Option Pricing and Risk Management : Analytic Approaches with GARCH-Lévy Dynamics

Md. Sharif Ullah Mozumder

MS(Actuarial Mathematics), M.Sc.(Mathematics)

A Thesis

for

The Division

of

Accounting and Finance

Presented in Partial Fulfillment of the Requirements

for the degree of Doctor of Philosophy at

Business School, The University of Nottingham

Nottingham, U.K.

August 2011

Abstract

This Ph.D. thesis considers making some contributions to the asset pricing and financial risk management literature. First of all it offers some dynamics in the area of asset pricing which are practically implementable for pricing European style options. More precisely it considers blending GARCH type non-Markovian dynamics with Lévy type Markovian innovations to offer analytic valuation of European style derivatives(at this initial stage). Revealing the mathematical underpinnings– required to replace conditional Gaussian innovations in GARCH option pricing models by innovations coming from some Lévy processes(with one sided and both sided jumps)–is the main focus. The necessity for this arises from the fact that the non-normal(Lévy) innovations are crucial as heteroskedasticity alone doesn't suffice to capture the option smirk and the analytic valuation is highly expected because it makes the model practically implementable. Thus besides incorporating non-normality particular attention is paid to analytic valuation as well; though the Monte Carlo techniques can be readily applied for the proposed dynamics. However an approximation is required to uphold the analytic pricing, especially for innovations coming from Lévy processes which are not Subordinator. These dynamics are capable of overcoming many deficiencies of benchmark Black-Scholes model and can be used to price other derivatives such as Credit, Interest rate, Commodity, Weather etc. The approach is built on a discrete time continuous state space and upholds the no-arbitrage principle of derivative pricing through the use of conditional Esscher transform to configure Equivalent Martingale Measure(EMM). Similar to the existing literature, established for GARCH with normal innovations, existence of EMM provides de-facto evidence in support of no-arbitrage argument. Besides the main focus this research has made some complementary contributions to the option pricing literature.

Since J.P.Morgan introduced RiskMetrics in 1994, the normal quantile based VaR has been considered as industry standard for risk management. However VaR itself has inherent inconsistencies which are exacerbated under the assumption of normality. The second part of this thesis considers two frequently referred approaches to non-normality in risk management : extreme value(EV) approach and Lévy approach. The idea is to reveal the relative performance of various risk measures under full density based Lévy approach and solely tail observation based EV approach. We provide empirical evidence which confirms that though purely tail based risk measures value-at-risk(VaR) and its coherent version expected shortfall(ES) are well comparable under both approaches, entire spectrum based spectral risk measure(SRM) is misleading for EV approach. Backtesting risk measure VaR is considered under both approaches. We plan to improve the computational efficiency of estimation of Lévy coherent risk measures through application of characteristic function based FRFT. Our ultimate goal is to see whether the conditional moment generating functions –developed for GARCH-Lévy models in the first part of this thesis– can be adapted to the characteristic function based FRFT technique in order to estimate the risk measures in analytic fashion.

Acknowledgement

First and foremost sincere acknowledgements are due to Nottingham University Business School for providing me the opportunities to carry out this research. The financial support from ESRC is wholeheartedly acknowledged which made this research possible. Appreciation goes to both my advisors. The time Dr. Ghulam Sorwar manages for me, especially when he was on sabbatical, is truly acknowledged. To Prof. Kevin Dowd it goes without saying. My frequent visits to Sheffield and the affections I enjoyed all the times were exactly what needed to encourage someone when s/he needs it most. Our frequent thoughtful discussions often made many things clear. The cordiality extended by his family members, as well, reflected a true sense of civility and affection. I enjoyed the lovely discussions with several speakers when they used to give talks in seminars. This seminar series is perhaps one of the important facets of Business School's Ph.D. program. All the Ph.D. students always remain glad for the administrative services Miss Andrea Tomlinson sincerely provides to them, there is no exception in my case either. Finally I pay my gratitude to almighty for giving me enough stamina to stay strong, throughout, for such a work which requires huge concentration.

Contents

List of Figures

List of Tables

Introduction

This introduction is in two sections, addressing respectively the motivation and the organisation of the thesis.

Motivation

In continuous time stochastic differential equation approaches it is often felt necessary to assume that the price process is Markovian, otherwise the equations may fail to produce any solution. However the no-memory property of a Markov process is not a good reflection of reality, and there is strong empirical evidence supporting the claim that stock price and interest rate processes are non-Markovian. For example, asset returns exhibit volatility clustering and strong time series structure, implying they are non-Markovian. Some critics even go to the length of saying that it is inappropriate to assume an unrealistic model in order to apply a theory which requires a Markovian modeling. A potential solution is the use of hybrid models which have been a mainstay in the time series literature since the GARCH model with student-t innovations was first proposed by Bollerslev(1987)[21]. However the latest time series analysis suggests that even these models may be inadequate to describe patterns in volatility evolution. Volatility estimates from intraday returns and high-low returns indicate long- lasting volatility shifts than are typically estimated in the ARCH framework, and suggest either a long-memory or multifactor volatility process.

The latest developments in the literature of hybrid models explore discrete time models but replace normal innovations to address skewness and kurtosis related deficiencies. Affine GARCH type models are appealing in this regard: conditional normal innovations can be replaced by innovations coming from Lévy family. In other words such approach blends

Markovian type innovations to non-Markovian type dynamics. GARCH type affine models with normal innovations, see Heston and Nandi(2000)[70], have two major advantages for econometric work. First European call and put options with any maturity T can be computed rapidly conditional upon the observed underlying asset price S_t; the value of relevant underlying latent variable σ_t; various model parameters and the market price of risk(λ) which determines the risk-neutral probability measure. Also the joint characteristic function associated with the joint conditional transition density $p(S_{t+T}, \sigma_{t+T}|S_t, \sigma_t, \Theta)$ has an analytic solution, implying objective transition densities can also be evaluated via Fourier inversion or other fast methods. As a result it becomes relatively straightforward to infer σ_t values from observed option prices and to test whether the observed time series properties of asset and/or option prices conditional on those values of σ_t are consistent with the predicted properties. Under this backdrops researchers developing pricing models along this mixture approach aspires to enrich the dynamics with useful stochastic properties of Lévy processes as innovations. After pioneering work of Heston and Nandi(2000)[70] some researchers tried some simple non-normal(Lévy) innovations in GARCH framework to analytically price European style options. This research attempts to answer:

- How the analytic GARCH approach with normal innovations fares among different approaches which are developed as alternatives to Black and Scholes model? As alternatives we consider Gram Charlier, jump-diffusion, pure jump and continuous time stochastic volatility approaches. Moreover we answer this question in a more realistic set-up where potential investors in derivatives market prefer using only the most recent information.

- Is it possible to mathematically trace GARCH option pricing, with innovations coming from standard both sided Lévy processes, in analytic fashion? If "yes", does their exist explicit relationship between statistical and risk-neutral dynamics?

- How to characterize the market price of risk under such dynamics?

- Is there any improved empirical evidence of pricing across various information aggregations which benefits from analytic valuations?

A research attempting these questions shed lights on improved pricing of other derivatives– e.g. credit, foreign exchange, interest rate etc– where rich stochastic properties blended with time series structures of the dynamics are expected to remove the sources of mispricing.

The relatively recent literature of financial risk management also offers nice scopes for contributions. One such scopes is the accurate quantification of VaR and coherent risk measures (expected shortfall, ES, and spectral risk measure, SRM). Implementing coherent risk measures ES and SRM for Lévy models in a comparative fashion, across leading indices, is not available in the literature. Moreover in case of SRM there are some issues to address. In this part we attempt to answer the question:

- Do the observations discarded by extreme value(EV) model and incorporated by Lévy models play any role in the performance of tail based risk measures VaR and ES? How about SRM?

Then the GARCH- Lévy dynamics come into the scenario which are expected to produce some improved empirical results in risk management, in addition to their improved pricing performance for derivatives. Moreover it will be of considerable practical help if analytic estimation of risk measures can be obtained for these rich dynamics.

Organisation of the Thesis

This thesis is organized as follows.

Chapter1 is an intuitive overview of GARCH features. Different characterizations of GARCH dynamics are studied, and some limitations of GARCH models are reported.

Chapter2 revisits the basics of Lévy modeling and is contributory in nature. This chapter makes a reformulation which demonstrates how the standard Lévy-Kintchine formula may be interpreted as a series of shocks superimposed on a normal distribution. This reformulation gives clear idea about how jumps come into the scenario and distort the basic path structures of Brownian motions in describing the returns of some financial asset. Considering the Variance-Gamma (VG) process as an example it then reveals the detailed mathematical intuitions which underlie the notion of time changing in finance. This in essence helps us recognize and correct a misspecification in Geman(2002)[62] .

Chapter3 is basically an empirical study based on chapter two. Recently Chourdakis (2005)[29] introduces fractional FFT(FRFT) in option pricing which is superior to traditional FFT. Using S&P500 index options we empirically focus on exposition of trade-off between models fitting performance and required calibration time for week by week dynamic calibration with FFT and FRFT specifications. In doing so we further investigate whether FRFT exhibits any distinctive features in addition to its substantial reduction in required computational time. More precisely for Black-Scholes and its time changed version the Variance Gamma model we investigate cross-maturity and cross-strike features of FRFT compared to those of FFT.

Chapter4 is another contributory chapter. In this chapter we consider number of available models which are developed as alternatives to the Black-Scholes model. The models considered in this chapter include Black-Scholes (1973)[19], the Gram-Charlier (GC) approach of Backus et al. (1997)[9], the stochastic volatility (HS) model of Heston (1993)[69], the closed-form GARCH process of Heston and Nandi (2000)[70] and a variety of Levy processes including the Variance Gamma (VG), Normal Inverse Gaussian (NIG), CGMY and Kou(2002)[75] jump-diffusion models. While most of the individual studies in the literature consider a cross-section of these models, we compare all these models using a common point-in-time data that reflects the perspective of a new investor who wishes to choose between models using only the most minimal recent data set. Moreover we compute the hedge factors delta and gamma for each of these models and then examine the accuracy of delta and delta-gamma approximations to the valuation of both individual options and an illustrative option portfolio. Based on the relative performance of Heston Nandi(2000)[70] model (CFG henceforth), in both pricing and approximation, we emphasize the necessity of exploring closed form GARCH approach with non-normal (Lévy) innovations.

Chapter5 is the main chapter of this thesis. In general for analytic derivative pricing the knowledge of the risk neutral distribution at maturity is essential. But the problem is that for the standard GARCH set up only the one step ahead distribution is available. Heston and Nandi(2000)[70] proposed a GARCH-like model with normal innovations where they were able to compute the characteristic function of the underlying using a recursive procedure and then used the Heston(1993)[69] approach to price option using Fourier inversion. For short maturity options, Christoffersen (2006)[33] observed some pricing biases in Heston and Nandi's(2000)[70] model and conjectured that this is due to the fact that single period innovations are Normal. We reveal the mathematical underpinnings required to replace conditional Gaussian innovations in GARCH option pricing models by innovations coming from Lévy processes with both one sided and two sided jumps. The necessity for this arises from the fact that the non-normal (Lévy) innovations are crucial as heteroskedasticity alone doesn't suffice to capture the option smirk and an analytic valuation is important because it makes the model practically implementable. Though we didn't explore this further, it is obvious that like Heston and Nandi(2000)[70] our approach is built on a discrete time continuous state space and upholds the no-arbitrage principle of derivative pricing through the use of a conditional Esscher transform to configure a Equivalent Martingale Measure(EMM). Similar to the one in existing literature, established for GARCH with normal innovations, the existence of EMM provides de-facto evidence in support of no-arbitrage argument.

Naturally, we realize that the innovations coming from Lévy processes with two sided jumps are mathematically cumbersome to deal with, as they require an approximation of volatility dynamics to uphold the analytic valuation methodology in GARCH-Lévy framework. All such cases considered are Brownian motions stochastically time changed by subordinators: VG–Brownian motion time changed by Gamma subordinator, NIG–Brownian motion time changed by inverse Gaussian subordinators, CGMY–Brownian motions time changed by tempered stable subordinators. We detailed the mathematical manipulations required to obtain semi-analytic option prices under GARCH dynamics with all these three innovations. However innovations coming from subordinated Lévy processes which can exhibit only positive jumps, are relatively easier to deal with. This is because in this case we do not require any approximation to obtain an analytic valuation. The case we examine most closely is "analytic GARCH option pricing with tempered stable innovations". The new GARCH-like processes with Lévy innovations, GARCH-Lévy model could be a plausible name, are capable of capturing the conditional skewness and conditional kurtosis.

Moreover it is possible to obtain recursive relations for the evaluation of the characteristic function multi-period ahead which can then yield the closed form prices, up to numerical integration, for European Derivatives. We implement one of the four dynamics, namely those of the GARCH-NIG model, which we examine in detail. The scale of involvement in implementation is the reason behind the selective implementation, especially as the coding involved requires enormous concentration and takes huge amount of calculation time. However once codes are developed and verified, procedures become implementable within manageable times. The point which we must stress here is that this is only possible because of the development of analytic valuation methodology coupled with application of FRFT.

Chapter6 comprises the risk management part of this thesis. In this chapter we revisit the basics of financial risk management. We found that tail based risk measures VaR and ES often forecast the risk almost equally well for both tail based EV model and full density based Lévy models; observations discarded by EV but incorporated by Lévy models do not make for any significant improvement in the performance of tail-based risk measures. It then investigates Lévy spectral risk measure as an alternative to Generalized Pareto spectral risk measure. In case of SRM, however, we observed that full density based Lévy models perform consistently better than solely tail based EV model. To the best of our knowledge this work is the first in its kind where coherent risk measures ES and SRM are implemented for Lévy models in comparative fashion for most of the leading indices over the world. As a consequence we provide clear empirical evidence against the use of SRM when investors prefer EV model to full density based models. On the other hand if for some or other reasons investors prefer to use SRM to quantify the underlying risk, it is better they use Lévy models instead of EV. For the empirical work of this chapter we used the same data as used by Cotter and Dowd(2006)[39][1]. This is because in Cotter and Dowd(2006)[39] they recommended using SRM with EV model to fix the margin requirement in clearing house, without noticing the subtle issue that extreme value model's calibration on few extreme observations often generate inconsistent values of quantiles outside the extreme tail. These quantiles, when used in estimation of SRM, provide a poor estimate of SRM. The idea behind using the same data as in Cotter and Dowd(2006)[39] is just to reveal how poor the EV SRM could be compare to Lévy SRM.

[1]The indices considered are S&P500, FTSE100, DAX, Hang-Seng and Nikkei225.

Part I

Option Pricing

Chapter 1

Basic Dynamics and GARCH

In quantitative Finance and Economics proper characterization of return dynamics is always a vibrant research topic. These dynamics are the fundamental tools for derivative research in general and option pricing in particular. Such dynamics play a pivotal role in currently much talked about financial risk management literature as well. The relative effectiveness of such dynamics is governed by stochastic characteristics of underlying process.

1.1 Short Background on Development of Return Dynamics

We start by referring to the very basic idea of Brownian perturbation:

$$W_{t_{k+1}} = W_{t_k} + \epsilon_{t_k}\sqrt{\Delta t} \tag{1.1}$$

where $t_{k+1} - t_k = \Delta t$ and $k = 0, \cdots, N$ with $t_0 = 0$. Here ϵ_{t_k}'s are independent and identically distributed random variables,(i.i.d henceforth), following $\epsilon_{t_k} \sim N(0,1)$. In the literature this is known as random walk. It follows that for $j < k$ we have

$$W_{t_k} - W_{t_j} = \sum_{i=j}^{k-1} \epsilon_{t_i}\sqrt{\Delta t}. \tag{1.2}$$

The right hand side is a sum of normal random variables , i.e. $W_{t_k} - W_{t_j}$ is always a normal random variable for any $j < k$. It follows immediately that

$$E(W_{t_k} - W_{t_j}) = 0$$

$$Var(W_{t_k} - W_{t_j}) = E(\sum_{i=j}^{k-1} \epsilon_{t_i}\sqrt{\Delta t}) = (k-j)\Delta t = t_k - t_j.$$

Also an immediate consequence of increments over non-overlapping intervals $t_i < t_j \leq t_{j+1} < t_k$, with $i < j < k$, is that $W_{t_k} - W_{t_{j+1}}$ and $W_{t_j} - W_{t_i}$ are independent and hence uncorrelated.

To see the intuitive relation of random walk and increments described in (1.2), consider partitioning $[0,1]$ into "n" subintervals each having length "$\frac{1}{n}$." Then for $t \in [0,1]$ and $\lceil nt \rceil$ being the greatest integer part of its argument define:

$$S_{\lceil nt \rceil} = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lceil nt \rceil} \epsilon_i. \tag{1.3}$$

where ϵ_i's are defined as before. Then clearly

$$S_{\lceil nt \rceil} = S_{\lceil nt \rceil - 1} + \epsilon_{\lceil nt \rceil} \frac{1}{\sqrt{n}}. \tag{1.4}$$

which is a special form of (1.1) with $\Delta t = \frac{1}{n}$ and $W_t = S_{\lceil nt \rceil}$. Furthermore for $t = 1$:

$$S_{\lceil nt \rceil} = S_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \epsilon_i \tag{1.5}$$

has a standard normal distribution. More importantly by central limit theorem, CLT henceforth, S_n tends in distribution to a standard normal variable even when "ϵ_i"'s are only i.i.d and not necessarily normally distributed. To rap things up the process $S_{\lceil nt \rceil}$ tends to a standard Brownian motion in distribution as $n \to \infty$. Equation (1.3) reveals the discrete time intuition behind simulating standard Brownian motion in continuous time:

$$dW_t = \epsilon_t \sqrt{dt} \tag{1.6}$$

The existence of such limit is well studied in the literature, see e.g. Billingsley(1999)[17]. Equation (1.6) plays an intuitive role in the derivation of celebrated Black-Schole-Merton's option pricing formula.

Definition 1.1 *A stochastic process W_t is said to be a standard Brownian motion, SBM henceforth, if it satisfies:*

[SBM1] for $t < s$ $W_t - W_s = W_{t-s} \sim N(0, t-s)$. i.e W_t is stationary.

[SBM2] for $0 \leq t_1 < t_2 \leq t_3 < t_4$, $W_{t_4} - W_{t_3}$ is uncorrelated with $W_{t_2} - W_{t_1}$. This is known as independent increments property.

[SBM3] $W_{t_0} = 0$.

The stationarity ,in particular in financial modeling, implies that the distribution of price appreciation doesn't depend on any particular time. As long as the length of the intervals, over which the the price appreciation is observed, remain same the distribution will remain same, no matter where it is observed. The independent increments property implies that the distribution governing such price fluctuations over non-overlapping observation periods are independent and hence uncorrelated.

Standard Brownian motions are not able to model the average tendency or drift of a process governing the price fluctuation of assets, since over any time interval it models the fluctuations by a zero mean distribution. Arithmetic Brownian motions, ABM henceforth, are thus considered to overcome this limitation of SBM. Under ABM the price fluctuation on an interval of length dt is governed by the stochastic differential equation, SDE henceforth,:

$$dS_t = \mu dt + \sigma dB_t \tag{1.7}$$

where dB_t is a SBM and μ and $\sigma > 0$ are constants. We will revisit the general structure of SDE. For the moment we just mention that SDE's describe the increments of a process, say X, which is driven by one or several governing random processes. When there is only one governing random process SDE's are, in general, described as:

$$dX_t = \mu[t, X_t]dt + \sigma[t, X_t]dB_t \tag{1.8}$$

where B_t is a SBM and $\mu \& \sigma$ are continuous functions of $t \& X$. When $\mu \& \sigma$ are functions of $t \& X_t$ only and doesn't depend on any of X_{t-h} values for $h > 0$, X_t in (1.8) is known as Markov process. So a diffusion process, represented by (1.8), is a Markov process. The drift rate and variance rate's are the limit's:

$$\mu[t, X_t] = \lim_{\Delta \to 0} \frac{E[X_{t+\Delta t} - X_t \mid \mathfrak{F}_t]}{\Delta t} \tag{1.9}$$

$$\sigma[t, X_t]^2 = \lim_{\Delta \to 0} \frac{Var[X_{t+\Delta t} - X_t \mid \mathfrak{F}_t]}{\Delta t} \tag{1.10}$$

respectively, which are also known as instantaneous drift and instantaneous volatility. A diffusion processes, hence a Markov process also, is not a martingale, unless the drift $\mu[t, X_t]$

is identically zero for all X_t and t. The value space and the distribution of future values depend on the function μ and σ.

Hence ABM, described in (1.7), is a diffusion process with $\mu[t, X_t] = \mu$ and $\sigma[t, X_t] = \sigma$. Clearly it is a Markov process and is not a martingale as long as $\mu \neq 0$. Its equivalent, and more intuitive, integral form is:

$$\int_0^t dS_s = \int_0^t \mu ds + \int_0^t \sigma dB_s \tag{1.11}$$

It now follows that

$$S_t = S_0 + \mu t + \sigma B_t. \tag{1.12}$$

So

$$E(S_t - S_0) = E(\mu t + \sigma B_t) = \mu t$$

and

$$Var(S_t - S_0) = Var(\mu t + \sigma B_t) = \sigma^2 t.$$

So mean and variance of price fluctuation over an interval of length t changes linearly with t. This model can be a suitable specification for an economic variable that grows, assuming $\mu > 0$, at a constant rate and is characterized by increasing uncertainty. But as the process can take negative values it is not suitable as a model for stock prices, since limited liability prevents stock prices from going negative. The remedy is what follows: Geometric Brownian Motion(GBM).

GBM is a model for describing the price fluctuation dS_t, on an interval of length dt, relative to the current value S_t. This proportional change, or rate of return, is modeled as an ABM. Consequently the governing SDE is:

$$\frac{dS_t}{S_t} = \mu dt + \sigma dB_t \tag{1.13}$$

where μ and $\sigma > 0$ are constants. Comparing with (1.8) it follows that GBM is a diffusion process with $\mu[t, S_t] = \mu S_t$ and $\sigma[t, S_T] = \sigma S_t$. Hence S_t governed by (1.13) is a Markov process and is not a martingale as long as $\mu \neq 0$. Drift co-efficient μS_t and diffusion co-efficient σS_t are both proportional to the latest known value of the price process S_t, and thus continuously changes. The higher the latest S_t the greater the drift co-efficient and the larger the perturbation. So a bigger random increment is more likely. Ito's formula, which we will discuss later, leads to the integral form of (1.13):

$$\int_0^t dln[S_u] = \int_0^t [\mu - \frac{1}{2}\sigma^2]du + \int_0^t \sigma B_u. \tag{1.14}$$

That is $ln[\frac{S_t}{S_0}] = [\mu - \frac{1}{2}\sigma^2]t + \sigma B_t$. So log returns are normally distributed with parameters:

$$E(ln[\frac{S_t}{S_0}]) = E([\mu - \frac{1}{2}\sigma^2]t + \sigma B_t) = [\mu - \frac{1}{2}\sigma^2]t \tag{1.15}$$

$$Var(ln[\frac{S_t}{S_0}]) = Var([\mu - \frac{1}{2}\sigma^2]t + \sigma B_t) = \sigma^2 t. \tag{1.16}$$

Thus return's $\frac{S_t}{S_0}$ are log-normally distributed:

$$S_t = S_0 \exp\left([\mu - \frac{1}{2}\sigma^2]t + \sigma B_t\right) \tag{1.17}$$

Having an exponential representation, S_t can never be negative. At $S_t = x$ it has the log-normal density:

$$f(x) = \frac{1}{x\nu\sqrt{2\pi}} e^{-\frac{1}{2}[\frac{ln(x)-m}{\nu}]^2} \tag{1.18}$$

where

$$m = E[ln(S_t)] = ln[S_0] + [\mu - \frac{1}{2}\sigma^2]t$$

and

$$\nu = \sqrt{Var[ln(S_t)]} = \sqrt{Var([\mu - \frac{1}{2}\sigma^2]t + \sigma B_t)} = \sigma\sqrt{t}$$

So GBM can be a suitable specification for an economic process which can not assume negative values and whose variability depends linearly on the level of the variable. Thus GMB is the traditional model for the stock prices. Celebrated Black-Schole-Merton ,BSM henceforth, idea capitalizes on GBM for asset return.

Theorem 1.1 *Consider a European option with pay-off $V(S)$ and expiration time T. Assume the continuously compounded rate of interest is r. Then the current European option price is determined by:*

$$\nu(0, S_0) = e^{-rT}\hat{E}[V(S_T)] \tag{1.19}$$

where $\hat{E}$ denotes the expectation under the risk neutral probability that is derived from the risk-neutral process:

$$\frac{dS_t}{S_t} = rdt + \sigma dB_t. \tag{1.20}$$

BSM call option pricing formula takes the explicit form:

Theorem 1.2 *Consider a European call option with strike price K and expiration time T. If the underlying option pays no dividends and continuously compounded risk-free rate is r, then the price of the contract at time t is given by:*

$$C(t, S_t) = S_t\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2) \tag{1.21}$$

where $\Phi(x)$ denotes the cumulative distribution function of standard normal random variable evaluated at the point x, $d_1 = \frac{[ln(\frac{S_t}{K})+(r+\frac{\sigma^2}{2})(T-t)]}{\sigma\sqrt{T-t}}$ and
$d_2 = \frac{[ln(\frac{S_t}{K})+(r-\frac{\sigma^2}{2})(T-t)]}{\sigma\sqrt{T-t}}$ *with $d_2 = d_1 - \sigma\sqrt{T-t}$.*

1.2 Imperfections in Black-Schole's Model

The seminal paper of Black-Schole was a break-through in option pricing literature. But empirical evidence suggests that the model is in conflicts with some of the stylized facts:

- The scale invariance property of Brownian motion leads to the fact that Brownian motion doesn't distinguish itself between time scales where as real price behavior does. Prices move essentially by jumps at intraday scales, at the scale of months they still manifest discontinuous behavior and only after coarse graining their behavior over longer time scale we get something resembling Brownian motion. Though Black-Schole's model can be chosen to give the right variance of return on a given time horizon, it doesn't behave properly under time aggregation, i.e. accross time scale. Since it is difficult to model the behavior of asset treturns equally well across all time scales,ranging from several minutes to several years, it is crucial to make the time scale explicit in various applications from very onset. Thus Black-Scholes's model is certainly not outperforming one on various time scales of practical interest.

- Looking into the early studies in literature, Mandelbrot(1963)[82] and Fama(1965)[55] had indicated that short-run returns in commodity and stock markets are not normally distributed but have fat tailes and are peaked i.e. they have leptokurtic distribution. However for longer investment horizons of a month or more the return distribution seems to converge to a normal distribution.

- Relatively recent evidence has suggested that the assumption of constant volatility is completely inconsistent in financial markets. See Fama(1976)[56].

- There is strong evidence in support of changes in stock prices being negatively correlated with changes in volatility. The phenomenon which is often termed as "leverage effect." Black-Schole , assuming constant volatility, completely fails to report such subtle effect.

- The most resounding failure of Black-Scholes model is its inability to recognize the systematic pattern of impled volatilities exposed by market option prices. When the Black-Scholes formula is inverted to compute the implied volatilities from reported market option prices volatility estimates differ across exercise prices and time to maturity. Two distinct patters are observed when implied volatilities are plotted against strike -prices (or against moneyness, a function of strike price), see Cont and Tankov(2003)[38]. The patterns are "volatility smile" and "volatility skew". As time to maturity increases these curve typically flatten out. Volatility smile is associated with a "U" shaped pattern of impled volatilities where at the money options have the smallest implied volatility. This pattern is common with currency options and in stock index option this pattern has been reported in the period prior to '87 market crash,see Sheikh(1991)[107] and Rubinstein(1994)[96]. After the crash, however, skewed implied volatility patterns are often observed: using post crash S&P 500 index options and futures options Rubinstein(1994)[96] and Derman and Kani(1994)[41] showed that implied volatilities decreases monotonically as the exercise price rises relative to the index level. All these phenomena turn inconsistent with Black-Schole which suggests a flat volatility surface across strike and maturity. It has been conjectured that the underpricing by Black-Schole model, particularly in case of short-run options, is a consequence of disregarding skewness and kurtosis of the return distribution.

