuzione condizionale.

Adesso veddiamo le prestazioni analizzate in detaglio nel capitolo 3. L'analisi si basa sull'uso di una vasta collezione di modelli GARCH. In fatti,esamineremo la sinergia tra la scelta della specificazione GARCH, l'uso di alternative alla distribuzione gaussiana, la selezione di una SDF corretta cosi come la scelta di varie strategie di stima. In fatti, abbiamo testato 24 combinazioni associati con 4 specificazioni diverse di GARCH (HN, GJR, IG-GARCH, NGARCH), 3 distribuzioni (Gaussien, NIG, Gaussien inverse), 4 SDF. Notiamo che abbiamo potuto raggiungere questo obiettivo eseguendo un accurato confronto empirico basato non solo sulle prestazioni sui prezzi "in-sample" et "out of sample", ma soprattutto sull'uso di una strategia di finestra scorrevole settimanale in cui ogni modello stimato di mercoledÃň è effettuato per valutare le opzioni una settimana più tardi. Quindi, abbiamo calcolato la media Piazza della volatilità implicita (IVRMSE) attribuita a ciascun modello. Questo è stato fatto per confrontare gli errori di prezzo. Abbiamo ottenuto risultati empirici molto specifici, precisi e molto convincenti:

- Pertanto dimostra la superiorità del "Pricing kernel" esponenziale in U nel senso di ravvicinamento del prezzo delle opzioni per l'indice di S&P500 per il periodo studiato. Con attenzione abbiamo effettuato una prova GMM necessario per il controllo di ogni "Pricing kernel". La scelta di questo "Pricing kernel" ha migliorato le valutazione delle opzioni con il modello IG-GARCH associato.

- Per quanto riguarda le specifiche GARCH "affine" e "non-affine" per le innovazioni NIG, possiamo concludere dal Capitolo 2 che il modello GARCH-GJR associato alle distribuzioni NIG è considerato un candidato adatto se stiamo parlando del processo NIG prezzo dell'opzione.

- Inoltre, abbiamo ottenuto risultati chiari che mostrano che delle informazioni di indice VIX migliora le prestazioni dei modelli GARCH GJR con la distribuzione NIG per l'innovazione.

- Da questi risultati, possiamo dire che l'efficacita delle previsioni di valutazione di opzioni e loro tempo di calcolo è anche migliorato. Su questo tema, abbiamo scoperto che stima combinata con l'indice VIX e rendimenti dà errori di valutazione al costo di calcolo molto ragionevole rispetto alle procedure di valutazione di opzioni.

- Per il modello di tipo NGARCH, i risultati sono interessanti nel caso in cui utilizziamo il processo di neutralizzazione del rischio EGP con la strategia di stima in due fasi, che abbiamo appena sviluppato nel Capitolo 2.

Pertanto, l'analisi di diversi modelli che utilizzano informazioni dal VIX fornisce un'ottima indicazione delle prestazioni a un costo computazionale molto ragionevole. In questo modo, confrontare le prestazioni con le informazioni del VIX è sembrato un primo passo molto interessante e parsimonioso nel processo di valutazione per eliminare i peggiori modelli di pricing delle opzioni GARCH.