- Recent research,see Cont and Tankov(2003)[38], has convincing evidence regarding the presence of jumps in equity price dynamics. In fact inability to trade continuously implies de facto jumps in return dynamics. These jumps contribute to (or may be a source of) stochastic volatilty when they lead to finite variation trajectories in the absence of diffusion term, which is mostly the case in practice. Black-Schole assuming a continuous path with drift (and no other combination or superposition of processes with it) contradicts with de facto presence of jumps.

1.3 Possible Remedies of the Imperfections

Researchers already have substantial contributions to remedied the imperfections surrounding Black-Scholes model. As hinted above it is crucial to include jumps in the return to make the models more realistic across different time scales. As we will see, in this perspective, Lévy process can yield some more realistic models for return dynamics. Approximation of densities considering skewness and Kurtosis are found to improve the performance of Black-Scholes model, see Backus et al(1997)[9]. We will revisit it later with some details. Smile-Skew related remedies are still a vibrant research in empirical finance. Black-Schole's model is not the only continuous time model built on Brownian motion. Considering instantaneous volatility as a local function of price and time, the nonlinear Markov diffusion models are proposed in Dupire(1994)[49] and Derman et al(1994)[41]:

$$\frac{dS_t}{S_t} = \mu dt + \sigma(t, S_t) dB_t. \tag{1.22}$$

In the same line another proposal is the stochastic volatility model, see Heston(1993)[69] and Hull et al(1987)[71], where the price S_t is the component of a bivariate diffusion (S_t, σ_t) driven by a two-dimensional Brownian motion (B_t^1, B_t^2):

$$\frac{dS_t}{S_t} = \mu dt + \sigma(t, S_t) dB_t^1 \tag{1.23}$$

$$\sigma_t = f(Y_t) \qquad dY_t = \alpha_t dt + \gamma_t dB_t^2. \tag{1.24}$$

These models have more flexible statistical properties but as the uncertainty is modeled by Brownian motion the perennial problem of continuity is still there which doesn't seem to be evidenced by real prices over the time scales of interest. Since continuity of paths plays a crucial role in general properties of diffusion models question arises whether results obtained and conclusions drawn from by studying these models are robust to the removal of continuity hypothesis. Studies in quantitative finance in the framework of models with jumps reveal that many results obtained in diffusion models are actually not robust to the presence of jumps in the prices and thus deserve to be considered anew when jumps are taken into account. Thus jumps added to diffusion models are shown to perform much better. It was first introduced in Merton(1976)[85] and a recent reference showing nice empirical performance is Kou(2002)[75]. It is shown that the dynamics:

$$\frac{dS_t}{S_{t-}} = \mu dt + \sigma dB_t + d(\sum_{i=1}^{N_t}(V_i - 1)) \tag{1.25}$$

where V_i is a sequence of independent and identically distributed non-negative random variables such that $Y = log(V)$ has an asymmetric double exponential distribution and N_t is a Poissson process with rate λ. With this jump incorporated dynamics , stochastic volatility can also be incorporated which leads to stochastic volatility jump diffusion model. This kind of models are found performing very well. However jump diffusion processes are Lévy processes[1] and Lévy processes have much more flexibility of characterizing the jumps enhancing the performance of the model.

1.4 What is GARCH?

It has been observed for quite a long time that there are clustering in financial market volatility. A volatile period tends to persist for some time before the market returns to normality. Given that volatility is unlikely to remain constant over time, how could it be modeled so that it responds to time varying shocks? Engle answers this question in Engle(1982)[54] under the name ARCH (Autoresgressive Conditional Heteroskedasticity)and its generalization, see Bollerslev(1986)[20], is what known as GARCH. The ARCH approach was later found to fit many financial time series and its widespread impact on finance has led to Nobel Committee's recognition of Rob Engles work in 2003. GARCH is just another way of modeling the volatility dynamics, specially in discrete time settings. Modeling conditional volatility by GARCH has recently shown to perform much better in capturing empirically observed characteristics in financial return, when option pricing is concerned, and is intuitively more realistic in its approach. GARCH has the elaboration "Generalized Autoregressive Conditional Heteroscedasticity." Following the works Engle(1982)[54] and Bollerslev(1986)[20] a voluminous financial and econometric literature has developed on volatility estimation

[1]See Cont and Tankov(2003)[38], Kyprianou(2006)[76].

and forecasting. By now the GARCH volatility models have become an important tool-kit in empirical asset pricing and financial risk management. Most cited in this voluminous literature concerning empirical finance are Campbell et al(1992)[28], French et al(1987)[58], Glosten et al(1993)[67], and Pagan et al(1990)[90].

1.4.1 Intuition and Examples

Engel's, see Engle(1982)[54], proposal to model the conditional variance σ_t^2 as a linear function of p lagged squared innovations z_t^2 is what known in the literature as ARCH(p) model:

$$z_t \mid \mathfrak{F}_{t-1} \sim N(0, \sigma_t^2) \tag{1.26}$$

$$\sigma_t^2 = \beta_0 + \beta_1 z_{t-1}^2 + \cdots + \beta_p z_{t-p}^2 \tag{1.27}$$

where $\beta_i > 0$ for all i, $\sum_{i=1}^{q} \beta_i < 1$ and $\mathfrak{F}_{t-1}$ represents the information set of all information upto and including $t-1$. That is given the information $\mathfrak{F}_{t-1}$, the next observation z_t has normal distribution with conditional mean $\mathbb{E}(z_t \mid \mathfrak{F}_{t-1}) = 0$, and conditional variance of $\mathbb{V}(z_t \mid \mathfrak{F}_{t-1}) = \sigma_t^2$. Following the idea of general stochastic process we can think of these as the mean and variance of z_t, computed over all paths which agree with $\mathfrak{F}_{t-1}$. Equation (1.27), specifies the way in which the conditional variance σ_t, is determined by the available information. Note that σ_t, is defined in terms of square of past innovations. This together with the assumptions that $\beta_0 > 0$ and $\alpha_i \geq 0$ guarantees that $\sigma_t > 0$. Some common features of ARCH models are:

- Typically q is of high order because of persistence of volatility in financial markets. The way volatility σ_t is constructed in (1.27), it is known at time $t-1$. So one-step-ahead forecast is readily available. Multi-step ahead forecasts can be formulated by assuming $\mathbb{E}[z_{t+\tau}^2] = \sigma_{t+\tau}$.

- It is surprising that if instead of restricting to paths which agree with the available information $\mathfrak{F}_{t-1}$ we consider all possible paths, we have $\mathbb{E}[z_t] = 0$, $\mathbb{V}[z_t] = \frac{\beta_0}{1-\sum_{i=1}^{q} \beta_i}$, a finite constant. To see these consider ARCH(1) and observe that:

$$\mathbb{E}[z_t] = \mathbb{E}[\cdots [\mathbb{E}[\mathbb{E}(z_t \mid \mathfrak{F}_{t-1})] \mid \mathfrak{F}_{t-2}] \cdots \mid \mathfrak{F}_0] = 0 \tag{1.28}$$

since by definition $\mathbb{E}(z_t \mid \mathfrak{F}_{t-1}) = 0$. Similarly, since

$$\mathbb{E}(z_t^2 \mid \mathfrak{F}_{t-1}) = \sigma_t^2 = \beta_0 + \beta_1 z_{t-1}^2,$$
$$\mathbb{E}(\beta_0 + \beta_1 z_{t-1}^2 \mid \mathfrak{F}_{t-2}) = \beta_0 + \beta_1(\beta_0 + \beta_1 z_{t-2}^2)$$

repeatedly applying this argument we have:

$$\mathbb{E}[z_t^2] = \mathbb{E}[\cdots[\mathbb{E}[\mathbb{E}(z_t^2 \mid \mathfrak{F}_{t-1})] \mid \mathfrak{F}_{t-2}] \cdots \mid \mathfrak{F}_0]$$
$$= \beta_0(1 + \beta_1 + \beta_1^2 + \cdots + \beta_1^{t-1}) + \beta_1^t z_0^2$$

That is for large t and $\beta_1 < 1$, $\mathbb{V}[z_t] = \mathbb{E}[z_t^2] = \frac{\beta_0}{1-\beta_1}$. It then follows that $Cov(z_i, z_j) = 0$, if $i \neq j$. That is for an ARCH(q), $\mathbb{V}[z_t] = \frac{\beta_0}{1-\sum_{i=1}^{q}\beta_i}$, hence unconditionally the process is stationary as long as $\sum_{i=1}^{i=q} \beta_i < 1$, which is assumed in the definition of the model. **It is only the conditional volatility which changes with time, not the overall volatility.**

- Though it is in the name **ARCH model is not autoregressive**. However if we add $\eta_t = z_t^2 - \sigma_t^2$, (which, according to the definition of z_t is a zero mean white noise) to both sides of equation (1.27), we get:

$$z_t^2 = \beta_0 + \sum_{i=1}^{q} \beta_i z_{t-i}^2 + \eta_t.$$

  That is **the squared process z_t^2 is autoregressive** with non-zero mean and autoregressive parameters $\beta_1, \beta_2, \cdots, \beta_q$.

- **The ARCH(q) model is nonlinear.** If we could express z_t as $z_t = \sum_{i=1}^{\infty} a_i e_{t-i}$, (for some independent white noise e_t), then we would have $\mathbb{V}(z_t \mid \mathfrak{F}_{t-1}) = \mathbb{V}(z_t \mid e_{t-1}, e_{t-2}, ..) = \mathbb{V}(e_t)$, a constant. This contradicts equation (1.26). So z_t must be a non-linear process.

- **The observations z_t of an ARCH(q) model is not Gaussian though the conditional one is.** Roughly the reason for this is that the unconditional distribution is an average of the conditional distributions for each possible paths upto $t-1$. Although each of these conditional distribution is Gaussian, the variance σ_t's are not equal across 't'. So unconditional distribution is the mixture of normal distribution with unequal variances, which is not normal.

- The distribution of z_t, tends to be more long-tailed than normal, which allows outliers to occur relatively often. The kurtosis in ARCH(1) process is shown, see Bera and Higgins(1993)[13], to be:

$$\frac{\mathbb{E}[z_t^4]}{\sigma_z^4} = 3\left(\frac{1-\beta_1^2}{1-3\beta_1^2}\right) > 3. \tag{1.29}$$

  This is very important since it reflects models leptokurtic behavior which is consistent with the short-run returns in financial data. Moreover once an outlier is included, it will increase the conditional volatility for some time to come. The reason is that any of the larger z_{t-i}, being squared, will make an increasing impact on σ_t as it is defined.

- Since $\mathbb{E}(z_t \mid \mathfrak{F}_{t-1}) = 0$, we see that z_t are Martingale difference. Thus the best estimate of z_t, based on the available information is simply the trivial predictor, namely the series mean 0. However although z_t is not forecastable, the squared series z_t^2 is:

$$\mathbb{E}(z_t^2 \mid \mathfrak{F}_{t-1}) = \mathbb{V}(z_t \mid \mathfrak{F}_{t-1}) = \sigma_t^2 = \beta_0 + \beta_1 z_{t-1}^2 + \cdots + \beta_p z_{t-p}^2$$

- z_t **are not independent,** though they are uncorrelated.This is because if z_t were independent they would form a linear process however as we saw ARCH(q) is not linear.

In application of ARCH(p) type models it's often found that the required p is rather large and so for the shake of a parsimonious parametrization a generalized ARCH(p,q), known as GARCH(p,q), was introduced in Bollerslev(1986)[20] in such a way so that conditional variance is also a function of its own lags of all order upto q:

$$\sigma_t^2 = \beta_0 + \beta_1 z_{t-1}^2 + \cdots + \beta_p z_{t-p}^2 + \alpha_1 \sigma_{t-1}^2 + \cdots + \alpha_q \sigma_{t-q}^2. \tag{1.30}$$

For GARCH(1,1), the constraints $\alpha_1 > 0$ and $\beta_1 > 0$ are needed to ensure positivity of σ_t^2. For higher orders of GARCH the constraints on α_i and β_j are more complex, see Nelson and Cao(1992)[88]. Using the similar intuition as ARCH the unconditional variance can be shown to be:

$$\sigma^2 = \frac{\beta_0}{1 - \sum_{i=1}^{p} \beta_i - \sum_{i=i}^{q} \alpha_i} \tag{1.31}$$

Hence the covariance stationarity in GARCH(p,q) model holds if and only if $\sum_{i=1}^{p} \beta_i + \sum_{i=1}^{q} \alpha_i < 1$.

Convergence of Conditional Variances

Let us get some more insight of volatility forecast in GARCH(1,1) model. The one-step ahead forecast of conditional variance at time t is:

$$\hat{\sigma}_{t+1}^2 = \beta_0 + \beta_1 z_t^2 + \alpha_1 \sigma_t^2. \tag{1.32}$$

Making use of the fact that $\mathbb{E}(z_{t+1}^2 \mid \mathfrak{F}_t) = \sigma_{t+1}^2$, the forecast of σ_{t+2}^2 can be obtained as:

$$\begin{aligned} \hat{\sigma}_{t+2}^2 &= \beta_0 + \beta_1 z_{t+1}^2 + \alpha_1 \sigma_{t+1}^2 \\ &= \beta_0 + (\beta_1 + \alpha_1)\sigma_{t+1}^2 \quad . \end{aligned} \tag{1.33}$$

Similarly,

$$\begin{aligned} \hat{\sigma}_{t+3}^2 &= \beta_0 + (\beta_1 + \alpha_1)\sigma_{t+2}^2 \\ &= \beta_0 + \beta_0(\beta_1 + \alpha_1) + (\beta_1 + \alpha_1)^2 \sigma_{t+1}^2 \\ &= \beta_0 + \beta_0(\beta_1 + \alpha_1) + \beta_0(\beta_1 + \alpha_1)^2 + (\beta_1 + \alpha_1)^2[\beta_1 z_t^2 + \alpha_1 \sigma_t^2] \end{aligned} \tag{1.34}$$

Hence for a large arbitrary forecast horizon τ , we get:

$$\hat{\sigma}_{t+\tau}^2 = \frac{\beta_0}{1 - (\beta_1 + \alpha_1)} + (\beta_1 + \alpha_1)^{\tau - 1}[\beta_1 z_t^2 + \alpha_1 \sigma_t^2]. \tag{1.35}$$

If $\alpha_1 + \beta 1 < 1$, as is assumed in the definition, and τ gets larger and larger then the second term on the right hand side of (1.35) dies out eventually and $\hat{\sigma}_{t+\tau}^2$ converges to the unconditional variance $\frac{\beta_0}{1-(\beta_1+\alpha_1)}$.

Exponential GARCH or EGARCH

To ensure positivity of conditional variance ARCH and GARCH models need to impose non-negativity restrictions on the α_i's and β_j's. Moreover in this early characterization of innovations the GARCH model assumes that the impact of news on the conditional volatility depends only on the magnitude , but not on the sign, of the innovations. But as mentioned above stylized facts suggests that changes in stock prices are negatively correlated with changes in volatility; thus the primitive characterization can't capture the so called leverage effect. To overcome these drawbacks the exponential GARCH , or EGARCH in short, was introduced in Nelson et al(1991)[87] such that logarithm of the conditional variance is specified as:

$$ln\sigma_t^2 = \beta_0 + \beta_1^1 \frac{z_{t-1}}{\sigma_{t-1}} + \beta_1^2 \left(\frac{\mid z_{t-1} \mid}{\sigma_{t-1}} - \mathbb{E}\left[\frac{\mid z_{t-1} \mid}{\sigma_{t-1}} \right] \right) + \alpha_1 ln\sigma_{t-1}^2. \tag{1.36}$$

with $z_t \sim N(0, \sigma_t^2)$. Then $\frac{z_t}{\sigma_t} \sim N(0,1)$ and consequently $\mathbb{E}\left[\frac{|z_t|}{\sigma_t}\right] = \sqrt{\frac{2}{\pi}}$. It is easy to see that the leverage effect is captured by β_1^1. For "good news", i.e. for $\frac{|z_{t-1}|}{\sigma_{t-1}} > 0$, the impact of the innovation z_{t-1} is $(\beta_1^2 + \beta_1^1)\frac{|z_{t-1}|}{\sigma_{t-1}}$ and for "bad news", i.e. for $\frac{|z_{t-1}|}{\sigma_{t-1}} < 0$, it is $(\beta_1^2 - \beta_1^1)\frac{|z_{t-1}|}{\sigma_{t-1}}$. Hence if $\beta_1^1 = 0$, $ln\sigma_t^2$ responds symmetrically to $\frac{|z_{t-1}|}{\sigma_{t-1}}$, in other words non zero β_1^1 captures the leverage effects. Furthermore, since conditional volatility, σ_t, is characterized in terms of log it is always positive and consequently there is no restriction on the sign of the model parameters.

Integrated GARCH or IGARCH

For a GARCH(p,q) process, when $\sum_{i=1}^p \alpha_i + \sum_{i=1}^q \beta_i = 1$ the unconditional variance in (1.31) blows up and the convergence of conditional variances in (1.35) is no longer meaningful. Conditional variance is then described as an integrated GARCH, or IGARCH, and there is no finite fourth moment. Conceptually an infinite variance is counter intuitive to real phenomena in Economics and Finance. However based on the empirical findings in support of GARCH(1,1) as the most popular model for many financial time series, a non-stationary version of GARCH(1,1) (where the persistence parameters α_1 and β_1 sum to 1) was incorporated to EWMA (exponentially weighted moving average) by $Riskmetrics^{TM}$. To see this incorporation first make repeated use of (1.30) to obtain:

$$\begin{aligned}\sigma_{t+2}^2 &= \beta_0 + \beta_1 z_{t+1}^2 + \alpha_1 \sigma_{t+1}^2 \\ &= \beta_0 + \beta_0\alpha_1 + \beta_1 z_{t+1}^2 + \beta_1\alpha_1 z_t^2 + \alpha_1^2\sigma_t^2\end{aligned} \tag{1.37}$$

So

$$\sigma_{t+\tau}^2 = \beta_0 \sum_{i=1}^{\tau} \alpha_1^{i-1} + \beta_1 \sum_{i=1}^{\tau} \alpha_1^{i-1} z_{t+\tau-i}^2 + \alpha_1^\tau \sigma_t^2. \tag{1.38}$$

Thus for $\tau \to \infty$ and $\alpha_1 < 1$, we can infer that:

$$\sigma_t^2 = \frac{\beta_0}{1-\alpha_1} + \beta_1 \sum_{i=1}^{\infty} \alpha_1^{i-1} z_{t-i}^2. \tag{1.39}$$

We have the EWMA model for sample standard deviation such that:

$$\hat{\sigma}_t^2 = \left(\frac{1}{1+\lambda+\lambda^2+\cdots+\lambda^n}\right)(\sigma_{t-1}^2 + \lambda\sigma_{t-2}^2 + \cdots + \lambda^n\sigma_{t-(n-1)}^2) \tag{1.40}$$

If $n \to \infty$ and $\lambda < 1$, we get:

$$\hat{\sigma}_t^2 = (1-\lambda)\sum_{i=1}^{\infty}\lambda^{i-1}\sigma_{t-i}^2. \tag{1.41}$$

When z_t^2 is taken as proxy for σ_t^2, (1.39) and (1.41) are both autoregressive series with long distributed lags, except that (1.39) has an additional constant term.

Nonlinear GARCH or NGARCH

A simple modification of GARCH(1,1) process, defined in (1.30), makes it possible for innovations and volatilities to be negatively correlated,for $\theta > 0$, a phenomenon described as leverage effects:

$$\sigma_{t+1}^2 = \beta_0 + \beta_1\sigma_t(z_t - \theta)^2 + \alpha_1\sigma_t^2. \tag{1.42}$$

In the literature this is known as nonlinear GARCH or NGARCH. The point here is that it is negative piece of news $z_t < 0$, which has more impact on variance than a positive piece of news $z_t > 0$ provided $\theta > 0$. The persistence of variance in this model is $\beta_1(1+\theta^2)+\alpha_1$ and the long-run unconditional variance is $\sigma^2 = \frac{\beta_0}{1-\beta_1(1+\theta^2)-\alpha_1}$.

GJR GARCH and TGARCH

Another GARCH(p,q) model allowing for asymmetric dependencies, i.e. incorporating leverage effects, is Glosten-Jagannathan-Runkle GARCH or GJR-GARCH model:

$$\sigma_t^2 = \beta_0 + \sum_{i=1}^{q}\alpha_i\sigma_{t-i}^2 + \sum_{j=1}^{p}\left(\beta_j z_{t-j}^2 + \gamma_j \mathbb{I}_{j,t-j} z_{t-j}^2\right). \tag{1.43}$$

where,

$$\mathbb{I}_{j,t-j} = \begin{cases} 1 & \text{if } z_{t-j} < 0 \\ 0 & \text{if } z_{t-j} > 0 \end{cases} \tag{1.44}$$

The positivity of conditional variance is ensured by the restrictions $\beta_0 > 0$, $\alpha_i \geq 0, \beta_j \geq 0$ and $\alpha_j + \gamma_j \geq 0$ for $i = 1\cdots q$ and $j = 1\cdots p$. Covariance stationarity holds if and only if:

$$\sum_{i=1}^{p}\alpha_i + \sum_{j=1}^{q}\left(\beta_j + \frac{1}{2}\gamma_j\right) < 1. \tag{1.45}$$

The TGARCH, i.e. threshold GARCH, is similar to GJR-GARCH except that it is formulated with absolute return:

$$\sigma_t^2 = \beta_0 + \sum_{i=1}^{q} \alpha_i \sigma_{t-i}^2 + \sum_{j=1}^{p} (\beta_j \mid z_{t-j} \mid + \gamma_j \mathbb{I}_{j,t-j} \mid z_{t-j} \mid) \,. \tag{1.46}$$

Positivity of the conditional variance is ensured with the restriction on the parameters as before and restriction on covariance stationarity now becomes complicated and in case of $p = q = 1$ it takes the form:

$$\alpha_1^2 + \frac{1}{2}\left[\beta_1^2 + (\beta_1 + \gamma_1)^2\right] + \frac{2}{\sqrt{2\pi}}\alpha_1(\beta_1 + \gamma_1) < 1. \tag{1.47}$$

1.5 GARCH Features for Derivative Pricing

At times when deterministic volatility in classical models was leading poor fit for options data, researchers started thinking to capitalize on GARCH models to fit options data. This was motivated by the success story of GARCH to fit return data. For the first time in literature , Duan(1995)[45] characterizes the relationship between market and risk-neutral probability distributions when the derivative under consideration follows a GARCH dynamic. That was the foundation of pricing European option using GARCH process. Subsequently the theoretical aspects of hedging in the GARCH option pricing model were considered in Garcia and Renault(1998)[61]. Jumps were incorporated in returns and volatility extending the GARCH option pricing model to give more realistic fit to real market option data. See Duan et al(2004)[46].

In scientifically developed time-continuous stochastic differential equation models the Markovian assumption of the underlying price process is required; otherwise the model may fail to produce a solution. See e.g. Shreve(2004)[104], Fusai and Roncoroni(2008)[59]. However the Markov property, from realistic point of view, turns out to be too strong to justify. Strong empirical evidence suggests that stock price processes and interest rate processes are non-Markovian. See e.g. Poon(2005)[93], Jondeau et al(2007)[73]. In fact it is now unanimously accepted that asset returns display the feature of volatility clustering and are of strong time series structure, implying the non-Markovian property. In order to apply sophisticated theory, it seems inappropriate to assume an unrealistic modeling assumption namely Markov structure of the dynamics. It is thus important to have a theory which, from practical point of view, allows non-Markovian modeling for asset prices. Discrete time GARCH processes have strong appeal under such background and can be more realistic candidates for asset price modeling under no-arbitrage condition.

1.5.1 GARCH Option Pricing

The GARCH option pricing model assumes that the per unit expected return of the underlying asset is equal to the risk free rate, r, plus a premium for taking the risk, λ, and a convexity adjustment term. Under such a specification the observed daily return is equal to the expected return plus a innovation term. The most common and starting assumption for the noise term is the conditional Gaussian distribution with mean zero and variance following a GARCH(1,1) process with leverage. That is:

$$R_{t+1} = ln(S_{t+1}) - ln(S_t) = r + \lambda\sqrt{\sigma_{t+1}} - \frac{1}{2}\sigma_{t+1} + \sqrt{\sigma_{t+1}}z_{t+1} \tag{1.48}$$

where $z_{t+1} \sim N(0,1)$ and the volatility dynamic is given by

$$\sigma_{t+1} = \beta_0 + \beta_1\sigma_t + \beta_2\sigma_t(z_t - \theta)^2. \tag{1.49}$$

With the assumption that specification (1.48) of the stock returns is under the physical, or market, measure P, the equation

$$\mathbb{E}^P[\exp(R_{t+1}) \mid \mathfrak{F}_t] = \exp(r + \lambda\sqrt{\sigma_{t+1}}) \tag{1.50}$$

signifies the role and meaning of λ as price of volatility risk. This model assumes that returns are drawn from a normal distribution with time varying volatility accommodating leverage effects. Because of this conditional heteroscedasticity or non-stationarity the unconditional distribution is fat-tailed. To ensure covariance stationarity of the innovation process z_t it is required that the parameter's satisfy

$$\beta_2(1 + \theta^2) + \beta_1 < 1,$$

which in turn ensures the positivity and finiteness of long run unconditional variance of the process given by:

$$\frac{\beta_0}{1-\beta_2(1+\theta^2)-\beta_1} \tag{1.51}$$

See Bollerslev(1986)[20], Berkes et al(2003)[14] and George(2001)[64].

It is easy to see that GARCH process defined by (1.48) and (1.49) reduces to the standard homoskedastic lognormal process of the Black-Schole's model if $\beta_1 = 0$ and $\beta_2 = 0$. That is Black-Schole's model is a special case of GARCH model. To utilize the GARCH approach for option pricing the conventional risk-neutral valuation relationship has to be considered in a local form(only one period ahead) which is known in the literature as local risk neutral valuation relationship, LRNVR, see Duan(1995)[45]. This essentially implies that under locally risk-neutral relationship we must have :

$$R_{t+1} = ln(S_{t+1}) - ln(S_t) = r - \frac{1}{2}\sigma_{t+1} + \sqrt{\sigma_{t+1}}z^*_{t+1} \tag{1.52}$$

$$\sigma_{t+1} = \beta_0 + \beta_1\sigma_t + \beta_2\sigma_t[z^*_t - (\theta+\lambda)]^2. \tag{1.53}$$

where $z^*_t \sim N(0,1)$. This risk-neutral version, corresponding to $\mathbb{Q}$ (say), is characterized in such a way so that is ensures:

$$\mathbb{E}^{\mathbb{Q}}[\exp(R_{t+1}) \mid \mathfrak{F}_t] = \exp\{r\} \tag{1.54}$$

$$\mathbb{V}^{\mathbb{P}}[R_{t+1} \mid \mathfrak{F}_t] = \mathbb{V}^{\mathbb{Q}}[R_{t+1} \mid \mathfrak{F}_t] = \sigma_{t+1}. \tag{1.55}$$

Denoting the new non-centrality parameter by $\theta^* = \theta + \lambda$, and assuming the interest rate r is a given constant, the risk neutral pricing measure is determined by four parameters $\beta_0, \beta_1, \beta_2$ and θ^*.

From (1.52), with MC(Monte Carlo) number of simulated hypothetical risk neutral asset price paths on each time period from present to maturity, we can obtain hypothetical asset price at maturity for each simulated path as :

$$\begin{aligned} S^*_{i,t+T} &= S_t \exp\left\{\sum_{j=1}^{T} R^*_{i,t+j}\right\}, \qquad i = 1,\cdots,MC. \\ &= S_t \exp\left\{r(T) - \frac{1}{2}\left\{\sum_{j=1}^{T}\sigma^*_{i,t+j}\right\} + \left\{\sum_{j=1}^{T}\sqrt{\sigma^*_{i,t+j}}z^*_{i,t+j}\right\}\right\} \end{aligned} \tag{1.56}$$

Here S_t is the present value of the underlying which is known. The equality in (1.56) follows because it is the characterization of a general sample path and model assumption provides a return specification of the form (1.52) for each such sample path and for each of the "T" future time step. See Christoffersen(2003)[34]. Then the option price, say European call, is calculated by taking the average over the future hypothetical payoffs and discounting them to the present:

$$\begin{aligned} C_{GH} &= \exp\{-rT\}\mathbb{E}^{\mathbb{Q}}[\max\{S^*_{t+T} - K, 0\}] \\ &\approx \exp\{-rT\}\frac{1}{MC}\sum_{i=1}^{MC}\max\{S^*_{i,t+T} - K, 0\} \end{aligned} \tag{1.57}$$

Put option prices can also be obtained in the same way. As the number of Monte Carlo replication, MC, gets infinitely large, the average will converge to the expectation. In practice MC=10000 suffices to obtain a good enough estimate. In addition, control variate technique can be used to reduce the variance of the option prices. Similarly option can be priced for other characterization of GARCH processes such as EGARCH, TGARCH etc.

1.5.2 Physical and Risk-neutral Measures

Since switching between physical(or market) measure and risk-neutral measures will be a frequent task, it is better to get some insight into it driven by Christoffersen(2003)[31]. It was basically introduced in Duan(1995)[45], under the name local risk neutral valuation relationship(LRNVR), and plays an important role in making GARCH theory more applicable.

Consider a general innovation function$f$ in (1.49).

$$\sigma_{t+1} = \beta_0 + \beta_1\sigma_t + \beta_2\sigma_t f(z_t). \tag{1.58}$$

The idea is that if f represents some sort of quantitative effect of innovation z_t on an economy in which the volatility process is driven by (1.58), then that quantitative effects should remain same under both physical and risk-neutral measures.

We notice that solving for z_{t+1} from risk-neutral dynamics (1.52) yields:

$$z^*_{t+1} = \frac{R_{t+1} - r + \frac{\sigma_{t+1}}{2}}{\sqrt{\sigma_{t+1}}} \tag{1.59}$$

and solving from physical dynamics (1.49) yields:

$$z_{t+1} = \frac{R_{t+1} - r + \frac{\sigma_{t+1}}{2} - \lambda\sqrt{\sigma_{t+1}}}{\sqrt{\sigma_{t+1}}} = \frac{R_{t+1} - r + \frac{\sigma_{t+1}}{2}}{\sqrt{\sigma_{t+1}}} - \lambda \qquad (1.60)$$

Thus using the intuition above about the role of f in the economy and assuming it is one-one , considering (1.59) and (1.60) we must have:

$$f(z^*_{t+1} - \lambda) = f(z_{t+1}) \qquad \forall f. \qquad (1.61)$$

So from now on we will switch between physical and risk-neutral measures by switching between their corresponding innovations according to (1.61), namely $z^*_{t+1} - \lambda = z_{t+1}$ or $z^*_{t+1} = \lambda + z_{t+1}$.

1.6 Success and Limitations of GARCH Models

A major contribution of the ARCH literature is the findings that apparent changes in the volatility of economic and financial time series may be predictable and possibly results from a specific type of non-linear dependence rather than exogenous structural changes in the variables, see e.g. Bera et al(1993)[13]. In case of financial data, however, large and small errors tend to occur in cluster i.e. large returns are followed by large returns and small by more small, see for example Christoffersen(2003)[34] for such empirical evidence. This suggests that returns are serially correlated. Thus it is logically inconsistent and statistically inefficient to use volatility measures that assume that volatility remains constant over some period when the resulting series moves through time. As argued earlier, unconditional distribution of z_t is always leptokurtic which makes the ARCH return dynamics consistent with the distributional properties of short-run returns in financial market. Furthermore ARCH type models are relatively simpler and easier to handle. They take care of clustered errors and non-linearities. Roughly speaking, such models can accommodate the changes in the econometricians ability to forecast. In Stock(1998)[109], ARCH approach is supported in an elegant way mentioning "any economic variable, in general, evolves on *operational* time scale, while in practice it is measured on a *calender* time scale. And this inappropriate use of calender time scale may lead to volatility clustering since relative to the calender time, the variable may evolve more quickly or slowly."

It has already been established that GARCH models consistently outperforms EWMA in all subperiods and under all evaluation measures. In Pagan et al(1990)[90] it has been established that EGARCH is best specially when compared with some non-parametric methods. To talk about limitation of GARCH approach it must be noted that since multi-period distribution can not be derived in closed form, asset prices must be simulated and parameter estimation involving such simulation is often time consuming. GARCH model features an exponential decay in the autocorrelation , however it has been noted that squared and absolute returns of financial asset typically have serial correlations and decay slowly. Some findings indicated that GARCH superiority is confined to the stock market and for forecasting volatility over shorter horizons only. In option pricing literature the simulation problem is tackled by Heston and Nandi(2000)[70], for the first time in literature. They derived recursive relations which are required to obtain multi-period ahead distributions. However the recursions were possible solely because of a classic relation involving a standard normal variate. This research mainly focuses on upholding similar recursive approach required in closed form valuation but incorporates innovations from much richer stochastic processes known as Lévy processes. In following chapters we study such processes with some details and find scopes in the relevant literature to make some complementary contributions. We consider relative performance of GARCH approach compare to other approaches to option pricing which justifies why further development in this approach should be of interest.

Chapter 2

Lévy Processes for Non-normality

Starting from Markowitz frontier analysis and Capital Asset Pricing Model(CAPM) of early periods to until recently with Value-at-Risk(VAR)computations, assumption of normality for asset returns has been dominant. A natural companion to such assumption is the continuity of paths. As discussed in chapter1, such assumption clearly contradicts many empirical findings leading to serious imperfections of classical Black-Scholes model. In this chapter we discuss the basics of Lévy modelling. In the following chapters some of these Lévy processes, which are well cited for considerable success in capturing more realistic and flexible modelling of real market data, will be investigated and compared with other approaches to option pricing. Later these models will be further explored to deal with risk-management issues. However some criticisms associated with Markov property of Lévy models could be circumvented by considering GARCH-Lévy type dynamics for even more realistic modelling of real market data. Such models blend the non-Markovian structure of GARCH dynamics with potentially non-normal innovation's coming from rich Lévy processes. This provides remedy to imperfections around normal innovations and offers a way to get rid-off strong Markov assumption. We need a concrete section to introduce the underpinnings of Lévy processes. However this literature on Lévy processes has become very vast and is under continuous up-gradation. In introducing the basics of Lévy processes our attention will be to gather working knowledge with some in depth intuitions of working tools. Even so we have to delve into some involved theoretical aspects.

2.1 Basics of Lévy Modelling

This section is intended to make an effective excursion into the theory of Lévy processes. Lévy processes belong to a particular family of stochastic processes with some natural properties giving them the flexibility to capture many important aspects inherent in time series data.

Definition 2.1 *A cadlag stochastic process $X = \{X_t; t \geq 0\}$, on $(\Omega, \mathcal{F}, \mathbf{P})$ with values in $\mathbb{R}$ is called a Lévy process if it satisfies the following properties:*

[L1] each $X_0 = 0$ a.s.

[L2] X_t has independent and stationary increments, i.e.

(i) for every increasing sequence of times $t_0 < t_1 < t_2 < \cdots < t_n$ the random variables $X_{t_0}, X_{t_1} - X_{t_0}, \cdots, X_{t_n} - X_{t_{n-1}}$ are independent.

(ii) $X_{t+h} - X_t \stackrel{D}{=} X_{t+h-t} = X_h$, i.e. the distribution of $X_{t+h} - X_t$ does not depend on t.

[L3] X_t is stochastically continuous, i.e.

$$\lim_{h \to 0} \mathbf{P}(|X_{t+h} - X_t| > \epsilon) = 0, \qquad \forall \epsilon > 0\,.$$

In no way does condition [$L3$] imply that the sample paths are continuous, as we will see in the case of the Poisson process. The intuitive meaning of [$L3$] is that for a given time t (deterministic) the probability of seeing a jump at t is zero, i.e. discontinuities (jumps) do not occur at deterministic times and so occur at random times. It serves to exclude processes with jumps at fixed times which can be regarded as "calender effects" and are not interesting for our modeling purposes. All these facts together with the notion of jumps yield the following result.

Proposition 2.1 *If $X = \{X_t; t \geq 0\}$ is a Lévy process then for fixed $t > 0$, $\triangle X_t = 0$ a.s..*

Proof. Consider a sequence $\{t_n, n \in \mathbb{N}\}$ in $\mathbb{R}^+$ with $t_n \nearrow t$ as $n \to \infty$. Since X has cadlag paths

$$\lim_{n \to \infty} X_{t_n} = X_{t^-}\,.$$

However by $[L3]$ the sequence $\{X_{t_n}; n \in \mathbb{N}\}$ converges in probability to X_t and so has a subsequence which converges almost surely to X_t. Hence the result follows from the definition of jumps and the uniqueness of the limits. $\square$

Remark 2.1 *The above proposition shows that $\triangle X$ is not a straightforward process to analyze.*

Many important intuitions in theory of Lévy processes are direct consequence of infinite divisibility of the underlying probability measure. Next section explains how one is related with other.

2.1.1 Notion of Infinitely Divisible Distributions(IDD)

Increments of a Lévy process are in one-to-one correspondence with infinitely divisible distributions. We present here a brief overview of this relation. For a more general discussion on IDD's we refer to Peter Major[81], Allun Gut(2005)[68] and Bulm and Rosenblat(1959)[24].

By sampling a Lévy process at times $0, \triangle, 2\triangle, 3\triangle, \cdots\cdots$ we simply obtain a random walk

$$S_n(\triangle) = \sum_{k=0}^{n-1} Y_k, \qquad \text{where each} \quad Y_k = X_{(k+1)\triangle} - X_{k\triangle},$$

are IID random variables whose distribution, by$[L2]$, is the same as that of

$$Y_k = X_{(k+1)\triangle} - X_{k\triangle} \stackrel{D}{=} X_{(k+1)\triangle - k\triangle} = X_{\triangle}, \qquad k = 0, 1, \cdots.$$

Since this can be done for any sampling interval $\triangle$ we say that by sampling a Lévy process with different $\triangle$ we specify a whole family of random walks $S_n(\triangle)$.

Choosing $n\triangle = t$, we see that for any $t > 0$ and any $n \geq 1$,

$$\begin{aligned} S_n(\triangle) &= \sum_{k=0}^{n-1} Y_k \\ &= (X_\triangle - X_0) + (X_{2\triangle} - X_\triangle) + \cdots \\ &\qquad\qquad + (X_{(n-1)\triangle} - X_{(n-2)\triangle}) + (X_{n\triangle} - X_{(n-1)\triangle}) \\ &= X_{n\triangle} = X_t. \end{aligned}$$

That is, X_t can be represented as the sum of n iid random variables whose distribution is that of $X_\triangle = X_{t/n}$. Otherwise said, X_t is divided into n iid parts. A distribution having this property is said to be infinitely divisible. Formally:

Definition 2.2 *A distribution function F (or an F distributed random variable X) is said to be infinitely divisible if for any positive integer n there exists independent and identically distributed random variables $Y_1, Y_2, \cdots, Y_n$ such that $Y_1 + Y_2 + \cdots + Y_n$ is F distributed. "Equivalently" a distribution function F is infinitely divisible if and only if its characteristic function*

$$\Phi_F(s) = \int e^{isx} dF(x), \qquad s \in \mathbb{R},$$

can be written for any integer n in the form

$$\Phi_F = [W]^n,$$

such that W is also a characteristic function of some distribution.

The following result characterizes IDD's.

Theorem 2.1 *The following are equivalent:*

[1] X is infinitely divisible.

[2] F_X has a convolution n^{th} root, for any n, that itself is the distribution function of a random variable.

[3] Φ_F has an n^{th} root, for any n, that itself is the characteristic function of a random variable.

For a detailed proof we refer to David Applebeum(2004)[2] and Sato(1999)[100]. If $M_1(\mathbb{R})$ denotes the set of all Borel probability measures on $\mathbb{R}$, a natural extension of the above theorem suggests us to generalize the definition of IDD to distributions that have a convolution n^{th} root in $M_1(\mathbb{R})$.

Proposition 2.2 *$F \in M_1(\mathbb{R})$ is infinitely divisible if and only if for each $n \in \mathbb{N}$ there exists*

$$F^{1/n} \in M_1(\mathbb{R}),$$

for which

$$\Phi_F(s) = \left[\Phi_{F^{1/n}}(s)\right]^n, \qquad \text{for all } s \in \mathbb{R}.$$

For a detailed intuitive proof see Mozumder(2007)[86].

Remark 2.2 *In general the convolution n^{th} root of a probability measure is not unique. However it is always unique when the measure is infinitely divisible. See Feller(1971)[57]. Thus if some IDD is used in modelling the random shocks of returns over a fixed time interval, then the infinite divisibility of the underlying probability measure implies that those shocks to returns are the convoluted sums of other shocks to returns over smaller subintervals of that particular interval. Furthermore the shocks to returns over the smaller subintervals are guaranteed to have unique distribution. The important fact is that those numerous shocks to returns over smaller subintervals, resulting the shocks to returns on the larger time interval, do not have to have the same distribution as the resulting one. In other words $\Phi_{F^{1/n}}(s)$ doesn't have to represent the same distribution as $\Phi_F(s)$, a condition required when $F^{1/n}$ are said to be closed under convolution. In fact we will see that a class of the extremely useful Lévy process, known as generalized hyperbolic Lévy process, is not closed under convolution, though many of its useful subclasses are. Also see example 2.6.*

Examples of IDD's

Example 2.1 Gaussian random variables

A standard result about random variables, see Huynh et al(2008)[72], states that if F has the underlying random variable $X \sim N(\eta, \sigma^2)$ then

$$\Phi_F(s) = e^{is\eta - \frac{1}{2}s^2\sigma^2}, \qquad s \in \mathbb{R}.$$

So we can write

$$\Phi_F(s) = \left[e^{is\frac{\eta}{n} - \frac{1}{2}s^2\frac{\sigma^2}{n}} \right]^n, \qquad s \in \mathbb{R},$$

and hence we can recognize $F^{1/n}$ as the distribution with underlying random variable $Y \sim N(\frac{\eta}{n}, \frac{\sigma^2}{n})$ having the characteristic function

$$\Phi_{F^{\frac{1}{n}}}(s) = e^{is\frac{\eta}{n} - \frac{1}{2}s^2\frac{\sigma^2}{n}}, \qquad s \in \mathbb{R}.$$

Then $\Phi_F(s) = [\Phi_{F^{1/n}}(s)]^n$ and hence $F = (F^{1/n})^{*n}$, which implies by Proposition 2.2, that Gaussian random variables are infinitely divisible.

Example 2.2 Gamma random variables

If F has the underlying random variable $X \sim G(\alpha, \beta)$ then

$$\Phi_F(s) = \int_0^\infty e^{isx} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} dx.$$

That is

$$\Phi_F(s) = \left(\frac{\beta}{\beta - is}\right)^\alpha = \left[\left(\frac{\beta}{\beta - is}\right)^{\frac{\alpha}{n}}\right]^n, \qquad s \in \mathbb{R}. \tag{2.1}$$

Since $\left(\frac{\beta}{\beta - is}\right)^{\frac{\alpha}{n}}$ is the characteristic function of $F^{1/n} \sim Gamma(\frac{\alpha}{n}, \beta)$, we get

$$\Phi_F(s) = [\Phi_{F^{1/n}}(s)]^n \qquad \textit{implying that} \qquad F = (F^{1/n})^{*n}.$$

So gamma random variables are infinitely divisible.

Example 2.3 Poisson random variables

In the univariate case, as shown in Appendix A, if F is from underlying random variable $X \sim Poisson(\lambda)$ its characteristic function is then

$$\Phi_F(s) = e^{\lambda(e^{is}-1)}, \qquad s \in \mathbb{R},$$

so we can write

$$\Phi_F(s) = \left[e^{\frac{\lambda}{n}(e^{is}-1)}\right]^n, \qquad s \in \mathbb{R}.$$

Thus we recognize $F^{1/n}$ as the distribution of a $Poisson(\frac{\lambda}{n})$ random variable with characteristic function $\Phi_{F^{\frac{1}{n}}}(s) = e^{\frac{\lambda}{n}(e^{is}-1)}$. Hence we get

$$\Phi_F(s) = [\Phi_{F^{1/n}}(s)]^n \qquad \textit{implying that} \qquad F = (F^{1/n})^{*n}.$$

So Poisson random variables are infinitely divisible.

Example 2.4 Compound Poisson (CP) random variables

Definition 2.3 *Suppose that $\{Z_n, n \in \mathbb{N}\}$ is a sequence of iid random variables taking values in $\mathbb{R}$ with common law F_Z and let $N \sim Poisson(\lambda)$ be independent of all Z_n. Then the compound Poisson random variable X, denoted $CP(\lambda, F_Z)$, is defined to be $X = Z_1 + Z_2 + \cdots + Z_N$, with $Z = 0$ if $N = 0$, so that we can think of X as a random walk with a random number of steps (jumps), controlled by a $Poisson(\lambda)$ random variable N and with random step sizes Z_i.*

Proposition 2.3 *For $X \sim CP(\lambda, F_Z)$ and each $s \in \mathbb{R}$*

$$\Phi_X(s) = \mathbb{E}\left[e^{is\sum_{i=1}^N Z_i}\right] = \exp\left[\int_{-\infty}^{\infty}(e^{isy}-1)\lambda F_Z(dy)\right].$$

Proof. Let Φ_Z be the common characteristic function of Z_n. By conditioning on the number of jumps and then using independence we get for any $s \in \mathbb{R}$,

$$\begin{aligned}\Phi_X(s) &= \mathbb{E}\left[e^{is\sum_{i=1}^N Z_i}\right] \\ &= \sum_{n=0}^{\infty}\mathbb{E}\left[e^{is(Z_1+\cdots+Z_N)} \mid N=n\right]\mathbf{P}(N=n) \\ &= \sum_{n=0}^{\infty}\mathbb{E}\left[e^{is(Z_1+Z_2+\cdots+Z_n)}\right]e^{-\lambda}\frac{\lambda^n}{n!} = e^{-\lambda}\sum_{n=0}^{\infty}\frac{[\lambda\Phi_Z(s)]^n}{n!}.\end{aligned}$$

That is

$$\Phi_X(s) = \exp\left[\lambda(\Phi_Z(s)-1)\right]. \tag{2.2}$$

Now with $\Phi_Z(s) = \int_{-\infty}^{\infty} e^{isy}F_Z(dy)$ it follows that

$$\Phi_X(s) = \exp\left[\lambda\left(\int_{-\infty}^{\infty} e^{isy}F_Z(dy) - 1\right)\right].$$

Using $\int_{-\infty}^{\infty} F_Z(dy) = 1$ we get that

$$\Phi_X(s) = \exp\left[\lambda\left(\int_{-\infty}^{\infty} e^{isy} - 1\right)F_Z(dy)\right],$$

so the proof is complete. □

Now from (2.2), above, it can be easily seen that

$$\Phi_X(s) = \left[\exp\left[\frac{\lambda}{n}\left(\Phi_Z(s)-1\right)\right]\right]^n, \qquad s \in \mathbb{R},$$

implying that the compound Poisson distribution is infinitely divisible with each division following a $CP(\frac{\lambda}{n}, F_Z)$.

Example 2.5 Inverse Gaussian (IG) random variables

If F has underlying random variable $X \sim IG(\mu,\theta)$ then for its density given by

$$f(x) = \frac{\mu}{\sqrt{2\pi}x^{3/2}}\exp\{\mu\theta\}\exp\Big\{-\frac{1}{2}\Big(\frac{\mu^2}{x}+\theta^2 x\Big)\Big\} \qquad x>0, \ \ and\ \mu,\ \theta>0,$$

the characteristic function can be obtained as:

$$\begin{aligned}\Phi_F(s) &= \int_0^{\infty} e^{isx}\frac{\mu}{\sqrt{2\pi}x^{3/2}}\exp\{\mu\theta\}\exp\Big\{-\frac{1}{2}\Big(\frac{\mu^2}{x}+\theta^2 x\Big)\Big\}dx \\ &= \exp\Big\{-\mu\big(\sqrt{-2is+\theta^2}-\theta\big)\Big\} \qquad s \in \mathbb{R}.\end{aligned} \tag{2.3}$$

So we can write

$$\Phi_F(s) = \left[e^{(-\frac{\mu}{n}(\sqrt{-2is+\theta^2}-\theta))}\right]^n \qquad s \in \mathbb{R}.$$

and hence we can recognize $F^{1/n}$ *as the distribution with underlying random variable* $Y \sim IG(\frac{\mu}{n}, \theta)$ *having the characteristic function*

$$\Phi_{F^{1/n}}(s) = e^{(-\frac{\mu}{n}(\sqrt{-2is+\theta^2}-\theta))} \qquad s \in \mathbb{R}.$$

Then $\Phi_F(s) = \left[\Phi_{F^{1/n}}(s)\right]^n$ *and hence* $F = (F^{1/n})^{*n}$*, which implies by Proposition 2.2 that inverse Gaussian random variables are infinitely divisible.*

We close this section with the following example which shows that the original random variable and the divisor need not necessarily have the same distribution:

Example 2.6 Infinite divisibility with different distributions

Assume that $G_1, G_2 \sim Geo(p)$ *are independent. Then*

$$\begin{aligned}
\mathbf{P}(G_1 + G_2 = n) &= \mathbb{E}\{\mathbf{P}(G_1 + G_2 = n \mid G_1 = k)\} \\
&= \sum_{k=0}^{n} \mathbf{P}(G_1 = k, G_2 = n-k) \\
&= \sum_{k=0}^{n} p^k(1-p)p^{n-k}(1-p) \\
&= \binom{n+1}{1} p^n(1-p)^2 = \binom{n+2-1}{2-1} p^n(1-p)^2.
\end{aligned}$$

Here G_1 *and* G_2 *have the same distribution it is* $G_1 + G_2$ *that is not geometric.*

For more details on the characteristic function and IDD's we refer to George Roussas(2005)[99] and Peter Major[81].

2.1.2 Important Results Concerning IDD's

Now we intend to discuss various results concerning IDD's which are essential for gaining working knowledge on applications of Lévy processes. The first result tells us what happen's when we add two IDD's (or consider the convolution of two IDD measures).

Theorem 2.2 *The sum of two infinitely divisible random variables is itself infinitely divisible.*

The proof results from a similar argument used for Lévy processes. The result implies that a finite sum of IDD's is itself IDD.

With this proposition and the examples of IDD's discussed above we now see how to construct a new IDD, which is in fact the corner stone of the application of Lévy processes. Let $X = X_1 + X_2$, where $X_1 \sim N(\eta, \sigma^2)$ and $X_2 \sim CP(\lambda, F_Z)$ are independent. Then

$$\Phi_X(s) = \exp\left[i\eta s - \frac{1}{2}s^2\sigma^2 + \int_{-\infty}^{\infty} \lambda(e^{isy} - 1)F_Z(dy)\right], \qquad s \in \mathbb{R}. \tag{2.4}$$

By the above definition, Example 2.1, Proposition 2.3, and Theorem 2.1 X is infinitely divisible. So IDD's can be constructed by convolution of Gaussian and compound Poisson random variables. So for time indexed IDD's(Lévy processes) sample paths can be seen as superposition of continuous Brownian motions and some jump processes.

The expression in (2.4) is close to the expression in the celebrated *Lévy Kintchine formula.* This is further explored in the following section.

2.1.3 Lévy-Kintchine formula

Theorem 2.3 *$F \in M_1(\mathbb{R})$ is infinitely divisible if there exists scalars $a, b \in \mathbb{R}$ and a measure ν satisfying $\nu(\{0\}) = 0$ and $\int_{\mathbb{R}\setminus\{0\}} \left(|x|^2 \wedge 1\right) \nu(dx) < \infty$ such that for all $s \in \mathbb{R}$:*

$$\Phi_F(s) = \exp\left[ias - \frac{1}{2}s^2b^2 + \int_{\mathbb{R}} \left[e^{isx} - 1 - isx\mathbb{I}_{\{-1,1\}}(x)\right] \nu(dx)\right] \tag{2.5}$$

Conversely any mapping of the above form is the characteristic function of an infinitely divisible probability measure on $\mathbb{R}$. Thus the parameters "a", "b^2" and the measure ν characterizes the distribution of the underlying infinitely divisible random variable and (a, b^2, ν) together is known as the characteristic triplet or Lévy triplet of the underlying infinitely divisible random variable.

A detailed proof can be found, for example, in Sato(1999)[100], David Applebeum(2004)[2], or Cont and Tankov(2004)[38]. We prefer the proof in David Applebeum(2004)[2] because of its constructive nature. As indicated in the proof, it is worth noting that all infinitely divisible distributions can be constructed as weak limits of the convolution between Gaussian and independent compound Poisson variables. This is precisely the reason why the expression in (2.4) is close to the expression of this Lévy Kintchine formula. As it is the central fact about versatility and predominant application of Lévy processes, we would like to gain more insight about this fact. Consider the last term in (2.5), this can be written as:

$$\begin{aligned} & \exp\left[\int_{\mathbb{R}} \left[e^{isx} - 1 - isx\mathbb{I}_{\{-1,1\}}(x)\right] \nu(dx)\right] \\ = \quad & \exp\left[\int_{|x|\leq 1} \left[e^{isx} - 1 - isx\right] \nu(dx)\right] \exp\left[\int_{|x|>1} \left[e^{isx} - 1\right] \nu(dx)\right] \end{aligned} \tag{2.6}$$

Now exploring the relation $\nu(A) = \lambda F(A)$, we can write the above equation as:

$$\begin{aligned} & \exp\left[\int_{\mathbb{R}} \left[e^{isx} - 1 - isx\mathbb{I}_{\{-1,1\}}(x)\right] \nu(dx)\right] \\ = \quad & \exp\left[\int_{\mathbb{R}} \lambda \left[e^{isx} - 1 - isx\mathbb{I}_{\{-1,1\}}(x)\right] F(dx)\right] \\ = \quad & \exp\left[\int_{|x|\leq 1} \lambda_{sj} \left[e^{isx} - 1 - isx\right] F(dx)\right] \exp\left[\int_{|x|>1} \lambda_{bj} \left[e^{isx} - 1\right] F(dx)\right] \end{aligned} \tag{2.7}$$

As we will see, here $\lambda = \lambda_{bj} + \lambda_{sj}$. Considering proposition 2.3, we can conclude that the last part in (2.7) is the characteristic function of random jumps, satisfying $| x |> 1$, coming from compound Poisson distribution with intensity

$$\begin{aligned} \lambda_{bj} &= \nu(| x |> 1) \\ &= \lambda F(| x |> 1) \\ &= \lambda \int_{|x|>1} F(dx) \end{aligned} \tag{2.8}$$

and distribution of jumps:

$$F_{|J|>1}(dx) = \frac{\nu(dx)\mathbb{I}_{|x|>1}}{\lambda_{bj}} = \frac{\nu(dx)\mathbb{I}_{|x|>1}}{\lambda \int_{|x|>1} F(dx)} \tag{2.9}$$

Then clearly, $F_{|J|>1}(\{| x |> 1\}) = \frac{\nu(\{|x|>1\})}{\lambda \int_{|x|>1} F(dx)} = 1$. The random variable, J(say), describing the jumps of all sizes(with intensity of jumps of all sizes $\nu(\mathbb{R}) = \lambda F(\mathbb{R}) = \lambda$) has distribution $F(x)$. Here "bj" stands for big jumps. That is the last part in (2.7) is the characteristic function of a "$CP\left(\lambda_{bj}, F_{|J|>1}(dx)\right)$".

However it is a bit tricky to get the idea of the first part in (2.7), which we now explore in full details. This part can be written as a limit corresponding to $\epsilon_i \to 0$, as $i \to \infty$:

$$\exp\left[\sum_i \left\{\lambda_{sj}^{\epsilon_i}\mathbb{E}^{F_{\epsilon_i<|J|\le 1}}\left[e^{isJ}-1\right] - is\lambda_{sj}^{\epsilon_i}\mathbb{E}^{F_{\epsilon_i<|J|\le 1}}[J]\right\}\right] \tag{2.10}$$

$$= \exp\left[\sum_i \left\{\lambda_{sj}^{\epsilon_i}\int_{\epsilon_i<|x|\le 1}\left[e^{isx}-1-isx\right]F(dx)\right\}\right] \text{ since J is F distributed}$$

$$= \exp\left[\sum_i \left\{\lambda_{sj}^{\epsilon_i}\int\left[e^{isx}-1-isx\right]F^{\epsilon_i<|x|\le 1}(dx)\right\}\right]$$

$$\overset{i\to\infty}{\longrightarrow} \exp\left[\int_{|x|\le 1}\lambda_{sj}\left[e^{isx}-1-isx\right]F(dx)\right] \tag{2.11}$$

where $\lambda_{sj} = \lambda_{sj}^{\epsilon_1} + \lambda_{sj}^{\epsilon_2} + \cdots + \lambda_{sj}^{\epsilon_i} + \cdots$ is the overall intensity of small jumps. The limit in (2.11) is the characteristic function of a compensated (mean subtracted) square integrable random variable, see Kyprianou(2006)[76]. For a general n each $\lambda_{sj}^{\epsilon_n}$ and $F_{\epsilon_n<|J|\le 1}(dx)$ are given by:

$$\begin{aligned}
\lambda_{sj}^{\epsilon_n} &= \nu(\epsilon_n \mid x \mid\le 1) \\
&= \lambda F(\epsilon_n <\mid x \mid\le 1) \\
&= \lambda\int_{\epsilon_n<|x|\le 1} F(dx) && (2.12)\\
F_{\epsilon_n<|J|\le 1}(dx) &= \frac{\nu(dx)\mathbb{I}_{\epsilon_n<|x|\le 1}}{\lambda_{sj}^{\epsilon_n}} = \frac{\nu(dx)\mathbb{I}_{\epsilon_n<|x|\le 1}}{\lambda\int_{\epsilon_n<|x|\le 1}F(dx)} && (2.13)
\end{aligned}$$

Overall intensity of small jumps with magnitude less than one is given by:

$$\begin{aligned}
\lambda_{sj} &= \nu(\epsilon_1 \mid x \mid\le 1) + \nu(\epsilon_2 \mid x \mid\le 1)\cdots + \nu(\epsilon_n \mid x \mid\le 1) + \cdots \\
&= \nu\left(\{\epsilon_1 \mid x \mid\le 1\} \cup \{\epsilon_2 \mid x \mid\le 1\} \cup\cdots\cup \{\epsilon_n \mid x \mid\le 1\}\cup\cdots\right) \\
&\overset{n\to\infty}{=} \nu(\mid x \mid\le 1) \\
&= \lambda F(\mid x \mid\le 1) \\
&= \lambda\int_{|x|\le 1} F(dx) && (2.14)
\end{aligned}$$

Consider an arbitrary summand in (2.10):

$$\begin{aligned}
&\exp\left\{\lambda_{sj}^{\epsilon_i}\mathbb{E}^{F_{\epsilon_i<|J|\le 1}}\left[e^{isJ}-1\right] - is\lambda_{sj}^{\epsilon_i}\mathbb{E}^{F_{\epsilon_i<|J|\le 1}}[J]\right\} \\
=\ &\exp\left[\lambda_{sj}^{\epsilon_i}\mathbb{E}^{F_{\epsilon_i<|J|\le 1}}\left[e^{isJ}-1\right]\right]\exp\left[-is\lambda_{sj}^{\epsilon_i}\mathbb{E}^{F_{\epsilon_i<|J|\le 1}}[J]\right] && (2.15)
\end{aligned}$$

Considering proposition 2.3, the first part of the expression in (2.15) corresponds to a random variable $\sum_{j=1}^{N} J_j \mathbb{I}_{\epsilon_i<|J_j|\le 1} \sim CP(\lambda_{sj}^{\epsilon_i}, F_{\epsilon_i<|J|\le 1}(dx))$, with $\lambda_{sj}^{\epsilon_i}$ and $F_{\epsilon_i<|J|\le 1}$ are given as (2.12) and (2.13) respectively. Here $N \sim Poisson(\lambda_{sj}^{\epsilon_i})$ and hence we obtain $\mathbb{E}\left[\sum_{j=1}^{N} J_j \mathbb{I}_{\epsilon<|J_j|\le 1}\right] = \mathbb{E}[N]\mathbb{E}^{F_{\epsilon_i<|J|\le 1}}[J] = \lambda_{sj}^{\epsilon_i}\mathbb{E}^{F_{\epsilon_i<|J|\le 1}}[J]$, implying that the second part in (2.15) is the characteristic function of a constant which is the mean of $CP(\lambda_{sj}^{\epsilon_i}, F_{\epsilon_i<|J|\le 1}(dx))$. Thus (2.15) is the characteristic function of a compensated(mean subtracted) compound Poisson random variable of small jumps, which we denote as $\sum_{j=1}^{N} J_j \mathbb{I}_{\epsilon_i<|J_j|\le 1} - \mathbb{E}^{F_{\epsilon_i<|J|\le 1}}[J] \sim CP^c(\lambda_{sj}^{\epsilon_i}, F_{\epsilon_i<|J|\le 1}(dx))$.

Hence applying the similar argument to each summand in (2.10), we see that (2.10) is the characteristic function of the sum of possibly infinite number of compensated compound Poisson random variables:

$$\begin{aligned} CP^c(\lambda_{sj}^{\epsilon_1}, F_{\epsilon_1<|J|\le 1}(dx)) \quad &+ \quad CP^c(\lambda_{sj}^{\epsilon_2}, F_{\epsilon_2<|J|\le 1}(dx)) + \cdots \\ &+ CP^c(\lambda_{sj}^{\epsilon_i}, F_{\epsilon_i<|J|\le 1}(dx)) + \cdots \end{aligned} \tag{2.16}$$

The compensation is required to obtain the convergence of numerous small jumps described by possibly infinite number of compensated compound Poisson random variables, as shown in (2.11), to a compensated square integrable random variable which characteristic function is exactly the first part of the expression in equation (2.7).

Before we close the intuitive discussion on Lévy-kintchine formula for IDD, we see how it generalizes the equation (2.4). When equation (2.4) characterizes the distribution of a random variable $X = \eta + N(0,\sigma^2) + CP(\lambda, F_Z)$, equation (2.5) characterizes the limiting distribution of a sum of $a + N(0, b^2) + CP\left(\lambda_{bj}, F_{|J|>1}(dx)\right) + \sum_i \left\{CP^c(\lambda_{sj}^{\epsilon_i}, F_{\epsilon_i<|J|\le 1}(dx))\right\}$, where the rates and distribution of big and small jumps are as defined earlier.

We will see in next section that Lévy-Kintchine formula of Lévy processes attaches such a limiting random variable at each time point constructing a general stochastic process, a general Lévy process, which can be used to model random evolution of asset prices. In such modelling approach randomly evolved sample paths of asset prices are the superposition of four types of, but possibly infinitely many, randomly evolved sample paths:(i) a linear drift such that on an unit time interval its change is described by the constant "a" ,(ii) a diffusion process such that on unit time interval its distribution is described by a "$N(0, b^2)$",(iii) a compound Poisson process of big jumps such that on unit time its distribution is described by a "$CP\left(\lambda_{bj}, F_{|J|>1}(dx)\right)$" and finally (iv) a limiting process of possibly infinitely many compensated compound Poisson process of small jumps such at on unit time interval each of them is described by "$CP^c(\lambda_{sj}^{\epsilon_i}, F_{\epsilon_i<|J|\leq 1}(dx))$."

We conclude that λ in (2.7) leads to the following intuition:

$$\begin{aligned} \lambda &= \lambda_{bj} + \lambda_{sj} \\ &= \nu(|\, x \,|> 1) + \nu(|\, x \,|\leq 1) \\ &= \lambda F(|\, x \,|> 1) + \lambda F(|\, x \,|\leq 1) \\ &= \lambda F(\{|\, x \,|> 1\} \cup \{|\, x \,|\leq 1\}) \\ &= \lambda F\left(\mathbb{R}\right) \\ &= \nu(\mathbb{R}) \end{aligned} \tag{2.17}$$

As λ is the intensity, i.e. expected number of jumps of all sizes, Equation (2.17) leads to the interpretation of Lévy measure as the expected number of jumps whose sizes belong to a certain Borel set. For example $\nu(A)$ is the expected number of jumps whose sizes belong to A. This intuition extends from IDD to for Lévy processes. That is *Lévy measure of Borel set A is the expected number of jumps per unit time provided jump sizes belong to A.*

Mathematical Fact about Lévy Measure

As introduced in Lévy-Kintchine formula ν is a Borel measure defined on $\mathbb{R} \setminus \{0\}$. We say that ν is a Lévy measure if

$$\int_{\mathbb{R}\setminus\{0\}} \left(|x|^2 \wedge 1\right) \nu(dx) < \infty \tag{2.18}$$

or equivalently

$$\int_{|x|\leq 1} |x|^2 \nu(dx) < \infty \qquad \text{and} \qquad \int_{|x|\geq 1} \nu(dx) < \infty. \tag{2.19}$$

Since $\left(|x|^2 \wedge \epsilon\right) \leq \left(|x|^2 \wedge 1\right)$ for all $0 < \epsilon \leq 1$ it follows that

$$\int_{\mathbb{R}\setminus\{0\}} \left(|x|^2 \wedge \epsilon\right) \nu(dx) \leq \int_{\mathbb{R}\setminus\{0\}} \left(|x|^2 \wedge 1\right) \nu(dx), \qquad \text{if} \quad 0 < \epsilon \leq 1,$$

and hence from (2.18) it follows that

$$\int_{\mathbb{R}\setminus\{0\}} \left(|x|^2 \wedge \epsilon\right) \nu(dx) < \infty \Rightarrow \nu\left[(-\epsilon, \epsilon)^c\right] < \infty \qquad \text{for,} \qquad 0 < \epsilon \leq 1. \tag{2.20}$$

So we have

$$\int_{|x|\leq\epsilon} |x|^2\nu(dx) < \infty \quad \text{and} \quad \int_{|x|\geq\epsilon} \nu(dx) < \infty, \qquad \text{if} \quad 0 < \epsilon \leq 1. \tag{2.21}$$

We will gain more insight on Lévy measure when we will study the results for Lévy process using those for IDD's.

The Lévy Exponent

Theorem 2.3 shows that the Lévy -Kintchine formula is related to the characteristic function of an infinitely divisible distribution F (or an F distributed random variable). It can be expressed, using two parameters "a", "b^2" and a measure "ν", as an exponential function with a complex exponent, i.e.

$$\Phi_F(s) = e^{\Psi(s)} \quad \text{where} \quad \Psi : \mathbb{R} \longrightarrow \mathbb{C}. \tag{2.22}$$

The complex function Ψ is known as the characteristic exponent or Lévy exponent of F (or an F distributed random variable).

Since we know that $|\Phi_F(s)| \leq 1$, see Allun Gut(2005)[68], then with the assumption that $\Psi = Re(\Psi) + iIm(\Psi)$ we have

$$\begin{aligned} \left|\Phi_F(s)\right| &= \left|e^{\Psi(s)}\right| = \left|e^{Re(\Psi)+iIm(\Psi)}\right| \\ &= \left|e^{Re(\Psi)}\right|\left|e^{iIm(\Psi)}\right| = e^{Re(\psi)}1 \leq 1. \end{aligned}$$

Hence $e^{Re(\Psi)} \leq 1$ implies that $Re(\Psi) \leq 0$, that is the characteristic exponent should always have a non positive real part.

The following theorems enhance the appreciation of Lévy processes in applications.

Theorem 2.4 *Any infinitely divisible probability measure can be constructed as the weak limit of a sequence of compound Poisson distributions.*

For a proof we refer to David Applebeum(2004)[2].

In a more general framework we have the following result.

Proposition 2.4 *If $\{F_n\}$ is a sequence of infinitely divisible distributions and $F_n \longrightarrow F$, then F is infinitely divisible, i.e. weak limits of sequences of infinitely divisible probability measures are infinitely divisible.*

Again for an intuitive proof we refer to David Applebeum(2004)[2].

2.2 IDD and Lévy Processes

The modeling intuitions described in section 2.1 needs to be incorporated into processes so the the richness of IDD can be extracted at each time point. As we described in subsection 2.1.3, usefulness of such models–from practical point of view–is highly likely. In other words the main target is to extract the richness of modeling through IDD in describing the evolution of the Lévy process at each time point. The first result is at the core of such possibility.

Theorem 2.5 *If $X = \{X_t; t \geq 0\}$ is a Lévy process, then X_t is infinitely divisible for each $t \geq 0$.*

The proof is based on David Applebeum(2004)[2], and is reproduced as it illuminates the intuition.

Proof. For each $n \in \mathbb{N}$, we can write

$$X_t = Y_1^n(t) + Y_2^n(t) + \cdots + Y_n^n(t),$$

where each

$$\begin{aligned} Y_k^n(t) &= X_{\frac{kt}{n}} - X_{\frac{(k-1)t}{n}}, && k = 1, 2, \cdots, n, \\ &\stackrel{D}{=} X_{(\frac{kt}{n} - \frac{kt}{n} + \frac{t}{n})}, && \text{by [L2]-(ii) of Definition 2.1}, \\ &= X_{\frac{t}{n}}. \end{aligned}$$

The last term in the above equality is independent of k, which shows that for all k, the $Y_k^n(t)$'s are iid with the common distribution $X_{\frac{t}{n}}$. Hence

$$\Phi_{X_t}(s) = \left[\Phi_{X_{\frac{t}{n}}}(s)\right]^n, \qquad s \in \mathbb{R},$$

which shows that X_t, for each $t \geq 0$, is infinitely divisible. □

Theorem2.5 ensures that X_t is infinitely divisible, for each $t \geq 0$. Hence by Lévy-Kintchine formula its distribution is described by a characteristic function of the form (2.5), through a set of parameters and a measure, at time $t = 1$. The following argument clarifies how to characterize the distribution for a general t, using the characterization at $t = 1$. This is one of the main facts about Lévy processes which tells us that characterizing the distribution of the whole process is equivalent to characterizing it at a single point in time $t = 1$.

In subsection 2.1.3, we saw that Lévy-Kintchine formula can be written in short as $\Phi_{X_t}(s) = e^{\Psi(t,s)}$, for each $t \geq 0$ and $s \in \mathbb{R}$, where $\Psi(t, \cdot)$ is a Lévy exponent of X_t. The following theorem shows that $\Psi(t, s) = t\Psi(s)$, for each $t \geq 0$ and $s \in \mathbb{R}$.

Theorem 2.6 *If X is a Lévy process then $\mathbb{E}[e^{isX_t}] = \Phi_{X_t}(s) = e^{t\Psi(s)}$, for each $t \geq 0$ and $s \in \mathbb{R}$, where Ψ is the Lévy exponent of X_1.*

For a proof, again, we refer to David Applebeum(2004)[2]. We now have the expression of Lévy Kintchine formula for the Lévy process $X = \{X_t; t \geq 0\}$:

$$\mathbb{E}[e^{isX_t}] = \exp\left\{t\left[ias - \frac{1}{2}s^2b^2 + \int_{\mathbb{R}\setminus\{0\}} \left[e^{isx} - 1 - isx\mathbb{I}_{\{-1,1\}}(x)\right] \nu(dx)\right]\right\}. \tag{2.23}$$

Comparing (2.23) with (2.5) we observe that the former is simply a version of the latter corresponding to $t = 1$. This, together with our intuitive interpretation of Lévy measure of an infinitely divisible distribution(in subsection 2.1.3 following equation (2.17)), leads to the following more convincing definition of the Lévy measure of a Lévy process.

Definition 2.4 *[Lévy measure of a Lévy process] For a Lévy process $X = \{X_t; t \geq 0\}$ on $\mathbb{R}$, the measure ν on $\mathbb{R}$ defined by:*

$$\nu(A) = \mathbb{E}\left[\sharp\{t \in [0,1] \mid \Delta X_t \neq 0, \Delta X_t \in A\}\right], \qquad A \in \mathcal{B}(\mathbb{R}), \tag{2.24}$$

is called the Lévy measure of X. Here $\nu(A)$ is the expected number, per unit time, of the jumps with sizes in the set A.

We are now in a position to relate the mathematical fact about the Lévy measure, as discussed in subsection 2.1.3, with its definition. With this definition, the fact in second part of (2.21) ensures that along any observation period the sample paths of Lévy processes exhibit finite number of big jumps of magnitude greater than ϵ, with $0 < \epsilon \leq 1$. However the first part of (2.21), together with the above definition of Lévy measure, doesn't ensure anything about the finiteness of number of small jumps of magnitude less than ϵ, with $0 < \epsilon \leq 1$. The first part in(2.21) ensures that even if it happens to be the case that $\sum_i \Delta X_i = \infty$, we will always have $\sum_i \mid \Delta X_i \mid^2 < \infty$.

Remark 2.3 *Characteristic triplet of the Lévy process is just the characteristic triplet of the infinitely divisible random variable X_1. It can be shown that this correspondence is unique. Thus given a Lévy process there corresponds a unique IDD which is the distribution of X_1. It follows that corresponding to every infinitely divisible distribution there exists a Lévy process so that the characteristics of the process evaluated at $t = 1$ coincide with the characteristics of the IDD.*

Now let us get more insight into how characteristic triplet of a Lévy process characterizes the distribution of innovations(and hence the distribution of return itself) in asset price models on an arbitrary interval, say $[t_1, t_2]$. When Brownian motion is replaced by an arbitrary Lévy process X_t in equation (1.13)(with $\sigma = 1$ for simplicity), we obtain:

$$\frac{dS_t}{S_t} = \mu dt + dX_t \tag{2.25}$$

Integrating on $[t_1, t_2]$ we obtain an infinitely divisible random variable describing the random evolution of log returns on $[t_1, t_2]$:

$$\begin{aligned} log\left(\frac{S_{t_2}}{S_{t_1}}\right) &= \mu(t_2 - t_1) + X_{t_2} - X_{t_1} \\ &\stackrel{D}{=} \mu(t_2 - t_1) + X_{(t_2 - t_1)} \qquad \text{by [L2]-(ii) of Definition 2.1} \end{aligned} \tag{2.26}$$

According to the Lévy-Kintchine formula for Lévy process, see equation (2.23), the distribution of $X_{(t_2-t_1)}$ is characterized by a characteristic function given by:

$$\mathbb{E}[e^{isX_{(t_2-t_1)}}] = \exp\left\{(t_2 - t_1)\left[ias - \frac{1}{2}s^2b^2 + \int_{\mathbb{R}\setminus\{0\}} \left[e^{isx} - 1 - isx\mathbb{I}_{\{-1,1\}}(x)\right] \nu(dx)\right]\right\}. \tag{2.27}$$

Analogous to our intuitive development, from subsection 2.1.3, for (2.27), we can write (2.26) as:

$$\begin{aligned} log\left(\frac{S_{t_2}}{S_{t_1}}\right) &\stackrel{D}{=} \mu(t_2 - t_1) + a(t_2 - t_1) + N(0, [\sqrt{(t_2 - t_1)}b]^2) + CP\left[(t_2 - t_1)\lambda_{bj}, F_{|J|>1}(dx)\right] \\ &\quad + \lim_{\substack{i\to\infty \\ \epsilon_i \to 0}} \sum_i \left\{CP^c\left[(t_2 - t_1)\lambda_{sj}^{\epsilon_i}, F_{\epsilon_i<|J|\leq 1}(dx)\right]\right\} \\ &\stackrel{D}{=} [\mu + a](t_2 - t_1) + N(0, [\sqrt{(t_2 - t_1)}b]^2) + CP\left[(t_2 - t_1)\lambda_{bj}, F_{|J|>1}(dx)\right] \\ &\quad + \lim_{\substack{i\to\infty \\ \epsilon_i \to 0}} \sum_i \left\{CP^c\left[(t_2 - t_1)\lambda_{sj}^{\epsilon_i}, F_{\epsilon_i<|J|\leq 1}(dx)\right]\right\} \end{aligned} \tag{2.28}$$

The distributions and rates , e.g. λ_{bj}, $F_{|J|>1}(dx)$ and $\lambda_{sj}^{\epsilon_i}$, $F_{\epsilon_i<|J|\leq 1}(dx)$, for each i, are defined in subsection2.1.3. As we saw in (2.11), in terms of distribution the compensation in last term ensures that the resulting process from the superposition of all the processes involved(in last term) is an square integrable martingale. A resonably large i with a reasonably small ϵ_i, can lead to a good approximate modelling tool with a large number of compensated compound Poisson processes. We can obtain expected total number of jumps of all sizes on $[t_2, t_1]$:

$$\begin{aligned}(t_2 - t_1)\lambda &= (t_2 - t_1)\lambda_{bj} + (t_2 - t_1)\left\{\lambda_{sj}^{\epsilon_1} + \lambda_{sj}^{\epsilon_2} + \cdots + \lambda_{sj}^{\epsilon_i} + \cdots\right\} \\ &= (t_2 - t_1)\lambda_{bj} + (t_2 - t_1)\lambda_{sj} \\ &= (t_2 - t_1)\left\{\nu(|\, x\,| > 1) + \nu(|\, x\,| \leq 1)\right\} \\ &= (t_2 - t_1)\nu\left(\{|\, x\,| > 1\} \cup \{|\, x\,| \leq 1\}\right) \\ &= (t_2 - t_1)\nu(\mathbb{R}). \end{aligned} \qquad (2.29)$$

This is in agreement with the definition 2.4 of Lévy measure of a Lévy process. That is considering jumps of all sizes, in other words the Borel set being $\mathbb{R}$, the Lévy measure is the expected number of jumps per unit time. Here λ_{bj} and λ_{sj} are as appeared in subsection 2.1.3.

So what equation (2.28) is all about? **In terms of distribution** it tells us that the random variable describing the log returns on $[t_2, t_1]$ is the sum of a constant and three different types of, but possibly infinitely many, random variables. The underlying Lévy measure ensures that the infinite(possibly) sum of compensated compound Poisson random variables converges to a compensated square integrable random variable as shown in (2.11). So the distribution of log-returns on $[t_2, t_1]$ is the convolution of the respective distributions of the summand random variables(considering the limiting one for the infinite sum). The characteristic function (2.27) characterizes such a convoluted distribution. However important idea is **in terms of sample paths**. Equation (2.28) describes that each sample path of log returns on $[t_2, t_1]$ is the superposition of four different types of but infinitely(possibly) many sample paths on $[t_2, t_1]$: a linear drift path, a Brownian motion path, a compound Poisson process path of big jumps and infinitely(possibly) many compensated compound Poisson process paths of small jumps. The infinite(possibly) superposition of sample paths of compensated compound Poisson processes converges to a sample paths of square integrable martingale. So now one can see the significance of replacing the Brownian motion B_t by a Lévy process X_t in (2.25). With Lévy process it can model the jumps of all sizes in the return paths, where as with Brownian motion it can only model the continuous evolution of returns.

2.2.1 An Example

Consider a compound poisson process with Poisson rate λ and jump sizes following a normal distribution with density:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{2.30}$$

Let us consider jumps with sizes in $A = (-\infty, 5]$. Then the compound Poisson process has the Lévy measure:

$$\nu(A) = \lambda \int_{-\infty}^{5} f(x)dx = \frac{\lambda}{\sqrt{2\pi\sigma}} \int_{-\infty}^{5} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \lambda N\left(\frac{5-\mu}{\sigma}\right) \tag{2.31}$$

Thus in general Lévy measure of jumps of sizes belonging to $(-\infty, k]$, for $k > -\infty$, is $\nu(\{(-\infty, k)\}) = \lambda N\left(\frac{k-\mu}{\sigma}\right)$, which we know is the expected number of jumps, per unit time, with sizes in $(-\infty, k]$. Clearly for all jump sizes $\nu(\{(-\infty, \infty)\}) = \lambda N\left(\frac{\infty-\mu}{\sigma}\right) = \lambda$, so λ is the average rate of jumps of the process. Since $\nu(\mathbb{R}) = \lambda < \infty$, this implies that the number of jumps on any time interval is finite, so expected number of jumps per unit time is finite. Here we use the fact that $N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}z^2} dz$. Thus

$$\begin{aligned} \lambda_{sj} &= \lambda \int_{-1}^{1} f(x)dx \\ &= \frac{\lambda}{\sqrt{2\pi\sigma}} \int_{-1}^{1} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \end{aligned} \tag{2.32}$$

which, for a given set of parameters λ, μ, σ, is a number. Similarly:

$$\begin{aligned} \lambda_{bj} &= \lambda \int_{|x|>1} f(x)dx \\ &= \frac{\lambda}{\sqrt{2\pi\sigma}} \left(1 - \int_{-1}^{1} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx\right) \end{aligned} \tag{2.33}$$

Distribution of big jumps can be obtained from (2.9). For example the probability of jumps in (1.5, 2) is:

$$\begin{aligned}
F_{|J|>1}(\{(1.5,2)\}) &= \frac{\int_{1.5}^{2}\nu(dx)\mathbb{I}_{|x|>1}}{\lambda_{bj}} \\
&= \frac{\frac{\lambda}{\sqrt{2\pi}\sigma}\int_{1.5}^{2} e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx}{\lambda\int_{|x|>1}F(dx)} \\
&= \frac{\frac{1}{\sqrt{2\pi}\sigma}\int_{1.5}^{2} e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx}{1-\frac{1}{\sqrt{2\pi}\sigma}\int_{-1}^{1} e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx} \qquad (2.34)
\end{aligned}$$

For a general summand under the sum in (2.28) the rate and distribution are respectively given by (2.12) and (2.13). For example for $\epsilon_n = 0.05$ the rate and the probability of jump sizes in $[0.07, 1]$ from the corresponding distribution are:

$$\begin{aligned}
\lambda_{sj}^{0.05} &= \lambda\int_{0.05<|x|\leq 1}F(dx) \\
&= \frac{\lambda}{\sqrt{2\pi}\sigma}\left(\int_{0.05}^{1} e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx\right) \qquad (2.35)
\end{aligned}$$

$$\begin{aligned}
F_{0.05<|J|\leq 1}(\{[0.07,1]\}) &= \frac{\int_{0.07}^{1}\nu(dx)\mathbb{I}_{|x|\leq 1}}{\lambda_{sj}^{0.05}} \\
&= \frac{\frac{1}{\sqrt{2\pi}\sigma}\int_{0.07}^{1} e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx}{\frac{1}{\sqrt{2\pi}\sigma}\int_{0.05}^{1} e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx} \\
&= \frac{\int_{0.07}^{1} e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx}{\int_{0.05}^{1} e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx} \qquad (2.36)
\end{aligned}$$

These kind of values for each summand, together with $F_J(x) \sim N(\mu, \sigma^2)$, can be used in (2.28) to simulate jumps in sample paths of log return on $[t_1, t_2]$(or on any other interval).

As discussed in the above example Lévy measure helps us decide whether the number of jumps of the underlying process is finite or infinite. It is in fact a general property of Lévy process:

Proposition 2.5 *Let X_t is a Lévy process with X_1 having characteristic triplet (a, b^2, ν). Then:*

- *if $\nu(\mathbb{R}) < \infty$ then almost all paths of X_t have a finite number of jumps on every compact interval. In that case the Lévy process is said to be of finite activity.*

- *if $\nu(\mathbb{R}) = \infty$ then almost all paths of X_t have an infinite number of jumps on every compact interval. In that case the Lévy process is said to be of infinite activity.*

A proof can be found e.g. in Sato(1999)[100].

Before we rap this section up, we need to further emphasize the usefulness of LKF for our study. In one sense LKF is the cornerstone of the entire thesis. Apart from it's intuitive appeal in revealing the path structure of general Lévy process, as explained in section2.1.3, LKF has more significant use in option pricing and risk management. It provides the characterization of a particular Lévy process through it's characteristic function. We then have the elementary result in Statistics which ensures the existence of a unique distribution function corresponding to a characteristic function. Since such a distribution is the fundamental tool used in estimation of risk measures, we revisit the technique of extracting the distribution from the characteristic function. This technique is known as Fourier inversion. Moreover the characteristic function of a model can be utilized to obtain the prices of European style derivatives, a technique frequently referred as Carr-Madan formula, see Carr and Madan(1999)[27].

We can obtain the probability density $f(x)$ by inverting the characteristic function, see Whit and Abate(1992)[113]:

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iux}\Phi(u)du \tag{2.37}$$

The probabilities can be expressed using Fourier integral theorem. For example:

$$\begin{aligned} P(a < X < b) &= \int_a^b f(x)dx \\ &= \frac{1}{2\pi}\int_a^b \left\{\int_{-\infty}^{\infty} e^{-iux}\Phi(u)du\right\} dx \end{aligned} \tag{2.38}$$

Assuming the continuity of the underlying distribution, the Fubini's theorem allow's us to interchange the integrals:

$$P(a < X < b) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left\{\frac{e^{-iub}}{-iu} - \frac{e^{-iua}}{-iu}\right\}\Phi(u)du \tag{2.39}$$

A more simplified form of such a probability is given by Gil-Pelaez(1951)[66]:

$$\begin{aligned} P(X \le a) &= \frac{1}{2} + \frac{1}{2\pi}\int_0^{\infty} \frac{e^{iua}\Phi(-u) - e^{-iua}\Phi(u)}{iu}du \\ &= \frac{1}{2} + \frac{1}{2\pi}\int_0^{\infty}\frac{e^{iua}\Phi(-u)}{iu}du - \frac{1}{2\pi}\int_0^{\infty}\frac{e^{-iua}\Phi(u)}{iu}du \end{aligned} \tag{2.40}$$

We realize that for an arbitrary complex number z, $z + \bar{z} = 2Re[z]$. We know the characteristic function is a hermitian function for which $\Phi(-u) = \overline{\Phi(u)}$.

$$
\begin{aligned}
& \frac{1}{2\pi}\int_0^\infty \frac{e^{iua}\Phi(-u)}{iu}du - \frac{1}{2\pi}\int_0^\infty \frac{e^{-iua}\Phi(u)}{iu}du \\
= & \frac{1}{2\pi}\int_0^\infty \frac{\overline{e^{-iua}\Phi(u)}}{-iu}du - \frac{1}{2\pi}\int_0^\infty \frac{e^{-iua}\Phi(u)}{iu}du \\
= & -\frac{1}{2\pi}\int_0^\infty \left\{ \frac{e^{-iua}\Phi(u)}{iu} + \frac{\overline{e^{-iua}\Phi(u)}}{iu} \right\} du \\
= & -\frac{1}{\pi}\int_0^\infty Re\left[\frac{e^{-iua}\Phi(u)}{iu}\right] du \qquad \text{since } z + \bar{z} = 2Re[z]. \qquad (2.41)
\end{aligned}
$$

Using equation (2.41) into equation (2.40), we obtain the particular type of probability which we will use in our pricing:

$$
\begin{aligned}
P(X > a) &= 1 - P(X \leq a) \\
&= \frac{1}{2} + \frac{1}{\pi}\int_0^\infty Re\left[\frac{e^{-iua}\Phi(u)}{iu}\right] du. \qquad (2.42)
\end{aligned}
$$

Thus probabilities can be extracted no matter how involved the expression of the characteristic function is. Only requirement is to calibrate the parameters of the characteristic function using market data before applying (2.42).

2.3 LKF: Jump Processes for Non-normality

Theorem 2.5 shows that for a Lévy process X_t, infinite divisibility allows us to write the random variable X_t(for each fixed $t > 0$) as the sum of n independent and identically distributed random variables. Remembering Central limit theorem, CLT henceforth, the infinite divisibility is the well-cited motivation for modeling stock return by the Gaussian distribution, namely that this distribution is the limiting distribution of sum of n independent random variables , upon some scaling, which usually represents the effects of various shocks in the economy. More precisely, Lévy processes are premised on similar argument of accommodating infinite economic shocks in the returns but not necessarily confining within the world of Gaussianity and continuous paths of return dynamics. Lévy-Kintchine representation ensures the infinite divisibility of the distribution of the underlying process at each$t > 0$ but at the same time it lays a foundation where we can see a general return dynamics as the superposition of different processes;such that limiting distribution needn't necessarily be Gaussian and can have jumps, of both finite and infinite number, in the path. So this is easy to see how powerful this class of processes for modeling non-normality in option pricing and addressing various other imperfections surrounding the benchmark Black-Schole's model. More interestingly we will now see that LKF implies that deviation from normality, i.e. from the Brownian paths, makes it mandatory to consider jumpy paths for the return dynamics.

We now revisit random measure oriented structural properties of Lévy processes; this revisit is required to see another intuitive form of LKF. Also the ideas go into next sections where subordinators and time changing are considered as tools to report and correct a misspecification in a classic work of Geman(2002)[62]; which is the complementary contribution of this chapter. As discussed in proposition2.5 the infinite activity Lévy process can exhibit infinite number of jumps per unit time along every sample path. At the same time properties of Lévy measure ensure that the number of big jumps per unit time is finite along every sample path. So infinite activity Lévy process ensures infinitely many small jumps per unit time along sample paths. This indicates that the small movements along the paths are so frequent that it adequately allows us to exclude the necessity of considering an additional and unrelated diffusion component. Thus along with finitely many big jumps(jumps larger than ϵ in magnitude, for a very small ϵ) the small jumps locally attribute to diffusion. And the continuity requirement of the diffusion process forces the rate of local arrival of jumps of all sizes to zero thus reduces the local variation of uncertainty in the price dimension to be explained with a single *instantaneous volatility* parameter. That is where the words pure jump find the justification: they are mutually exclusive of diffusion processes.

In our intuitive development in section 2.1.3 we saw that any Lévy process can be expressed as the sum of three processes: (i) a Brownian motion with drift, (ii) a compound Poisson process of big jumps and (iii) a limiting process of compensated compound Poisson processes of small jumps. Each of these processes being semimartingale, see Kyprianou(2006)[76] and Shiryaev(1999)[106], the superimposed resultant processes is again a semimartingale. Thus any Lévy process is a semimartingale. We know stock price processes have to be semimartingale under real probability measure and Lévy processes appear as a wide natural class of candidates for stock price dynamics. Now clearly continuity of trajectories requires the components corresponding to the processes (ii) and (iii) above to be zero. According to our development in subsection 2.1.3 and Lévy-Kintchine formula of a general Lévy process (2.23) , this means that:

$$X_t = at + bB_t.$$

Thus we arrive to an important conclusion: *The only Lévy process with the continuous paths is the Brownian motion (with drift).* The consequence is that if we use Lévy process to describe the return or natural log of stock we obtain normality together with continuity. In other words if the data exhibits deviation from normality we have no continuous process left for modelling (as Brownian motion is the only continuous Lévy process). **So non-normality needs to be modeled using discontinuous(jumpy) Lévy processes.** Furthermore to obtain a finite quadratic variation(which is a better representation of stock price dynamics) the diffusion component must be zero and the process must be a pure jump Lévy process.

2.4 Random Measure of Jumps

We need this mathematical section for intuitive development in sections to follow. Since a Lévy process is cadlag, the number of jumps ΔX_s such that $|\Delta X_s| \geq \epsilon$, before some time t, has to be finite for all $\epsilon > 0$. Hence if $B \in \mathcal{B}(\mathbb{R})$, is bounded away from 0 (i.e. $0 \notin \bar{B}$, the closure of B), then for $t \geq 0$

$$N_t^B = \sharp\{s \in [0,t] : \Delta X_s \in B\} = J_X([0,t] \times B) \tag{2.43}$$

is well defined and a.s. finite. The process N^B is clearly a counting process, called a counting process of B . It inherits the Lévy properties from X . Since the Poisson process is the only non-trivial counting process which is Lévy then N_t^B is a Poisson process with a certain intensity $\nu^X(B) < \infty$. If B is a disjoint union of Borel sets B_i, then $N_t^B = \sum_i N_t^{B_i}$. Hence considering (2.24), $\nu^X(B) = \mathbb{E}N_1^B = \sum \mathbb{E}N_1^{B_i} = \sum \nu(B_i)$. ν is a Borel measure and as indicated in (2.20) it holds that $\nu(\mathbb{R}\setminus(-\epsilon,\epsilon)) < \infty$ for all $\epsilon > 0$. In particular ν is $\sigma-$finite.

Now to discuss the term $J_X([0,t] \times B)$ in (2.43) we need the idea of Poisson random measure. However we keep our discussion about this measure at an introductory level. For more details see Bertoin(1996)[15] and Applebeum(2004)[2].

From the definition of the Poisson process we recall that the jump times $T_1, T_2 \cdots$ form a random configuration of points on $[0, \infty)$ and that the Poisson process N_t counts the number of such points in the interval $[0, t]$ see Cont and Tankov(2004)[38].This counting procedure defines a measure on $[0, \infty)$.

Definition 2.5 *For any measurable set $A \subset \mathbb{R}^+$ a positive integer valued counting measure $M(\omega, \cdot)$ defined as*

$$M(\omega, A) = \sharp\{i \geq 1; T_i(\omega) \in A\}, \qquad \omega \in \Omega, \tag{2.44}$$

is a random measure.

The very first property of a Poisson process ensures that $M(A)$, for any bounded measurable set A, is almost surely finite. The intensity λ of the Poisson process determines the average value of the random measure M, i.e. $\mathbb{E}[M(A)] = \lambda|A|$ where $|A|$ is the Lebesgue measure of A. M is also known as a random jump measure associated to the Poisson process N. The Poisson process can be expressed in terms of the random measure M in the following way:

$$N_t(\omega) = M(\omega, [0, t]) = \int_{[0,t]} M(\omega, ds).$$

The properties of the Poisson process , see Cont and Tankov(2004)[38], can be translated into properties of the measure M. Some of the important ones are as follows.

[1] For disjoint intervals $[t_1, t_1^{'}], \cdots, [t_n, t_n^{'}]$, $M([t_k, t_k^{'}])$ is the number of jumps of the Poisson process in $[t_k, t_k^{'}]$. It is a Poisson random variable with parameter $\lambda(t_k^{'} - t_k)$. Generally for any measurable set A, $M(A)$ follows a Poisson distribution with parameter $\lambda|A|$, where $|A| = \int_A dx$ is the Lebesgue measure of A.

[2] For two disjoint intervals $[t_i, t_i^{'}]$ and $[t_j, t_j^{'}]$ where $i \neq j$, $M([t_i, t_i^{'}])$ and $M([t_j, t_j^{'}])$ are independent random variables.

A natural extension of this notion of random measure is the Poisson random measure ,where $\mathbb{R}^+$ is replaced by any $E \subset \mathbb{R}$ and the Lebesgue measure by any Radon measure μ on E.

Definition 2.6 *Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space, $E \subset \mathbb{R}$ and μ a given positive Radon measure on $(E, \mathcal{E})$. A Poisson random measure on E with intensity measure μ is an integer valued random measure*

$$M : \Omega \times E \to \mathbb{N}, \quad (\omega, A) \mapsto M(\omega, A)$$

such that:

[1] for almost all $\omega \in \Omega, M(\omega, \cdot)$ is an integer valued Radon measure on E, i.e. for any bounded measurable $A \subset E$, $M(A) < \infty$ is an integer valued random variable.

[2] for each measurable set $A \subset E, M(\cdot, A) = M(A)$ is a Poisson random variable with parameter $\mu(A)$, i.e.

$$\mathbf{P}\{M(A) = k\} = e^{-\mu(A)} \frac{[\mu(A)]^k}{k!}, \qquad \forall k \in \mathbb{N}. \tag{2.45}$$

[3] for disjoint measurable sets $A_1, \cdots, A_n \in \mathcal{E}$, the corresponding random variables $M(A_1), \cdots, M(A_n)$ are independent.

We state the following proposition, without proof, which ensures the existent of the Poisson random measure.

Proposition 2.6 *For any Radon measure μ on $E \subset \mathbb{R}$, there exists a Poisson random measure M on E with intensity μ.*

For a proof see see Cont and Tankov(2004)[38].

It can be shown that to every cadlag process and in particular to every compound Poisson process $X = \{X_t; t \geq 0\}$ on $\mathbb{R}$ we can associate a random measure on $\mathbb{R} \times [0, \infty)$ describing the jumps of X. For every measurable set $B \subset \mathbb{R} \times [0, \infty)$

$$J_X(B) = J_X(A \times [t_1, t_1']) = \sharp\{t \in [t_1, t_1']; \Delta X_t \in A\}, \tag{2.46}$$

that is $J_X(A \times [t_1, t_1'])$:=number of jumps in X occurring between time t_1 to t_1' whose amplitude belongs to A.

The random measure J_X contains all information about the discontinuities (jumps) of X. It tells us when the jumps occur and how big they are. J_X does not give any information regarding the continuous part of X. It is easy to see that X has continuous sample paths if and only if $J_X = 0$ a.s. which implies that there are no jumps.

The following proposition shows that J_X is a Poisson random measure in the sense defined above.

Proposition 2.7 *Let $X = \{X_t; t \geq 0\}$ be a compound Poisson process with intensity λ and jump size distribution F. Its jump measure J_X is a Poisson random measure on $\mathbb{R} \times [0, \infty)$ with intensity measure $\mu(dx \times dt) = \nu(dx)dt = \lambda dF(x)dt$.*

For a proof see [38].

Equation (2.45) implies that, $J_X(B)$ as defined in (2.46) satisfies:

$$\mathbb{E}[J(dx \times dt)] = \mu(dx \times dt) = \nu(dx)dt = \lambda dF(x)dt. \tag{2.47}$$

Equation (2.47) bears important intuition. It extends the intuition of the definition 2.4 of Lévy measure, as the expected number of jumps per unit time with jump sizes in a Borel set $A \in \mathcal{B}(\mathbb{R})$, to the arrival rates of jump sizes in a spatial domain. More precisely if $B \subset \mathbb{R} \times [0, \infty)$ is of the form $A \times [t_2, t_1]$, (2.47) means that integration of the Lévy density over this spatial domain provides the arrival rate of jumps in this domain.

We are now in a position to relate our intuitive development in subsection 2.1.3 to Lévy-Ito decomposition. Lévy-Ito decomposition states that every Lévy process X_t can be decomposed as:

$$X_t = at + B_t + X_t^l + \lim_{\epsilon \longrightarrow 0} \tilde{X}_t^\epsilon, \qquad t \geq 0, \tag{2.48}$$

where

$$X_t^l = \int_{\substack{s \in [0,t] \\ |x| > 1}} x J_X(ds \times dx) = \sum_{\substack{0 \leq s \leq t \\ |\Delta X_s| \geq 1}} \Delta X_s$$

corresponds to discontinuous large jump process and

$$\begin{aligned} \tilde{X}_t^\epsilon &= \int_{\substack{s \in [0,t] \\ \epsilon \leq |x| \leq 1}} x\{J_X(ds \times dx) - \nu(dx)ds\} \\ &= \int_{\substack{s \in [0,t] \\ \epsilon \leq |x| \leq 1}} x\{\tilde{J}_X(ds \times dx)\} \end{aligned}$$

corresponds to compensated small jump process.Clarly equation(2.48) bears the same intuition as we developed in subsection2.1.3 for $X_t = log(\frac{S_{t+1}}{S_t})$.

2.5 Subordinator and Its Application in Finance

Subordinators are essential tools to study the time changed processes in Finance. The definition as well as mathematical characterization of Subordinators rely on the idea of total variation of the process, among others.

Definition 2.7 (Bounded and unbounded variation) *The total variation of a right continuous function with left limit is defined as:*

$$\| f \|_{TV} := \sup \Big\{ \sum_{j=1}^{j=n} | f(t_j) - f(t_{j-1}) | ; 0 = t_0 < t_1 < \cdots < t_n = t, \ n \in \mathbb{N} \Big\}.$$

Clearly for an increasing function on $[0, t]$ *with* $f(0) = 0$ *this is just* $f(t)$ *and for a difference* $f = g - h$ *of two increasing functions with* $f(0) = g(0) = 0$ *the total variation is at most* $g(t) + h(t) < \infty$. *We will see that finite variation in one of the recent option pricing model, namely Variance Gamma, is a direct consequence of this fact. Such functions are known as functions of finite or bounded variation. A finite variation process is one such that each of its sample paths are of finite variation.*

In case the total variation of a function is infinite, the function is known as "of unbounded variation". Brownian motion, which is the only Lévy process with continuous sample paths, has unbounded variation though it has a finite quadratic variation.

The variation of a Lévy process is completely characterized as:

Proposition 2.8 *Let* X_t *is a Lévy process with* X_1 *having characteristic triplet* (a, b^2, ν). *Then:*

- *if* $b^2 = 0$ *and* $\int_{|x| \le 1} | x | \nu(dx) < \infty$ *then almost all paths of* X_t *have finite variation. The converse is also true.*
- *if* $b^2 \neq 0$ *and* $\int_{|x| \le 1} | x | \nu(dx) = \infty$ *then almost all paths of* X_t *have unbounded variation. The converse is also true.*

For a proof we refer to Sato(1999)[100]. The first part of this proposition simply tells us that if the underlying process is not diffusive and if mathematical structure of the process guarantees that the total contribution of the absolute movements caused by small jumps will be finite than almost all the sample paths of the process are guaranteed to be of finite variation. The second part of this proposition is rather more convincing. It tells us that if the paths of a process are observed to be of unbounded variation than there must be diffusion and absolute contribution of the small jumps will be infinite. This proposition implicitly implies that diffusion and total absolute contribution of the small jumps are closely associated with each other; which goes with our intuition. However in practice the compensated(mean subtracted) compound Poisson part in(2.28) yields the ground where unbounded variation from small jumps becomes less likely.

In concise form this implies:

$$\begin{aligned}\int_0^{\infty}(1\wedge \mid x\mid)\nu(dx) &= \int_0^1 \mid x\mid \nu(dx)+\int_1^{\infty}\nu(dx)\\ &= \begin{cases}\text{finite} & \text{if } X_t \text{ is of finite variation(so } b^2=0).\\ \infty & \text{if } X_t \text{ is of unbounded variation.}\end{cases}\end{aligned} \tag{2.49}$$

Since by the property of Lévy measure, see (2.21), $\int_1^{\infty}\nu(dx)<\infty$, unbounded variation can be seen as resulting from diffusion and small jumps.

Definition 2.8 (Subordinator) *Let $\{X_t; t\geq 0\}$ be a Lévy process such that X_1 has the Lévy triplet (a,b^2,ν). Then X_t is an increasing process in t if and only if $\nu(-\infty,0]=0$, $b^2=0$ $\int_0^1 x\nu(dx)<\infty$ and $d=a-\int_0^1 x\nu(dx)>0$. Such an increasing process is known as Subordinator.*

In this case X_t can be expressed as the sum of its jumps over times 0 to t and linear drift:

$$X_t = dt+\int_{[0,t]\times\mathbb{R}} xJ_X(ds\times dx) = dt+\sum_{\substack{0\leq s\leq t\\ |\Delta X_s|\geq 1}}\Delta X_s, \qquad t\geq 0, \tag{2.50}$$

and its characteristic function can be expressed as

$$\mathbb{E}\big[e^{isX_t}\big]=\exp\Big[t\big\{ids+\int(e^{isx}-1)\nu(dx)\big\}\Big] \tag{2.51}$$

where $d=a-\int_{|x|\leq 1}x\,\nu(dx)$.

The idea is that in case of Subordinator, jumps are the only source of randomness and finite variation ensures that small jumps are integrable. So usual compensation of small jumps in Lévy-Kintchine formula to ensure the integrability of Lévy measure is not needed and the compensation part can be adjusted with the drift of the process yielding a new drift.

We close this section with a proposition, putting all these facts together, which completely characterizes a Subordinator:

Proposition 2.9 *Let $\{X_t; t \geq 0\}$ be a Lévy process on $\mathbb{R}$. The following conditions are equivalent:*

[i] $X_t \geq 0$ *a.s. for some* $t > 0$.

[ii] $X_t \geq 0$ *a.s. for every* $t > 0$.

[iii] Sample paths of X_t are almost surely non-decreasing:

$$t \geq \acute{t} \qquad \Rightarrow \qquad X_t \geq X_{\acute{t}} \quad a.s.$$

[iv] The characteristic triplet of X_t satisfies $b^2 = 0$, $\nu(-\infty, 0] = 0$, $\int_0^\infty (1 \wedge x)\nu(dx) < \infty$ and $d = a - \int_{|x|\leq 1} x\nu(dx) \geq 0$, that is X_t has no diffusion component, only positive jumps of finite variation and positive drift.

For a proof see Cont and Tankov(2004) [38].

2.5.1 Time Change Through Subordinator

The main application of subordinator in finance is the so called "time change", i.e. to model the change from "calender time" to "business time". Such time axis modeling by a positive Lévy process(subordinators) has intuitive explanation in Finance. See Geman(2002)[62].

As we explained in section2.5 subordinator can only display positive jumps in positive direction. Thus in case of subordinator drift being positive and there being no negative jump, the diffusion component needs to be zero(otherwise there could be a negative change with positive probability). Hence positive jumps, positivity being required for time changing, are the only source of randomness and finite variation ensures that small jumps are summable. So usual compensation of small jumps in Lévy-Kintchine formula to ensure the integrability of Lévy measure is not needed and the compensation part can be adjusted with the drift of the process to give a new drift. See Cont and Tankov()2004[38], Sato(1999)[100] for details. The following theorem is very important in financial application. It shows that when a Lévy process(modelling return dynamics) is subordinated by a subordinator(modelling time change) the resulting process is still a Lévy process. Moreover it shows how to get the characteristics of the resulting process. The proof of the theorem can be found in Sato(1999)[100].

Theorem 2.7 *Let T_t be a subordinator with Lévy measure $\tilde{\nu}$, drift d . Its distribution at time t, P_{T_t} is characterized by the equation (2.51) and let $\lambda = P_{T_1}$. Further assume X_t is a $\mathbb{R}-$valued Lévy process with Lévy triplet (a, b^2, ν). Its distribution P_{X_t}, at $t > 0$ is characterized by equation (2.23) and let $\mu = P_{X_1}$. Then provided the processes X_t and T_t are independent, the process defined as*

$$Y_t(\omega) = X_{T_t(\omega)}(\omega) \quad t \geq 0; \tag{2.52}$$

is also a Lévy process. The distribution of Y_t is given by:

$$P[Y_t \in B] = \int_0^\infty \mu^s(B)\lambda^t(ds), \;\; B \in \mathfrak{B}(\mathbb{R}). \tag{2.53}$$

The Lévy triplet (a_y, b_y^2, ν_y) of Y_t is given by:

$$a_y = d.a + \int_0^\infty \tilde{\nu}(ds) \int_{|x|\leq 1} x\mu^s(dx), \tag{2.54}$$

$$b_y^2 = d.b^2, \tag{2.55}$$

$$\nu_y(B) = d.\nu(B) + \int_0^\infty \mu^s(B)\tilde{\nu}(ds), \quad B \in \mathfrak{B}(\mathbb{R} \setminus \{0\}). \tag{2.56}$$

Remark 2.4 *We will see in details how to make use of each of these equations. We must mention that all the time changed Lévy processes in finance have to be analytically developed based on the above theorem. We would like to see such analytic development for one of the successful time changed Lévy models in finance.*

2.5.2 Analysis of Variance-Gamma Process

Variance gamma process was first introduced by Madan and Senata(1990)[79]. Subsequently it was adapted to option pricing by Madan et al(1998)[80]. Brownian motion was time changed by an increasing gamma process. We present the detailed derivations in VG model as a Brownian motion time changed by gamma subordinator. These derivations reveal the stochastic intuitions behind time changed processes. Furthermore we realize that Geman (2002)[62] had misspecified the expressions of some parameters in the paper when discussing variance gamma process. Since Geman (2002)[62] didn't go for the details derivation of the Lévy measure of variance gamma process Geman (2002)[62] even didn't realize that the specification of the parameters do not yield the correct expression of the Lévy measure used in the paper. We will go for the detailed derivations with the correct specification of those parameters and will show how our specification of the parameters yields the expression of the Lévy measure used in the work Geman (2002)[62]. This is the reason why we considered VG model for illustration. However other Lévy processes also have the facets of time changing embedded. e.g. Normal Inverse Gaussian(NIG) process has the mathematical treatment as Brownian motion time changed by IG subordinator; CGMY process has the mathematical illustration as Brownian motion time changed by tempered stable(TS) subordinators.

We saw in example 2.2 that gamma random variable is infinitely divisible. Thus according to the Lévy-Kintnchine formula of Lévy processes we can attach the gamma variable to get a gamma process such that:

$$T_{t+s} - T_t = \Gamma\left(\frac{s}{\gamma}, \frac{1}{\gamma}\right) \text{ in general } T_t = \Gamma\left(\frac{t}{\gamma}, \frac{1}{\gamma}\right) \tag{2.57}$$

where $\Gamma(\alpha, \beta)$ has the density:

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{\beta x}; \;\; x > 0, \; \alpha > 0, \; \beta > 0. \tag{2.58}$$

We now show that gamma process is a subordinator.

Lemma 2.1 *The generating triplet for the* $\Gamma(\alpha, \beta)$ *distribution is* $(0, 0, \nu^s)$, *where Lévy measure* ν^s *is given by:*

$$\nu^s(dx) = \frac{\alpha}{x} e^{-\beta x} dx, \; x > 0. \tag{2.59}$$

It then follows that gamma process T_t, with $\alpha = \frac{1}{\gamma}$ and $\beta = \frac{1}{\gamma}$, in (2.57) is a subordinator with triplet $(0, 0, t\nu^s)$.

Proof. If F is the probability measure with density (2.58), then (2.1) shows that:

$$\Phi_F(s) = \mathbb{E}[e^{isx}] = \left[\frac{\beta}{\beta - is}\right]^\alpha = \left[\frac{\beta - is}{\beta}\right]^{-\alpha} \tag{2.60}$$

However:

$$\begin{aligned}
log\left[\frac{\beta-is}{\beta}\right] &= \int_0^{is}\frac{-dy}{\beta-y}\\
&= -\int_0^{is}\left[\int_0^{\infty}e^{-\beta x+yx}dx\right]dy\\
&= -\int_0^{\infty}e^{-\beta x}\left[\int_0^{is}e^{yx}dy\right]dx\\
&= -\int_0^{\infty}\frac{e^{-\beta x}}{x}\left[e^{isx}-1\right]dx
\end{aligned}$$

So:

$$\begin{aligned}
\Phi_F(s) &= \left[\frac{\beta-is}{\beta}\right]^{-\alpha}\\
&= \exp\left[\int_0^{\infty}\left[e^{isx}-1\right]\frac{\alpha}{x}e^{-\beta x}dx\right]
\end{aligned} \tag{2.61}$$

Comparing (2.61) with (2.23) the result follows. □

The following results, about Bessel function of third kind, will be used in next derivation.

Lemma 2.2 *If K_p is the modified Bessel function of third kind, then:*

$$K_p(x)=\frac{1}{2}\left(\frac{x}{2}\right)^p\int_0^{\infty}e^{-t-\frac{x^2}{4t}}t^{-p-1}dt;\ \ x>0,p\in\mathbb{R}. \tag{2.62}$$

$$K_{n+\frac{1}{2}}(x)=\sqrt{\left(\frac{\pi}{2}\right)}x^{-\frac{1}{2}}e^{-x}\left(1+\sum_{i=1}^{n}\frac{(n+i)!}{(n-i)!i!}(2x)^{-i}\right) \tag{2.63}$$

Proof's can be found in Watson(1944)[112].

Consider the process $X_t=\sigma B_t+\theta t$, where B_t is a standard Brownian motion and $\sigma>0,\theta\in\mathbb{R}$ are volatility and drift parameters,respectively. The Variance Gamma, VG henceforth, process is defined as the process Y_t subordinated to X by the $\Gamma-$subordinator T:

$$Y_t:=X_{T_t}=\sigma B_{T_t}+\theta T_t \tag{2.64}$$

Gamma process is characterized as in (2.57) so that it ensures mean rate t and variance γt with the probability density:

$$f_{T_t}(x)=\frac{\left(\frac{1}{\gamma}\right)^{\frac{t}{\gamma}}}{\Gamma\left(\frac{t}{\gamma}\right)}x^{\frac{t}{\gamma}-1}e^{-\frac{x}{\gamma}} \tag{2.65}$$

With this parametrization, (2.60) ensures that the Laplace transform of this gamma subordinator is:

$$\mathbb{E}[e^{-sT_t}] = (1+s\gamma)^{-\frac{t}{\gamma}}. \tag{2.66}$$

Equation (2.64) shows that conditional on a jump of size $T_t = s$ in the time axis, the move of the process Y_t is normally distributed with mean θs and variance $\sigma^2 s$. Applying conditioning we can now compute the characteristic function of compound random variable (here normal compounded by gamma):

$$\begin{aligned}
\Phi_{Y_t}(s) &= \mathbb{E}[e^{is(\sigma B_{T_t}+\theta T_t)}] \\
&= \int_0^\infty \mathbb{E}\left[e^{is(\sigma B_{T_t}+\theta T_t)} \mid T_t = u\right] f_{T_t}(u)du \\
&= \int_0^\infty e^{is\theta u - \frac{1}{2}s^2\sigma^2 u} f_{T_t}(u)du \\
&= \int_0^\infty e^{-[\frac{1}{2}s^2\sigma^2 - is\theta]u} f_{T_t}(u)du \\
&= \left(1+\left[\frac{1}{2}s^2\sigma^2 - is\theta\right]\gamma\right)^{-\frac{t}{\gamma}} \qquad \text{[using (2.66)]} \\
&= \left(\frac{1}{1-is\theta\gamma+\frac{1}{2}s^2\sigma^2\gamma}\right)^{\frac{t}{\gamma}}
\end{aligned} \tag{2.67}$$

We now use theorem 2.7 to obtain the Lévy triplet of VG process.

From (2.54):

$$\begin{aligned}
a_{vg} &= \text{[drift subordinator].[drift subordinate]} + \int_0^\infty \tilde{\nu}(dx). \int_{-1}^1 y P^s_{\sigma B_1+\theta}(dy). \\
&= 0.\text{[drift subordinate]} + \int_0^\infty \frac{\left(\frac{1}{\gamma}\right)}{x} e^{-\frac{x}{\gamma}} dx. \int_{-1}^1 y P^s_{\sigma B_1+\theta}(dy). \quad \text{[using Lemma2.1]} \\
&= \int_0^\infty \frac{\left(\frac{1}{\gamma}\right)}{x} e^{-\frac{x}{\gamma}} dx. \int_{-1}^1 y \frac{1}{\sigma\sqrt{2\pi s}} e^{-\frac{(y-\theta s)^2}{2\sigma^2 s}} dy \in \mathbb{R}.
\end{aligned} \tag{2.68}$$

From (2.55):

$$b^2_{vg} = \text{[drift subordinator].[diffusion subordinate]} = 0 \qquad \text{[using lemma 2.1]}. \tag{2.69}$$

From (2.56):

$$\begin{aligned}
\nu_{vg}(dx) &= \text{[drift subordinator].[Levy measure subordinate]} + \int_0^\infty P^s_{\sigma B_1+\theta}(dx)\tilde{\nu}(ds) \\
&= 0 + dx \int_0^\infty \frac{1}{\sigma\sqrt{2\pi s}} e^{-\frac{(x-\theta s)^2}{2\sigma^2 s}} \frac{\alpha}{s} e^{-\beta s} ds \ \text{[with } \alpha = \frac{1}{\gamma},\ \beta = \frac{1}{\gamma} \text{ and using Lemma2.1]} \\
&= \frac{\alpha}{\sigma\sqrt{2\pi}} dx \int_0^\infty e^{-\frac{(x-\theta s)^2}{2\sigma^2 s}} s^{-\frac{3}{2}} e^{-\beta s} ds
\end{aligned} \tag{2.70}$$

Now:

$$
\begin{aligned}
& e^{-\frac{(x-\theta s)^2}{2\sigma^2 s}} e^{-\beta s} \\
= & \; e^{-\frac{x^2-2xs\theta+\theta^2 s^2}{2s\sigma^2}-\beta s} \\
= & \; e^{\frac{x\theta}{\sigma^2}} e^{\left[-\left(\beta+\frac{\theta^2}{2\sigma^2}\right)s-\left(\frac{x^2}{2\sigma^2}\right)\frac{1}{s}\right]}
\end{aligned}
$$

So from (2.70) we obtain:

$$
\begin{aligned}
\nu_{vg}(dx) &= \frac{\alpha}{\sigma\sqrt{2\pi}} dx \int_0^\infty e^{-\frac{(x-\theta s)^2}{2\sigma^2 s}} s^{-\frac{3}{2}} e^{-\beta s} ds \\
&= \frac{\alpha}{\sigma\sqrt{2\pi}} e^{\frac{x\theta}{\sigma^2}} dx \int_0^\infty s^{-\frac{3}{2}} \exp\left[-\left(\beta+\frac{\theta^2}{2\sigma^2}\right)s-\left(\frac{x^2}{2\sigma^2}\right)\frac{1}{s}\right] ds \qquad (2.71)
\end{aligned}
$$

To evaluate the integral in (2.71) we need to use (2.62). Assume $\kappa = (\alpha + \frac{\theta^2}{2\sigma^2})$ and $s' = \kappa s$. Thus rearranging the integrand in (2.71) we obtain:

$$
\begin{aligned}
& \int_0^\infty \left(\frac{s'}{\kappa}\right)^{-\frac{3}{2}} \exp\left[-s' - \left(\frac{\frac{2x^2\kappa}{\sigma^2}}{4s'}\right)\right] \frac{1}{\kappa} ds' \\
= & \; \sqrt{\kappa} \int_0^\infty (s')^{-\frac{3}{2}} \exp\left[-s' - \left(\frac{\left(\sqrt{\frac{2x^2\kappa}{\sigma^2}}\right)^2}{4s'}\right)\right] ds' \\
\overset{(2.62)}{=} & \; \sqrt{\kappa} K_{\frac{1}{2}}\left(\sqrt{\frac{2x^2\kappa}{\sigma^2}}\right) \left[\frac{1}{2}\left(\frac{\sqrt{\frac{2x^2\kappa}{\sigma^2}}}{2}\right)^{\frac{1}{2}}\right]^{-1} \qquad (2.72)
\end{aligned}
$$

Thus (2.71) turns into:

$$
\begin{aligned}
\nu_{vg}(dx) &= \frac{\alpha}{\sigma\sqrt{2\pi}} e^{\frac{x\theta}{\sigma^2}} \frac{\sqrt{\kappa} K_{\frac{1}{2}}\left(\sqrt{\frac{2x^2\kappa}{\sigma^2}}\right)}{\left[\frac{1}{2}\left(\frac{\sqrt{\frac{2x^2\kappa}{\sigma^2}}}{2}\right)^{\frac{1}{2}}\right]} dx \\
&\overset{(2.63)}{=} \frac{\alpha}{\sigma\sqrt{2\pi}} e^{\frac{x\theta}{\sigma^2}} \frac{\sqrt{\kappa}\sqrt{\frac{\pi}{2}} \cdot \frac{1}{\left(\sqrt{\frac{2x^2\kappa}{\sigma^2}}\right)^{\frac{1}{2}}} e^{-\sqrt{\frac{2x^2\kappa}{\sigma^2}}}}{\left[\frac{1}{2}\left(\frac{\sqrt{\frac{2x^2\kappa}{\sigma^2}}}{2}\right)^{\frac{1}{2}}\right]} dx \\
&= \frac{\alpha}{\mid x \mid} \exp\left(\frac{x\theta}{\sigma^2} - \frac{\mid x \mid}{\sigma}\sqrt{2\kappa}\right) dx \\
&= \frac{1}{\gamma \mid x \mid} \exp\left(\frac{x\theta}{\sigma^2} - \frac{\mid x \mid}{\sigma}\sqrt{\frac{2}{\gamma} + \frac{\theta^2}{\sigma^2}}\right) dx \qquad (2.73)
\end{aligned}
$$

The last equality follows by plugging back the value of κ with $\alpha = \frac{1}{\gamma}$.

2.5.3 Geman's Misspecification and Our Correction

We gathered enough knowledge to report that Geman's(2002)[62] identificationof the parameters of the VG model is wrong. Such identification do not yield the Lévy measure mentioned in that paper. We identify the parameters correctly and show that our identification produces the expression of the Lévy measure which Geman(2002)[62]used in the paper for numerical works. According to the theorem 2.7, the expression of Lévy measure ought to be unique.

Geman et al(2001)[63] show that the VG process may be expressed as the difference of two independent gamma processes:

$$Y_t = G_t^p - G_t^n \tag{2.74}$$

where G_t^p and G_t^n are interpreted as price changes from positive and negative shocks respectively. This is clearly argued in Geman(2001)[63] and Geman(2002)[62]. The idea is simple : the difference of two positive jump processes can describe a general stock price path under the assumption that all possible movements in price are caused by frequent tiny and occasional big jumps of both positive and negative types. Positive moves are caused by a positive Gamma process and negative moves are caused by the negative of a positive Gamma process. For the validity of (2.74), according to (2.66) and (2.67), it suffices to have:

$$\begin{aligned} \frac{1}{1 - is\theta\gamma + \frac{1}{2}s^2\sigma^2\gamma} &= \left(\frac{1}{1 - is\eta_p}\right) \cdot \left(\frac{1}{1 - is\eta_n}\right) \\ &= \frac{1}{1 - is(\eta_p - \eta_n) + s^2\eta_p\eta_n} \end{aligned} \tag{2.75}$$

This is equivalent to:

$$\eta_p - \eta_n = \theta\gamma \tag{2.76}$$

$$\eta_p\eta_n = \frac{\sigma^2\gamma}{2} \tag{2.77}$$

Equation (2.75) follows from the fact that VG process is characterized by VG characteristic function. Since VG has equivalent characterization involving two Gamma processes; it's characteristic function has equivalent characterization involving characteristic functions of corresponding Gamma processes.

Geman(2002)[62] specified the solutions of (2.76) and (2.77) as:

$$\eta_p = \frac{1}{\sqrt{\frac{\theta^2\gamma^2}{4} + \frac{\sigma^2\gamma}{2}} - \frac{\theta\gamma}{2}} \tag{2.78}$$

$$\eta_n = \frac{1}{\sqrt{\frac{\theta^2\gamma^2}{4} + \frac{\sigma^2\gamma}{2}} + \frac{\theta\gamma}{2}} \tag{2.79}$$

Then Geman(2002)[62] mentioned that these specifications the Lévy measure of the VG process can be written as:

$$\nu_{vg}(dx) = \begin{cases} C\frac{e^{-Mx}}{x}dx & \text{if } x > 0, \\ C\frac{e^{-Gx}}{|x|}dx & \text{if } x < 0, \end{cases} \tag{2.80}$$

with $C = \frac{1}{\gamma}$, $G = \frac{1}{\eta_n}$, $M = \frac{1}{\eta_p}$.

However using our derived form of Lévy measure, see (2.73), we checked that solutions (2.78) and (2.79) do not yield the expression of the Lévy measure (2.80). Moreover (2.78) and (2.79) don't even satisfy (2.76) and (2.77). We now solve equations (2.76) and (2.77) separately for η_p and η_n.

For η_p we write (2.76) as $\eta_p = \eta_n + \theta\gamma$. Then from (2.77) we obtain:

$$\begin{aligned} (\eta_n + \theta\gamma)\eta_n &= \frac{\sigma^2\gamma}{2} \\ \Longrightarrow \eta_n &= \frac{-2\theta\eta \pm \sqrt{4\theta^2\gamma^2 + 8\sigma^2\gamma}}{4} \\ &\overset{(+)}{=} \sqrt{\frac{\theta^2\gamma^2}{4} + \frac{\sigma^2\gamma}{2}} - \frac{\theta\gamma}{2} \end{aligned}$$

Then again from (2.77) we obtain:

$$\eta_p = \frac{1}{\frac{2}{\sigma^2\gamma}\left(\sqrt{\frac{\theta^2\gamma^2}{4} + \frac{\sigma^2\gamma}{2}} - \frac{\theta\gamma}{2}\right)} \tag{2.81}$$

For η_n we write (2.76) as $\eta_n = \eta_p - \theta\gamma$. Then (2.77) implies:

$$\begin{aligned} (\eta_p - \theta\gamma)\eta_p &= \frac{\sigma^2\gamma}{2} \\ \Longrightarrow \eta_p &= \frac{2\theta\eta \pm \sqrt{4\theta^2\gamma^2 + 8\sigma^2\gamma}}{4} \\ &\overset{(+)}{=} \sqrt{\frac{\theta^2\gamma^2}{4} + \frac{\sigma^2\gamma}{2}} + \frac{\theta\gamma}{2} \end{aligned}$$

Then again from (2.77) we obtain:

$$\eta_n = \frac{1}{\frac{2}{\sigma^2\gamma}\left(\sqrt{\frac{\theta^2\gamma^2}{4}+\frac{\sigma^2\gamma}{2}}+\frac{\theta\gamma}{2}\right)} \tag{2.82}$$

Our solutions satisfy (??). Moreover using (2.73), we now prove that our solutions yield the form of Lévy measure used by Geman(2002)[62].

From (4.14) for $x > 0$ we obtain:

$$\begin{aligned}\nu_{vg}(dx) &= \frac{1}{\gamma\mid x\mid}\exp\left[x\left(\frac{\theta}{\sigma^2}-\frac{1}{\sigma}\sqrt{\frac{2}{\gamma}+\frac{\theta^2}{\sigma^2}}\right)\right]dx\\ &= \frac{1}{\gamma\mid x\mid}\exp\left[-x\frac{2}{\sigma^2\gamma}\left(\sqrt{\frac{\theta^2\gamma^2}{4}+\frac{\sigma^2\gamma}{2}}-\frac{\theta\gamma}{2}\right)\right]dx\\ &= C\frac{e^{-Mx}}{x} \quad x>0. \end{aligned} \tag{2.83}$$

where $C = \frac{1}{\gamma}$ and $M = \frac{1}{\eta_p}$ with η_p given by (2.81).

Similarly from (2.73), since for $x < 0$; $\mid x \mid= -x$ i.e. $x = -\mid x \mid$, we obtain:

$$\begin{aligned}\nu_{vg}(dx) &= \frac{1}{\gamma\mid x\mid}\exp\left[-\mid x\mid\left(\frac{\theta}{\sigma^2}+\frac{1}{\sigma}\sqrt{\frac{2}{\gamma}+\frac{\theta^2}{\sigma^2}}\right)\right]dx\\ &= \frac{1}{\gamma\mid x\mid}\exp\left[-\mid x\mid\frac{2}{\sigma^2\gamma}\left(\sqrt{\frac{\theta^2\gamma^2}{4}+\frac{\sigma^2\gamma}{2}}+\frac{\theta\gamma}{2}\right)\right]dx\\ &= C\frac{e^{-G|x|}}{x} \quad x<0. \end{aligned} \tag{2.84}$$

where $C = \frac{1}{\gamma}$ and $G = \frac{1}{\eta_n}$ with η_n given by (2.82). Equation (2.83) and (2.84) together imply (2.80).

2.6 Choosing a Pricing Measure In Incomplete Market

One of the downside of most of the otherwise sophisticated models is that these models render the market incomplete. See Schouten(2003)[102]. So for Lévy models we need to choose a pricing(martingale) measure out of many possibilities. Existence of a martingale measure is related to the absence of arbitrage while uniqueness of a martingale measure is related to the market completeness i.e. perfect hedging. In Lévy market there are many equivalent martingale measures under which discounted asset price process is a martingale;

so perfect hedging can not be obtained i.e. there always remain residual risk that can not be hedged. In this section we provide detailed technical analysis to show how to choose a risk-neutral distribution(martingale measure) corresponding to a physical(statistical) measure underlying a Lévy process.

One approach to to find an equivalent martingale measure, which is analytically more tractable and hence frequently used in the literature, is premised on conditional Esscher transform proposed by Gerber and Shiu(1994)[65]. Given a statistical distribution P, describing the evolution of the true underlying return process ,the conditional Esscher transform identifies an equivalent probability measure Q describing a corresponding martigale process. The Esscher parameter plays the prominent role in identifying the measure Q so that the discounted price process becomes martingale.

We consider $S_t = S_0 e^{X_t}$, where X_t is a Lévy process which can be seen as a continuously compounded rate of return over a period of length t. According to section2.2 X_t has an infinitely divisible distribution with probability density function given by $f(x,t)$. Assuming that the moment generating function(mgf) of this density exists for each t, it is defined as:

$$\begin{aligned} M(u,t) &= \mathbb{E}[e^{uX_t}] \\ &= \int_{-\infty}^{\infty} e^{ux} f(x,t)dx. \end{aligned} \tag{2.85}$$

Provided $M(u,t)$ is continuous at $t = 0$ it follows from infinite divisibility, see section2.2, that $M(u,t) = [M(u,1)]^t$. Let θ be a real number such that $M(\theta) = \int_{-\infty}^{\infty} e^{\theta x} f(x)dx$ exists, then the Esscher transform(with parameter θ) of the process $\{X_t\}_{t\geq 0}$ is defined to be a process with new probability density, for each $t > 0$, given by:

$$f^q(x,t;\theta) = \frac{e^{\theta x} f(x,t)}{\int_{-\infty}^{\infty} e^{\theta y} f(y,t)dy} = \frac{e^{\theta x} f(x,t)}{M(\theta,t)} \tag{2.86}$$

This implies that the Esscher equivalent measure is given by:

$$\begin{aligned} \frac{dQ}{dP} &= \frac{e^{\theta X_t}}{\mathbb{E}[e^{\theta X_t}]} \\ &= \frac{e^{\theta X_t}}{M(\theta,t)} \\ &= \frac{e^{\theta X_t}}{[M(\theta,1)]^t} \\ &= \exp\left(\theta X_t - t\log\left(M(\theta)\right)\right) \end{aligned} \tag{2.87}$$

The idea of equivalence comes through the fact that for a null event{}, when $\int_{\{\}} f(x,t) = 0$, (2.86) shows that $\int_{\{\}} f^q(x,t) = 0$ too. Similarly for a whole event when $\int_{-\infty}^{\infty} f(x,t) = 1$, (2.86) shows that $\int_{-\infty}^{\infty} f^q(x,t) = \int_{-\infty}^{\infty} \frac{e^{\theta x} f(x,t)}{M(\theta,t)} = 1$ too. Furthermore for all $u \in \mathbb{R}$, $\frac{e^{\theta X_t}}{\mathbb{E}[e^{\theta X_t}]}$ is a martingale with constant expectation of 1 for each t. Equation (2.86) is the core of all Esscher manipulation.

The mgf corresponding to f^q is:

$$\begin{aligned} M^q(u,t;\theta) &= \int_{-\infty}^{-\infty} e^{ux} f^q(x,t;\theta) \\ &= \int_{-\infty}^{-\infty} e^{ux} \frac{e^{\theta x} f(x,t)}{M(\theta,t)} \\ &= \frac{1}{M(\theta,t)} \int_{-\infty}^{-\infty} e^{(u+\theta)x} f(x,t)dx \\ &= \frac{M(u+\theta,t)}{M(\theta,t)} \\ &= \frac{[M(u+\theta,1)]^t}{[M(\theta,1)]^t} \\ &= \left[\frac{M(u+\theta,1)}{M(\theta,1)}\right]^t \\ &= [M^q(u,1;\theta)]^t \end{aligned} \tag{2.88}$$

Esscher parameter θ is selected in a way so that the modified(actually shifted) probability measure Q is a martingale measure which is equivalent to the statistical probability measure P. The idea is to find $\theta = \theta^*$, so that the discounted stock price process $\left\{e^{-rt}S_t\right\}_{t\geq 0}$ is a martingale with respect to the probability measure corresponding to θ^*. Since the martingale condition is $S_0 = \mathbb{E}^{\mathbb{Q}}[e^{-rt}S_t] = e^{-rt}\mathbb{E}^{\mathbb{Q}}[S_t]$ this translates into finding θ^* which is a solution to:

$$\begin{aligned} S_0 &= e^{-rt} S_0 \mathbb{E}^{\mathbb{Q}}[e^{X_t}] && [\text{since} S_t = S_0 e^{X_t}] \\ &= e^{-rt} S_0 \int_{-\infty}^{\infty} e^x f^q(x,t;\theta)dx \\ &= e^{-rt} S_0 \int_{-\infty}^{\infty} e^x \frac{e^{\theta x} f(x,t)}{M(\theta,t)} dx && [\text{using}(2.86)] \\ &= e^{-rt} S_0 \int_{-\infty}^{\infty} \frac{e^{(\theta+1)x} f(x,t)}{M(\theta,t)} dx \\ &= e^{-rt} S_0 \frac{M(\theta+1,t)}{M(\theta,t)} \\ &= e^{-rt} S_0 [M^q(1,1;\theta)]^t && [\text{using}(2.88)] \end{aligned} \tag{2.89}$$

The solution does not depend on t, so considering $t = 1$ we obtain:

$$e^r = M^q(1, 1; \theta) = \frac{M(\theta + 1, 1)}{M(\theta, 1)} \tag{2.90}$$

Finally the solution θ^* of equation (2.90) provides the risk neutral density f^q of the log returns over an interval of length t through the real density f, as shown in (2.86). We will frequently use equation (2.90) to select a pricing measure in the thesis.

2.7 Conclusion

In this chapter we have revisited the basic aspects of Lévy processes. We then demonstrated how the standard Lévy-Kintchine formula may be interpreted as a series of shocks superimposed on a normal distribution. Using this derivation we have been able to offer a correct solution to the mis-specification in the characterization of the Lévy measure for the VG model derived by Geman (2002)[62]. We analyzed LKF to characterize the distributional aspects of log-returns in a way which is suitable to infer the time changing effects of return processes in finance. This requires revealing the detailed theoretical underpinnings in order to replace diffusion by jumps. In other words analytic development of VG process, time-changed version of Black-Scholes model, is revisited from the general theoretical perspective of time changing which is motivated by identifying and correcting a misspecification in Geman(2002)[62].

Chapter 3

Pricing with FFT and FRFT : Dynamic Views

Most of the processes in finance and economics belong to a rich family of stochastic processes known as Lévy processes. As we saw in previous chapter these processes can essentially include jumps of various sizes, arriving at various rates, along randomly evolved paths and Brownian motion is the only continuous member of this family. Complete characterization of processes in this family comes through the celebrated Lévy -Kintchine formula(LKF). Besides drift and diffusion components in such characterization, we saw that a measure(commonly known as Lévy measure) plays pivotal role in suiting these processes to different needs e.g. modeling jump effects. We reformulated the underlying random variables, embedded in celebrated Lévy-Kintchine formula, in a form which is suitable to infer time changing effects of Lévy processes used in financial modeling. In addition we saw that this form of LKF clarifies the fact that Lévy measure of a process alone characterizes both the rate and distribution of jumps of a particular size. Lévy triplet of all the time changed processes can be characterized using a common theoretical framework and we explored this framework to revisit VG process as a time changed process. This recognizes a simple misspecification in an earlier work of Geman(2000)[62].

In this chapter we conduct an empirical investigation using the VG process, which we rigorously revisited in previous chapter. Recently Chourdakis(2005)[29] introduces fractional FFT(FRFT) in option pricing. Chourdakis(2005)[29] presents detailed analysis to show how the computational efficiency improves with respect to various FRFT parameters compare to those in FFT procedure. FRFT and FFT basically differ on required computational time as one has the flexibility to choose a parameter more than the other. In this empirical chapter we focus on exposition of trade-off between models fitting performance and required calibration time for week by week dynamic calibration with FFT and FRFT specifications. Parameters time-varying feature in dynamic calibration is investigated which provides information about the stability of the model across time. Furthermore we investigate whether FRFT exhibits any distinct feature in addition to substantial reduction in required computational time. Saying otherwise, for Black-Scholes and its time changed version Variance Gamma model we investigate cross-maturity and cross-strike features of FRFT compare to those of FFT. To distinguish the effects of time changing under FRFT and FFT we pretend that neither Black-Scholes nor VG model has closed form solutions and models under FRFT and FFT are different. We obtain the Black-Scholes values for both FRFT and FFT in the same way as we obtain for VG-FRFT and VG-FFT models. After all VG is a pure jump Brownian motion for a change of calender time to business time, as detailed in previous chapter. We consider weekly S&P500 index option, unlike most of the studies considering daily prices.

The drifted Brownian motion, without time change, describes the assets log return through two parameters μ and b as in (2.28)(without compound Poisson parts). As the equation (2.64) shows VG is a Brownian motion with a change of calender time to business time by a gamma process. Thus when Brownian motion encounters a time change, LKF through (2.28), exemplifies thats jumps come into scenario and diffusion disappears. The parameters a_{vg} and b_{vg} play the same role for the VG process as a and b in (2.28) for a general Lévy process. Thus b_{vg} being zero, (2.28) shows that the dynamics of the log returns has no diffusion. That is how equation (2.28) establishes that *Brownian motion is pure jump only in business time.* Furthermore ν_{vg} completely describes the rate and distribution of both small and big jumps, as explained in section 2.1.3, when jumps come into the scenario as a consequence of time changing.

3.1 Risk-neutral Specifications

The characteristic function of VG model under real measure if given by (2.67), which can be written as:

$$\Phi_{Y_t}(s) = \exp\left\{-\frac{t}{\gamma}ln\left(1 - is\theta\gamma + \frac{1}{2}s^2\sigma^2\gamma\right)\right\} \tag{3.1}$$

We can extract two parts from (3.1). One is the drift part $\mu = 0$ and another is the non-drift part $\phi(s) = -\frac{t}{\gamma}ln\left(1 - is\theta\gamma + \frac{1}{2}s^2\sigma^2\gamma\right)$ .The drift part under risk-neutral measure can now be obtained, see Shiryaev(1999)[106], as:

$$\mu^{rn}(s) = i\left[r - \frac{\phi(-i)}{t}\right] st = i\left[r + \frac{1}{\gamma}ln\left(1 - \theta\gamma - \frac{1}{2}\sigma^2\gamma\right)\right] st \tag{3.2}$$

Finally the risk-neutral characteristic function can be obtained as:

$$\begin{aligned} \Phi^{rn}_{Y_t}(s) &= \exp\left\{\mu^{rn}(s) + \phi(s)\right\} \\ &= \exp\left\{i\left[r + \frac{1}{\gamma}ln\left(1 - \theta\gamma - \frac{1}{2}\sigma^2\gamma\right)\right] st - \frac{t}{\gamma}ln\left(1 - is\theta\gamma + \frac{1}{2}s^2\sigma^2\gamma\right)\right\} \end{aligned} \tag{3.3}$$

Similarly the risk-neutral characteristic function of Black-Scholes model can be obtained as:

$$\Phi^{rn}_{B_t}(s) = \exp\left\{i\left(r - \frac{1}{2}\sigma^2\right) st - \frac{1}{2}s^2\sigma^2 t\right\} \tag{3.4}$$

where the Brownian motion $B_t \sim N(\tilde{\mu}t, \sigma t)$ has the following characteristic function under the real measure:

$$\Phi_{B_t}(s) = \exp\left\{is\tilde{\mu}t - \frac{1}{2}s^2\sigma^2 t\right\}. \tag{3.5}$$

Our empirical study is conducted under the risk-neutral measures utilizing the characteristic functions (3.3) and (3.4).

3.2 Pricing with FFT and FRFT

We consider logarithm of the prices, $s_t = log(S_t)$ and $k = log(K)$ where K is the strike price of the option. As in Carr and Madan(1999)[27] the value of an European call with maturity T can be expressed as a function of k:

$$C_T(k) = \int_k^\infty e^{-rT}(e^s - e^k)q_T(s)ds \tag{3.6}$$

Here $q_T(s)$ is the risk-neutral density of the log prices. To ensure square integrability of $C_T(k)$, Carr and Madan(1999)[27], introduced modified call prices:

$$c_T(k) = e^{\alpha k} C_T(k), \qquad \alpha > 0 \tag{3.7}$$

where α is known as the dampening factor. Following Carr and Madan(1999)[27] an analytic expression for the pricing formula (3.6) can be obtained as:

$$C_T(k) = \frac{e^{-\alpha k}}{\pi} \int_0^\infty e^{-iuk} \psi_T(u) du \tag{3.8}$$

where $\psi_T(u)$ has an analytic expression:

$$\psi_T(u) = \frac{e^{-rT} \Phi_T(u - (\alpha+1)i)}{\alpha^2 + \alpha - u^2 - i(2\alpha+1)u} \tag{3.9}$$

Here Φ is the characteristic function of the model for which prices are computed. As mentioned earlier in our empirical study we will consider Φ for Black-Scholes and VG models under risk-neutral dynamics, given by (3.4) and (3.3) respectively.

Using numerical integration technique, e.g. trapezoidal rule, the integral appearing in (3.8) can be approximated as:

$$\int_0^\infty e^{-iuk} \psi_T(u) du \approx \sum_{j=0}^{N-1} e^{-iu_j k} \tilde{\psi_T}(u_j) \eta \tag{3.10}$$

where $\tilde{\psi_T}$ is same as ψ_T with weights attached by integration rule. η is grid spacing such that $u_j = \eta j$ and upper limit of integration is ηN.

For some integrable function f, the spirit of FFT lies in approximating the continuous Fourier Transform by its discrete version:

$$\int_0^\infty e^{-iuk} f(u) du \approx \sum_{j=0}^{N-1} e^{-i\frac{2\pi}{N} kj} f(u_j) \tag{3.11}$$

Usual approach in the literature is to fine-tune (3.10) to (3.11) and then obtain the option prices through (3.8). The technique is to consider only the useful log-strikes near log-spots:

$$k_\Delta = -\frac{N\lambda}{2} + \lambda\Delta + log(S_0) \qquad \Delta = 0, \cdots, N-1. \tag{3.12}$$

For Lévy models $S_0 = 1$, and then assuming $b = \frac{N\lambda}{2}$ equation (3.12) ensures that log-strikes range is $-b$ to b. Here λ is the grid length of equidistant log-strikes. We can write the sum in (3.10) as:

$$\sum_{j=0}^{N-1} e^{-iu_j k_\Delta}\tilde{\psi_T}(u_j)\eta = \sum_{j=0}^{N-1} e^{-iu_j(-\frac{N\lambda}{2}+\lambda\Delta)}\tilde{\psi_T}(u_j)\eta$$
$$\stackrel{(u_j=\eta j)}{=} \sum_{j=0}^{N-1} e^{-i\eta j\lambda\Delta}e^{i\eta j\frac{N\lambda}{2}}\tilde{\psi_T}(u_j)\eta \quad (3.13)$$

With the following notation we obtain equation (3.13) in the form of (3.11) which is particularly suitable to apply FFT on the vector f with components $f(u_i)$:

$$f(u_j) = e^{i\eta j\frac{N\lambda}{2}}\tilde{\psi_T}(u_j)\eta \quad (3.14)$$
$$\eta\lambda = \frac{2\pi}{N} \quad (3.15)$$

So out of three parameters η, λ, N, two can be chosen arbitrarily and the other should satisfy (3.15), the so called FFT condition. For better accuracy both η and λ have to be small thus N is required to be large. So there is a trade-off between accuracy and number of strikes(hence computational time). In our empirical study we use FFT parameters as in Carr and Madan(1999)[27].

FRFT is developed to get rid of condition (3.15), providing the flexibility to choose all three parameters. So we can choose smaller N to consider only effective strikes around spots, significantly reducing the computational time, in addition to choosing appropriate grid spacing parameters η and λ for satisfactory accuracy. It was first introduced in Bailey and Swartztrauber(1991)[6] and is recently incorporated into option pricing in Chourdakis(2005)[29]. FRFT is a fast and easy way to compute sums of the form:

$$\sum_{j=0}^{N-1} e^{-i2\pi kj\epsilon} f_j \quad (3.16)$$

Here ϵ is the fractional parameter. Clearly $\epsilon = \frac{1}{N}$ yields the usual FFT. Upon choice of the parameter N, upper integration limit a and log-strike bound b, the grid spacing and fractional parameters can be obtained as:

$$\eta = \frac{a}{N} \quad (3.17)$$
$$\lambda = \frac{2b}{N} \quad (3.18)$$
$$\epsilon = \frac{1}{N} = \frac{\eta\lambda}{2\pi} \quad (3.19)$$

In our empirical study we use $a = 64$, $b = 0.3$ and $N = 32$. Consistent way of choosing FRFT parameters and related issues are discussed in Lee(2004)[78]. To compute $N-$point FRFT for a vector x, the algorithm suggest, see Bailey and Swartztrauber(1991)[6], defining $2N-$ point vectors as:

$$y_j = \begin{cases} x_j e^{-i\pi j^2 \epsilon} & 0 \leq j < m \\ 0 & m \leq j < 2m. \end{cases} \tag{3.20}$$

$$z_j = \begin{cases} e^{i\pi j^2 \epsilon} & 0 \leq j < m \\ e^{i\pi (j-2m)^2 \epsilon} & m \leq j < 2m \end{cases} \tag{3.21}$$

where ϵ is as given by (3.19). The FRFT is then computed as:

$$G_k(x, \epsilon) = e^{-i\pi k^2 \epsilon} \odot D_k^{-1}[D_j(y) \odot D_j(z)] \tag{3.22}$$

Here $\odot$ stands for element wise multiplication, $D_j(.)$ is the discrete fourier transform computed with the usual FFT procedure as in (3.11) and D^{-1} is the inverse fourier transform.

Our closed form Black-Scholes prices, used in comparison, are calculated under the risk-neutral measure using the following celebrated result, see Black and Scholes(1973)[19]:

Theorem 3.1 *Consider a European call option with strike price K and expiration time T. If the underlying option pays no dividends and continuously compounded risk-free rate is r, then the price of the contract at time t is given by:*

$$C(t, S_t) = S_t \Phi(d_1) - K e^{-r(T-t)} \Phi(d_2) \tag{3.23}$$

where $\Phi(x)$ denotes the cumulative distribution function of standard normal random variable evaluated at the point x, $d_1 = \frac{[ln(\frac{S_t}{K}) + (r + \frac{\sigma^2}{2})(T-t)]}{\sigma\sqrt{T-t}}$ and $d_2 = \frac{[ln(\frac{S_t}{K}) + (r - \frac{\sigma^2}{2})(T-t)]}{\sigma\sqrt{T-t}}$ with $d_2 = d_1 - \sigma\sqrt{T-t}$.

3.3 Empirical Study

Chourdakis(2005)[29] used some selected values of parameters and didn't calibrate the models with real market data. We calibrate the models separately assuming FRFT and FFT as different models. For this we consider options traded on S&P500 for the sample period of

January'2007 to November'2007. For out-of-sample assessment we consider market prices of options traded on last week of December 07[1]. Though we present our in-sample analysis only for one in-sample week, we investigate several other in-sample weeks as well.

Calibration Results					
Specifications	RMSE	Average time (second)	σ	θ	η
VG(FFT)	2.6931	20.97	0.1294 (0.0393)	-0.1802 (0.0268)	0.0786 (0.0221)
VG(FRFT)	2.7234	0.45	0.1232 (0.0505)	-0.1837 (0.0313)	0.0839 (0.0276)
BS(FFT)	3.1765	11.27	0.1320 (0.0360)		
BS(FRFT)	3.2447	0.29	0.1308 (0.0362)		
BS(closed form)	3.1764	0.063	0.1320 (0.0360)		

Figure 3.1: *Calibration results under different specifications of Black-Scholes and Variance-Gamma. We consider weekly traded options on S&P500 from January'07 to November'07. The estimates reported are the average of dynamic weekly calibrations over this sample period. The standard error of each estimate appears in parenthesis. The average(over 44 weeks)weekly calibration time is also reported.*

The parameters reported in table3.1 are the average of weekly estimates over the sample period.

For in-sample prices figure3.2 shows the Black-Scholes fit and figure3.3 shows the VG fit both for FRFT and FFT. For out-of-sample prices the corresponding fits are presented in figure3.4 and figure3.5 respectively.

[1]Though we present the out-sample fiton last week of December,2007; we investigate earlier weeks of December as well. We observe that as we move from 1st week to last week, the out-of-sample fits get worse.

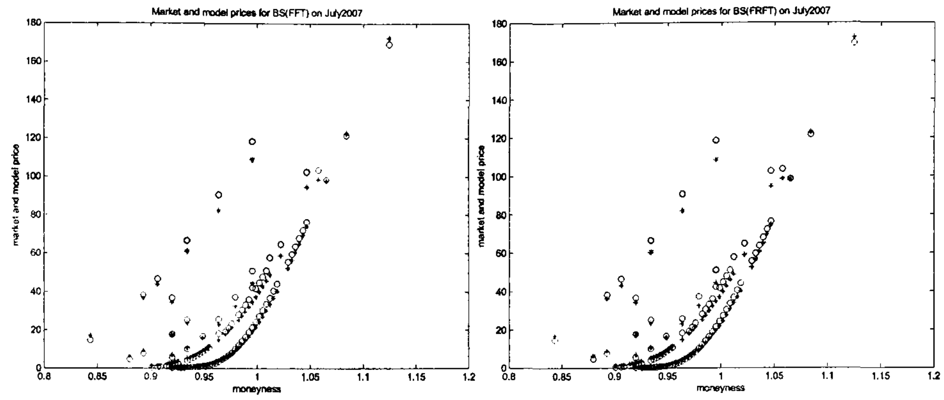


Figure 3.2: *In-sample Black-Scholes fit under FFT(left) and FRFT(right). 0(market), *(model) and different colors are for different maturities as red(23dtm), blue(58dtm), green(86dtm), ceylon(149dtm), yellow(240dtm) and black(331dtm).*

3.3.1 Dynamic Distinction Between FFT and FRFT

We focus whether specifications with FFT and FRFT exhibit any distinctive feature for dynamic weekly calibration over the sample period of January'07 to November'07. However in figures we present the case with third weeks of each month, i.e. mid month. Figure3.6 presents the number of options used in such dynamic calibration. Volatility estimates at each week under different specifications are shown in figure3.7(left). It shows that it is not FFT and FRFT which cause difference in dynamic volatility estimation, rather it is time change which systematically estimates slightly higher level of volatility. On the right hand side, of figure3.7, we show that VG model exhibits better calibration performance than BS model. We conjecture that VG estimate of dynamic volatility is a better reflection of true volatility than the BS estimate, which possibly leaves a favorable calibration for the VG model. After all VG model captures the volatility through all its three parameters where as in case of BS model it is captured by its sole parameter. We see that for both BS and VG, FFT performs slightly better than FRFT throughout the months. However figure3.8 shows the difference in time requirements for the dynamic calibrations with FFT and FRFT

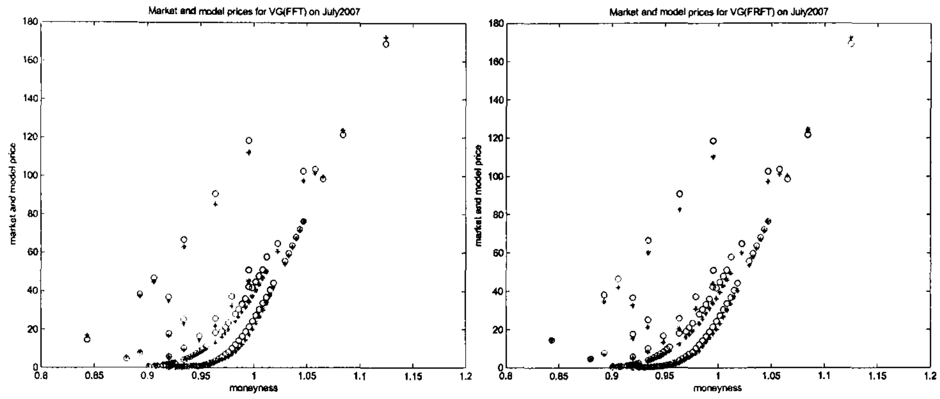


Figure 3.3: *In-sample Variance Gamma fit under FFT(left) and FRFT(right). 0(market), *(model) and different colors are for different maturities as red(23dtm), blue(58dtm), green(86dtm), ceylon(149dtm), yellow(240dtm) and black(331dtm).*

specifications. It is now a trade-off between slightly favorable, often negligible,calibration performance and the requirement of significantly longer calibration time.

For weekly dynamic calibration the average of estimates are found to reflect models inherent stability over the entire calibration period, for both FFT and FRFT specifications. More specifically parameters time-varying tendency are found to be negligible under both FFT and FRFT specifications, implying that the means of such estimates are rather a good "make-do" approach to decide on the final parameter values over a long period. Though we do not report, we observe that other choices such as median and mode of dynamic weekly estimates are found to undermine the potentiality of time change, namely for such choice it is observed that VG model is not necessarily performing better than Black-Scholes model. In figures3.9 and 3.10 we graphically argue in favor of using the average of weekly parameter estimates in pricing. For both BS and VG these figures show that mean deviation of the dynamic estimates are roughly close to zero throughout the months.

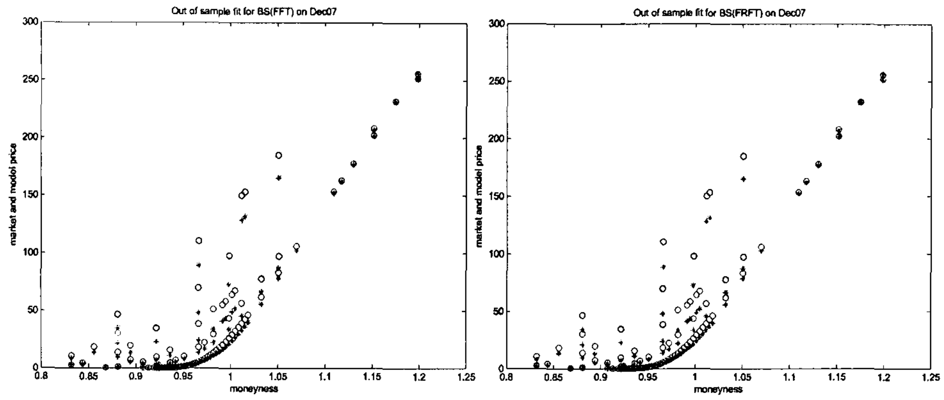


Figure 3.4: *Out-of-sample Black-Scholes fit under FFT(left) and FRFT(right). 0(market), *(model) and different colors are for different maturities as red(23dtm), blue(51dtm), green(86dtm), ceylon(114dtm), yellow(177dtm), black(268dtm), magenta(359dtm).*

3.3.2 Cross-maturity and Cross-strike investigation

We investigate the pricing errors for four model specifications BS(FRFT), BS(FFT), VG(FRFT) and VG(FFT), across maturity and strike, relative to the closed form Black-Scholes prices. Our motivation is to examine the impact of the FRFT and FFT valuation methods and the impact of the underlying models(BS vs. VG) on the option prices.

To reveal the cross-strike features of FRFT and FFT under time changed and original process we express pricing errors as function of strikes only, holding the maturity constant. We consider three different maturities observed in the market: minimum, mean and maximum corresponding to short, medium and long term options respectively:

$$ERROR^{model}(K_i) = P^{model}(K_i, t) - P^{BS}(K_i, t). \tag{3.24}$$

Similarly to reveal cross-maturity features of FRFT and FFT we express pricing errors as function of maturities only, holding the strike constant. Three different strikes are considered: minimum, equal to asset and maximum of the observed strikes in the market; these correspond to ITM, ATM and OTM options respectively:

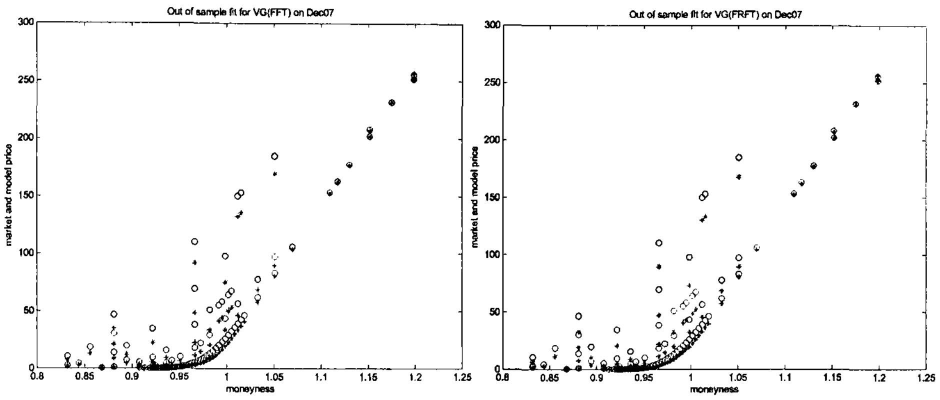


Figure 3.5: *Out-of-sample Variance Gamma fit under FFT(left) and FRFT(right). 0(market), *(model) and different colors are for different maturities as red(23dtm), blue(51dtm), green(86dtm), ceylon(114dtm), yellow(177dtm), black(268dtm), magenta(359dtm).*

$$ERROR^{model}(t_i) = P^{model}(K, t_i) - P^{BS}(K, t_i). \tag{3.25}$$

We plot cross-strike(left) and cross-maturity(right) errors, for different specifications, in figure3.11, using illustrative market data for the last week of July'07. Figure3.12 plots error surfaces across all ranges of strikes and maturities.

The first empirical observation is that when Fractional parameter of FRFT induces some unsystematic price fluctuations across strike(left panel in figure3.11) across-strike, its influence across-maturity is rather systematic(right panel in figure3.11). For any fixed maturity, across-strike prices under FRFT and FFT(for both BS and VG) eventually converge to closed form Black-Scholes prices. The higher the fixed maturity is, the slower the rate of convergence. For short term options FRFT fluctuations are closely around FFT fluctuations; for medium and long term options they systematically get deviated from each other. Over all across-strike the effect of time changing is rather systematic. Considering equation(2.28) and the discussion following equation(2.73) this means that changing the source of randomness from diffusion to jumps causes the prices to be higher for ITM and lower for OTM options. This provides some remedy to the Black-Scholes models deficiency

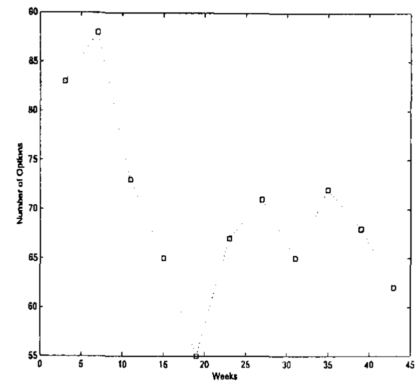


Figure 3.6: *Number of options used in dynamic calibration. The case presented is for third weeks of each month(mid-month). However in calibration we considerd all 44 week's for the sample period of Jan'07 to Nov'07*

in 'under pricing the ITM' and 'over pricing the OTM' options. This remedy is apparently the reason behind the VG models superior performance over BS model.

Across-maturity error patterns under VG model encounter a gradual reversal with respect to moneyness criteria. For ITM options VG prices are ,on average, higher than BS prices; for ATM it is ,on average, lower for short term options and higher for long term options. Finally in case of OTM options it is lower than BS prices. FFT and FRFT prices increasingly differ with the change of moneyness criteria. The greatest deviation is observed in case of OTM options. See figures3.11(right panel) and 3.12.

3.4 Conclusion

We calibrate the VG and BS models for weekly recorded option contracts using both FFT and FRFT methods. We observe that fractional parameter of FRFT causes some unsystematic price fluctuation across-strike. For short maturities FRFT prices fluctuate closely around FFT prices. However as the maturities increases two specifications give deviated prices. Across maturities FFT and FRFT prices increasingly differ with the change in mon-

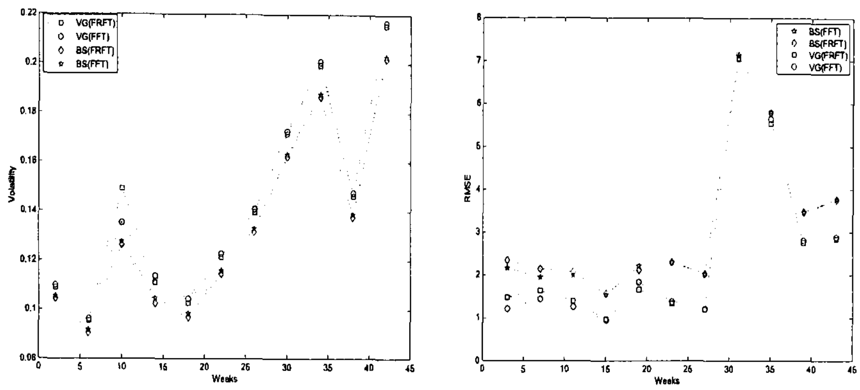


Figure 3.7: *Dynamic distinction in volatility(left) and RMSE(right) estimation with FFT and FRFT. The case presented is for third weeks of each month(mid-month). However in calibration we considerd all 44 week's for the sample period of Jan'07 to Nov'07.*

eyness status. These are related with characteristic function and moneyness grids in some complicated ways. More importantly like other studies we found that FRFT is much faster than FFT, economizing on 97-98% of the calculation time at a cost of small pricing errors. These findings have important implications for the calibration of options models and for options risk-management in general. We also observe that there are important differences between BS and VG option values, implying that inappropriate use of BS in the context where the true process was VG can lead to major pricing errors. Otherwise said, assuming the market is under regular diffusive shocks can lead to major pricing errors when the true market exhibits frequent small and big jump shocks. Models inherent stability in dynamic calibration is found to be similar for both FFT and FRFT specifications. Consequently mean values of dynamic weekly estimates are found to work well in out-of-sample as well.

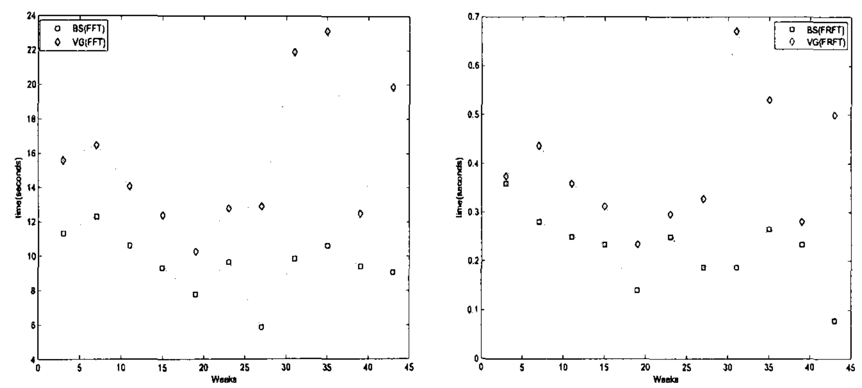


Figure 3.8: *Dynamic distinction in required time for calibration with FFT(left) and FRFT(right). The case presented is the required time for the calibration at third weeks of each month(mid-month). However in calibration we considerd all 44 week's for the sample period of Jan'07 to Nov'07.*

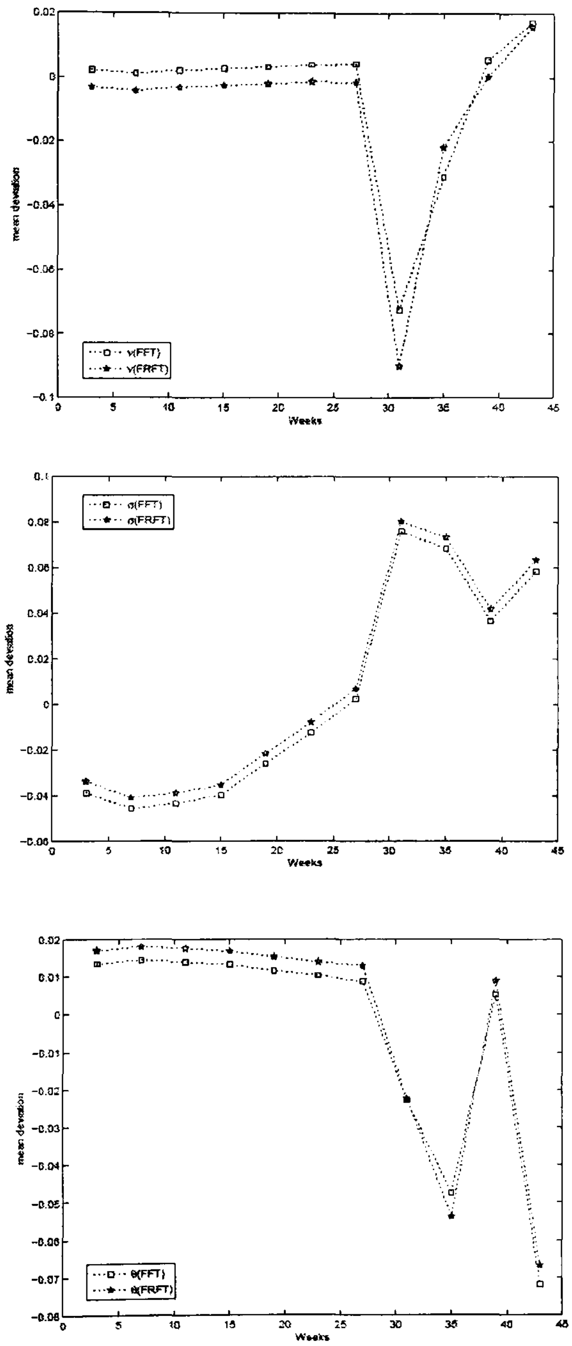


Figure 3.9: *Mean deviation of Variance Gamma parameter estimates under FFT and FRFT(ν, σ, θ;from top to bottom). The case presented is for third weeks of each month(mid-month). However in calibration we consider all 44 week's for the sample period of Jan'07 to Nov'07.*

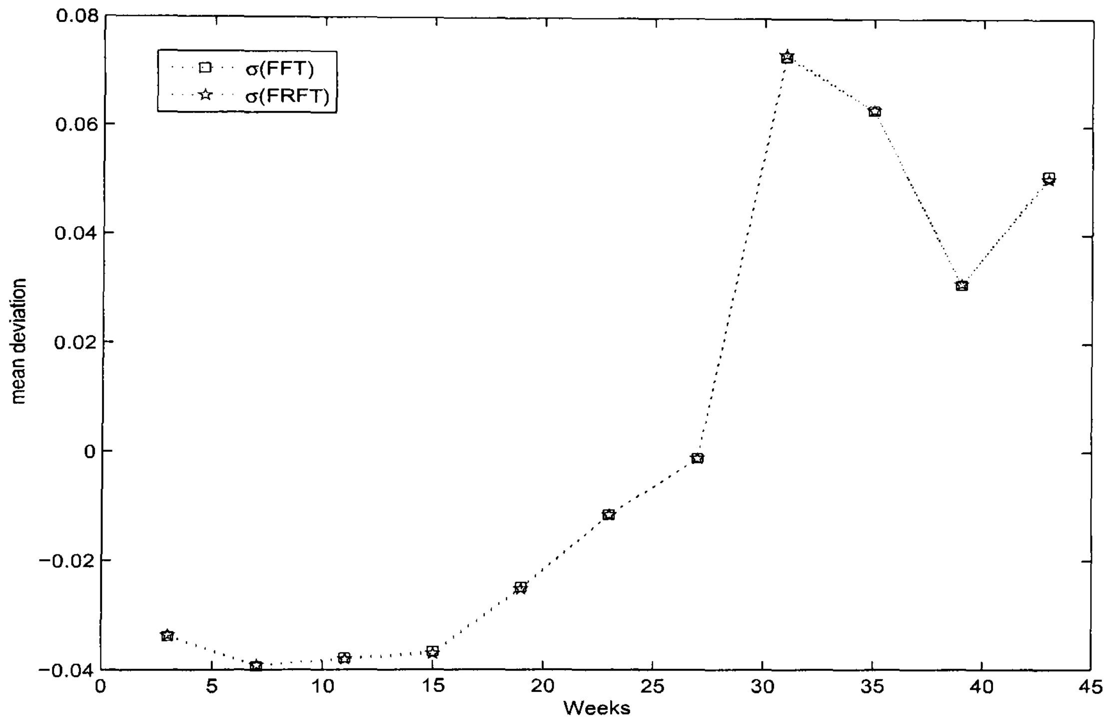


Figure 3.10: *Mean deviation of Black-Scholes parameter estimate under FFT and FRFT. The case presented is for third weeks of each month(mid-month). However in calibration we considerd all 44 week's for the sample period of Jan'07 to Nov'07.*

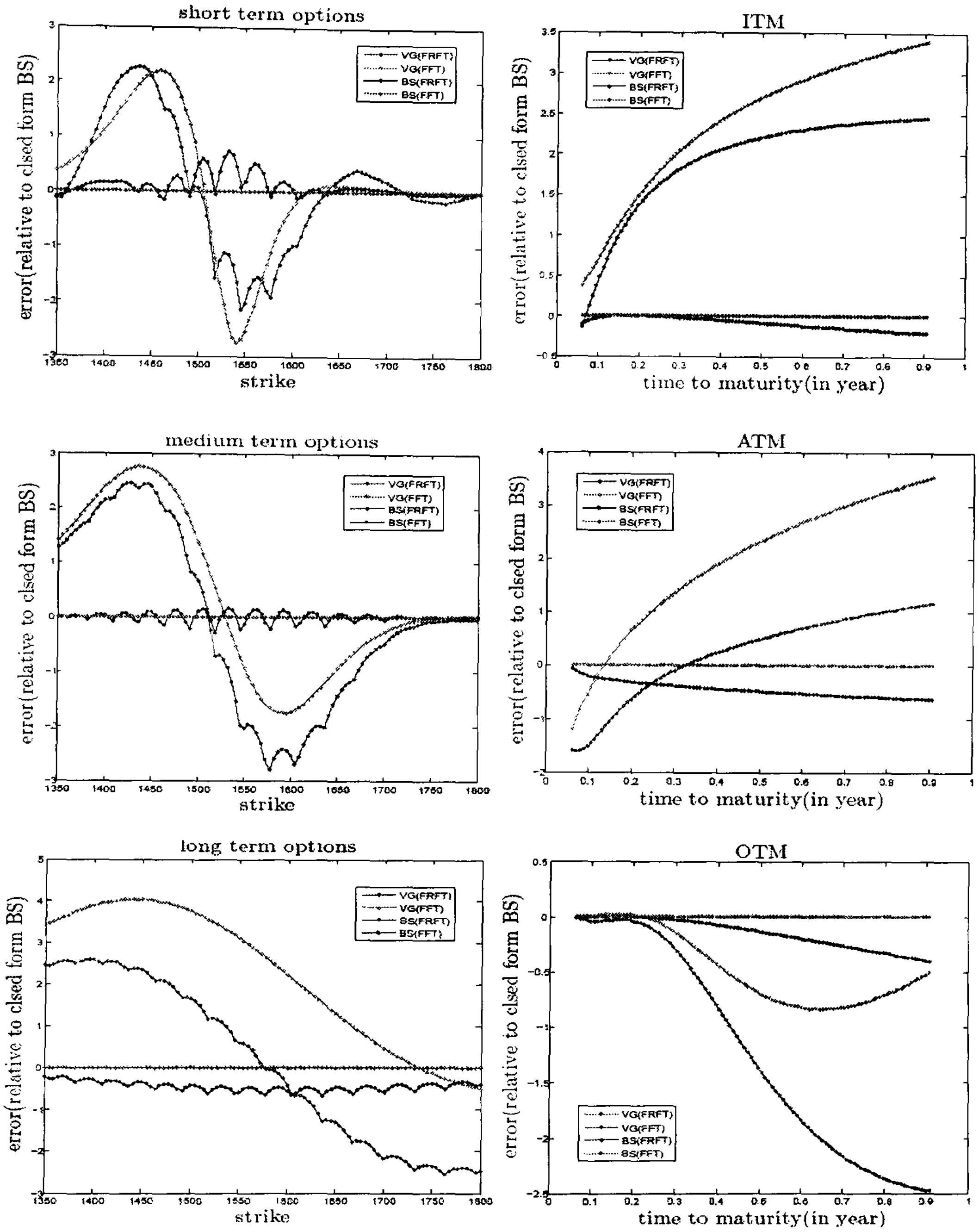


Figure 3.11: *Cross strike(left) and cross-maturity(right) features of FRFT and FFT under Black-Scholes and Variance Gamma models. We used the market information of last week of July'07. The models are calibrated over the sample period of Jan'07 to Nov'07. Cross strike features are presented for short(top), medium(middle) and long(bottom) term options. Cross maturity features are presented for ITM(top), ATM(middle) and OTM(bottom) options. The spot was 1518.09.*

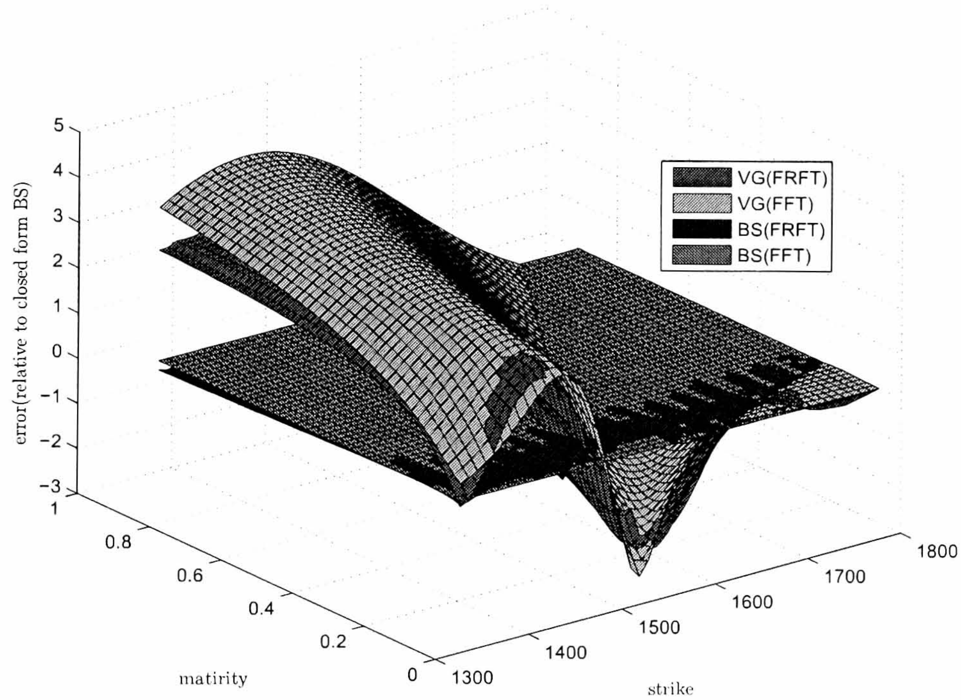


Figure 3.12: *FRFT and FFT features under Black-Scholes and Variance Gamma models. We use last week of July'07 market information. The spot was 1518.09. The models are calibrated over the sample period of Jan'07 to Nov'07*

Chapter 4

Existing Approaches to Nonnormality: Pricing and Approximation

This chapter carries out a comparative analysis of the calibration and performance of a variety of options pricing models. These include Black-Scholes (1973)[19], the Gram-Charlier (GC) approach of Backus et al. (1997)[9], the stochastic volatility (HS) model of Heston (1993)[69], the closed-form GARCH process of Heston and Nandi (2000)[70] and a variety of Lvy processes including the Variance Gamma (VG), Normal Inverse Gaussian (NIG), CGMY and Kou(2002)[75] jump-diffusion models. Unlike most studies of option pricing, we compare these models using a common point-in-time data that reflects the perspective of a new investor who wishes to choose between models using only the most minimal recent data set. For each of these models, we also examine the accuracy of delta and delta-gamma approximations to the valuation of both individual options and an illustrative option portfolio. Based on the relative performance of Heston Nandi(2000)[70] model (CFG henceforth), in both pricing and approximation, we emphasize the necessity of exploring similar closed form GARCH approach with nonnormal(Lévy) innovations.

4.1 Introduction

Empirical performance of alternative option pricing models was systematically documented in Bakshi et al(1997)[7]. Authors alternative selections in Bakshi et al(1997)[7], were basically around continuous time stochastic volatility, with and without jumps, and stochastic interest rate characterization. It is concluded that simple stochastic volatility(SV) has more pronounced effect than stochastic volatility stochastic interest rate(SVSI) and its jump included variant SVSI-J. However SV is just one way of incorporating skewness. We consider empirical assessment of SV with other approaches of incorporating skewness and kurtosis: Gram-Charlier approach, GARCH approach, pure jump Lévy approach and jump-diffusion approach. Unlike Bakshi et al.[7], we consider single day traded options of weekly data for our empirical investigation.

Gram-Charlier is a continuous time approach which explicitly incorporates skewness and kurtosis to bench-mark Black-Schole(1973)[19] model. Since it doesn't incorporate market evidence of jumps into the return dynamics, it is a continuous path approach as well. Moreover this approach still considers volatility as constant. On the other hand GARCH is a discrete time approach which allows jumps to be incorporated in return dynamics. However most attractive feature of GARCH models is the realistic modelling of volatility, replacing the constant volatility phenomenon of the bench-mark model. Going with the frequent reference in recent literature we consider a GARCH version which does not include jumps in return dynamics. So our representative from the GARCH family is the Heston-Nandi closed form GARCH model(CFG) with discrete time continuous path approach to incorporating skewness and kurtosis to return dynamics. This model uses a GARCH(1,1) structure to update daily volatilities. So when Gram-Charlier model incorporates skewness and kurtosis without incorporating stochastic volatility and jumps, CFG model incorporates skewness and kurtosis by incorporating stochastic volatility into the return dynamics. Since our GARCH representative is the discrete time stochastic volatility model, we consider the continuous time stochastic volatility model of Heston(HS) as well; given the fact that it is another commonly used model. In this continuous volatility adjustment approach, volatility is driven by a separate Stochastic Differential Equation(SDE) namely CIR process.

Our jump incorporating models are the frequently refereed Lévy models in option pricing literature, see Schouten(2003)[102], Kyprianou(2006)[76] and the references therein. So continuous time jumpy path approach to incorporating skewness and kurtosis is examined by Variance Gamma(VG), Normal Inverse Gaussin(NIG) and CGMY, a further extension of VG, models. All these Lévy models are infinite activity pure jump (with suitable parametrization of CGMY), where small jumps are so frequent that it renders diffusion redundant, see Geman(2002)[62]. Nonetheless we consider a sole representative of finite activity jump-diffusion Lévy model, namely Kou's double exponential(DE) model, see Kou(2002)[75].

Some of the Lévy models we consider(namely VG, NIG and CGMY) introduce skewness and kurtosis through stochastic time changing feature of asset pricing. Excess kurtosis in such models may result from the implicit stochastic volatility induced by time changing, see Geman(2002)[62]. We didn't consider Lévy stochastic volatility models which incorporate stochastic volatility through separate dynamics. These models are rarely used in the market because of their involved mathematical manipulation for marginal improvement in pricing. So in this chapter by stochastic volatility we either mean Heston's stochastic volatility(SV) model or GARCH stochastic volatility(CFG) model. Stochastic time changing is believed to improve the pricing performance significantly as documented in Geman(2002)[62] for intra day tic data. This chapter assesses the time changed models compare to other common approaches to pricing subject to an investor's willingness to use most recent minimum market information contained in a single day traded options.

The feasibility of approximating option portfolio and inherent pitfall in such approximation under BS pricing model, caused by non-linearity of realistic composition of the portfolio, was investigated in Britten and Schaefer(1999)[22]. It was further explored in Christoffersen(2003)[31]. It is concluded that from the standpoint of reality, there is hardly any alternative to the full valuation of the portfolio. However under normal market situation such approximation is justified in short period. We attempt to enrich the literature by exploring whether various approaches to incorporating skewness and kurtosis to pricing models discriminate the performance of such approximations. We consider complete pay-off profile of the portfolio instead of considering any particular risk measure.

The performance of approximation, under various models, depends on models pricing performance as well as their Delta and Gamma values. Though some models have closed form Greeks, most of the sophisticated models don't have any. There are several ways to numerically compute the Greeks for those models, namely tree based approach, Monte-Carlo approach and finite-difference approach. However for uniformity in comparison we disregard the closed form formulas for Greeks, whenever available, and compute Greeks by finite difference approach for all the models under investigation.

For empirical study we consider options traded on S&P500 index on Wednesday 23rd January,2008. We consider the immediate next Thursday data for out-of-sample assessment of the models.

This chapter is structured in the following way. Section 4.2 provides short description of the models which consider different approaches to incorporating skewness and kurtosis. Risk-neutral dynamics are revisited which are required for pricing options. Then in section 4.3 we discuss issues around the implementation of Greeks($\Delta, \Gamma,$) and approximation of option prices. Comparative look into approximation pitfall is presented in section 4.4. Section 4.5 deals with data and calibration issues. Our empirical findings are discussed in section 4.6. We present pricing and approximation analysis separately. Finally the last section concludes.

4.2 The Models and the Dynamics

Celebrated Black-Schole-Merton ,BS hereafter, idea capitalizes on Geometric Brownian motion(GBM) for asset return. BS provides a closed form solution to European option; a simple derivative with non-linear pay-off. The basic idea of European style derivative pricing is captured in the following central result, which proof can be found in any classic finance book, e.g. Shiryaev(1999)[106]:

Theorem 4.1 *Consider a European option with pay-off $V(S)$ and expiration time T. Assume the continuously compounded rate of interest is r. Then the current European option price is determined by:*

$$\nu(0, S_0) = e^{-rT}\hat{E}[V(S_T)] \tag{4.1}$$

where $\hat{E}$ denotes the expectation under the risk neutral probability that is derived from the risk-neutral process:

$$\frac{dS_t}{S_t} = rdt + \sigma dB_t. \tag{4.2}$$

BS considered normal distribution for log-returns and therefore fails to incorporate empirical evidence of smile-skew effects resulting from skewness and kurtosis. Gram-Charlier explicitly incorporates skewness and kurtosis in BS framework therefore ensures substantial improvement in pricing performance. However more realistic pricing requires replacing the Brownian motion B_t by characteristically more rich Lévy processes. Lévy processes can incorporate the empirical evidence of jumps in return in addition to structural feasibility of allowing the return distributions to have skewness and kurtosis. This could often improve the pricing performance significantly. A comprehensive survey of Lévy processes in finance can be found in Schouten(2003)[102], Cont and Tankov(2004)[38], Liuren(2006)[77] and accessible theoretical treatment of Lévy processes can be found in Kyprianou(2006)[76], Sato(1999)[100], Applebeum(2004)[2]. In practice for most of these Lévy models prices have to be computed through numerical inversion of characteristic functions which is obviously time consuming. This introduces some kind of trade off between quick implementation of BS model obtaining more consistent prices from otherwise sophisticated models which require considerable time to implement. Nonetheless despite this obvious drawback a plethora of alternative option pricing models are developed in recent times and option pricing is still a vibrant research area in its own merit. Moreover these alternative approaches can possibly shed further lights on other aspects of the models e.g. hedging performance.

We briefly revisit the pricing models which are premised on diverse approaches of development. The skewness and kurtosis in these models are incorporated and characterized quite differently. For pricing the options we use risk-neutral characterization of each model.

4.2.1 Gram-Charlier model

The Gram-Charlier approach was first introduced in Backus et al(1997)[9]. An extension of BS density was considered allowing for skewness and kurtosis:

$$f(x) = \phi(x) - \zeta_{1t}\frac{1}{3!}\phi^3(x) + \zeta_{2t}\frac{1}{4!}\phi^4(x) \tag{4.3}$$

where superscripts on ϕ indicate the order of derivative of BS density, $\zeta_{1t} = \frac{\zeta_{11}}{\sqrt{t}}$ and $\zeta_{2t} = \frac{\zeta_{21}}{t}$ are skewness and kurtosis on a horizon of t and ζ_{11} and ζ_{21} are per unit skewness and kurtosis. In Backus et al(1997)[9] it was shown that with this density the call option price can be written as:

$$\begin{aligned} c_{GC} &\simeq S_t\Phi(d) - Ke^{-rt}\Phi\left(d - \sqrt{t}\sigma\right) \\ &\quad + S_t\phi(d)\sigma\left[\frac{\zeta_{11}}{3!}(2\sqrt{t}\sigma - d) - \frac{\frac{\zeta_{21}}{\sqrt{t}}}{4!}(1 - d^2 + 3d\sqrt{t}\sigma - 3t\sigma^2)\right] \end{aligned} \tag{4.4}$$

4.2.2 Heston Stochastic Volatility model

Stochastic volatility model of Heston, see Heston(1993)[69], assumes a diffusion process for the stock price given by:

$$\frac{dS_t}{S_t} = \mu dt + \sqrt{\vartheta_t}dB_t^1 \tag{4.5}$$

$$\tag{4.6}$$

and a CIR process for the volatility $\sqrt{\vartheta_t}$ given by:

$$\begin{aligned} d\vartheta_t &= \kappa[\theta - \vartheta_t]dt + \sigma\sqrt{\vartheta_t}dB_t \\ dB_t dB_t^1 &= \rho dt. \end{aligned} \tag{4.7}$$

The SV model has flexible distributional structure in which the correlation(ρ) between volatility and asset returns serves to control the level of asymmetry and the volatility variation coefficient(σ) serves to control the level of kurtosis. The risk-neutral specification is similar to one given in (4.7) but κ and θ replaced by $\kappa^* = \kappa + \lambda$ and $\theta^* = \frac{\kappa\theta}{\kappa+\lambda}$, see Heston(1993)[69], Rouah and Vainberg(2007)[97]. Here λ is the market price of volatility risk. The closed form solution, up to numerical integration, in Heston model is given by:

$$\begin{aligned} c_{HS} = &= S_t\left(\frac{1}{2} + \frac{1}{\pi}\int_0^\infty Re\left[\frac{K^{-iz}f_1}{iz}\right]dz\right) \\ &\quad - Ke^{-rt}\left(\frac{1}{2} + \frac{1}{\pi}\int_0^\infty Re\left[\frac{K^{-iz}f_2}{iz}\right]dz\right) \end{aligned} \tag{4.8}$$

where $f_j = \exp\left\{C_j + D_j\vartheta + izx\right\}$ with

$$
\begin{aligned}
x &= x_t = log(S_t) \\
\vartheta &= \vartheta_t \\
C_j &= irz(T-t) + \frac{\kappa\theta}{\sigma^2}\left\{(b_j - i\rho\sigma z + d_j)(T-t) - 2log\left[\frac{1-g_j e^{d_j(T-t)}}{1-g_j}\right]\right\} \\
D_j &= \frac{b_j - iz\rho\sigma + d_j}{\sigma^2}\left[\frac{1-e^{d_j(T-t)}}{1-g_j e^{d_j(T-t)}}\right] \\
g_j &= \frac{b_j - iz\rho\sigma + d_j}{b_j - iz\rho\sigma - d_j} \\
d_j &= \sqrt{(iz\sigma\rho - b_j)^2 - (2iu_j z - z^2)\sigma^2} \\
u_1 &= \frac{1}{2}, \quad u_2 = -\frac{1}{2} \\
b_1 &= \kappa + \lambda - \rho\sigma, \quad b_2 = \kappa + \lambda
\end{aligned}
$$

4.2.3 Heston-Nandi GARCH model

Heston and Nandi(2000) provide a closed form pricing formula for a European option, where the underlying follows the non-linear GARCH process:

$$
\begin{aligned}
log(\frac{S_{t+1}}{S_t}) &= r + \lambda\sigma_{t+1}^2 + \sigma_{t+1}z_{t+1}; \qquad z_{t+1} \sim N(0,1) \\
\sigma_{t+1}^2 &= \omega + \alpha(z_t - \theta\sigma_t)^2 + \beta\sigma_t^2
\end{aligned}
\tag{4.9}
$$

From GARCH characterization (4.9) the variance persistence of return process can be derived to be $\beta + \alpha\theta^2$; so the process will be mean-reverting if $\beta + \alpha\theta^2 < 1$. It is shown in Heston and Nandi(2000)[70] that the risk-neutral characterization can be obtained by plugging $\lambda = -\frac{1}{2}$ and $\theta^* = \theta + \lambda + \frac{1}{2}$. Furthermore Heston and Nandi(2000)[70] argued that in this model α determines kurtosis and θ determines skewness. This model has a moment generating function(mgf) of the form:

$$
f(z) = S_t^z \exp\left\{A(t;t+T,z) + B(t;t+T,z)\sigma_{t+1}^2\right\} \tag{4.10}
$$

where $A(t;t+T,z)$ and $B(t;t+T,z)$ are given by the recursive relations:

$$\begin{aligned} A(t;t+T,z) &= A(t+1;t+T,z) + zr + B(t+1;t+T,z)\omega \\ &\qquad -\frac{1}{2}ln\left(1-2\alpha B(t+1;t+T,z)\right) \\ B(t;t+T,z) &= z(\lambda+\theta) - \frac{1}{2}\theta^2 + \beta B(t+1;t+T,z) \\ &\qquad +\frac{\frac{1}{2}(z-\theta)^2}{1-2\alpha B(t+1;t+T,z)} \end{aligned} \tag{4.11}$$

Heston and Nandi(2000)[70] then shows that the closed form GARCH(CFG) price can be obtained as:

$$\begin{aligned} c_{CFG} &= S_t\left(\frac{1}{2}+\frac{1}{\pi}\int_0^\infty Re\left[\frac{K^{-iz}f^*(iz+1)}{izf^*(1)}\right]dz\right) \\ &\qquad -Ke^{-rt}\left(\frac{1}{2}+\frac{1}{\pi}\int_0^\infty Re\left[\frac{K^{iz}f^*(iz)}{iz}\right]dz\right) \end{aligned} \tag{4.12}$$

Here f^* is the risk-neutral version of f.

4.2.4 Pure Jump Lévy models

The Lévy models we consider in this section assume that all possible movements in stock price are caused by jumps. The Lévy measure of such a process ensures frequent arrival of small jumps, so frequent that they render diffusion redundant, see Geman(2002)[62]. Hence they are known as pure jump processes. As an illustration in chapter2 and chapter3 we considered VG model to clarify how mathematics conforms with such an elegant intuition.

According to the Lévy-Kintchine formula the distribution of $X_{(t_2-t_1)} = log(\frac{S_{t_2}}{S_{t_1}})$ is characterized by the characteristic function of an infinitely divisible random variable given by:

$$\begin{aligned} &\mathbb{E}[e^{isX_{(t_2-t_1)}}] \\ &= \exp\left\{(t_2-t_1)\left[ias - \frac{1}{2}s^2b^2 + \int_{\mathbb{R}\setminus\{0\}}\left[e^{isx} - 1 - isx\mathbb{I}_{\{-1,1\}}(x)\right]\nu(dx)\right]\right\} \end{aligned} \tag{4.13}$$

where t_1 can naturally be zero. Scalars $a, b \in \mathbb{R}$ and the measure ν satisfies $\nu(\{0\}) = 0$ and $\int_{\mathbb{R}\setminus\{0\}}\left(|x|^2 \wedge 1\right)\nu(dx) < \infty$, which means that though numerous small jumps may not be integrable, square of those jumps are always integrable,a requirement which helps us extract a square integrable martingale process in the limit. In case of pure jump processes b is always zero. For further details see chapter2 and Cont et al(2004)[38], Kyprianou(2006)[76]. For example the variance gamma process characterizes the random variable X_1 through the parameters (σ, θ, γ) and the Lévy measure:

$$\nu_{vg}(dx) = \frac{1}{\gamma \mid x \mid} \exp\left(\frac{x\theta}{\sigma^2} - \frac{\mid x \mid}{\sigma}\sqrt{\frac{2}{\gamma} + \frac{\theta^2}{\sigma^2}}\right) dx \tag{4.14}$$

When integrated for jumps of all possible sizes (4.14) implies that the total rate is infinite,i.e. $\int_0^\infty \nu_{vg}(dx) = \infty$. However for any $\epsilon > 0$, we have $\int_\epsilon^\infty \nu_{vg}(dx) < \infty$, implying that it is small jumps which are numerous and jumps exceeding any threshold $\epsilon > 0$ are finite, arriving in compound Poisson fashion. The Lévy measure when used in (4.13) with $a = b = 0$ yields the following closed form characteristic function of the process X_t:

$$\Phi_{X_t}(s) = \left(\frac{1}{1 - is\theta\gamma + \frac{1}{2}s^2\sigma^2\gamma}\right)^{\frac{t}{\gamma}} \tag{4.15}$$

For this pure jump Lévy model the skewness and Kurtosis of log returns over an interval of length one is given by:

$$skew(X_1) = \frac{\theta\gamma(3\sigma^2 + 2\gamma\sigma^2)}{(\sigma^2 + \gamma\theta^2)^{\frac{3}{2}}} \tag{4.16}$$

$$Kurt(X_1) = 3\left(1 + 2\gamma - \frac{\gamma\sigma^4}{(\sigma^2 + \sigma\theta^2)^2}\right) \tag{4.17}$$

The risk-neutral version of the characteristic function (4.15) required in Carr Madan formula to price the options, see Carr and Madan(1999)[27], is given by:

$$\Phi_{X_t}^{(VG;rn)}(s) = \exp\left\{i\left[r + \frac{1}{\gamma}ln\left(1 - \theta\gamma - \frac{1}{2}\sigma^2\gamma\right)\right]st - \frac{t}{\gamma}ln\left(1 - is\theta\gamma + \frac{1}{2}s^2\sigma^2\gamma\right)\right\} \tag{4.18}$$

This risk-neutral form basically results from mean-correction of drift part, or introducing a drift to a driftless process, see Schouten(2003)[102].

The VG model has alternative characterization as difference of two Gamma processes, see Geman(2002)[62]. Using this characterization, VG model is generalized in Carr et al(2002)[26] which introduces an additional parameter and is known as CGMY model. In Carr et al(2002)[26], it has been shown that the success of VG model in explaining the smile effect of the market is likely due to the fact that the underlying process is pure jump with infinite activity and finite variation. The new parameter Y in CGMY model permits finite/infinite activity and finite/infinite variation. The CGMY process characterizes the distribution of X_1 through the parameters (C, G, M, Y) and the Lévy measure:

$$\nu_{cgmy}(dx) = \begin{cases} C\frac{e^{-Mx}}{x^{1+Y}}dx & \text{if } x > 0, \\ C\frac{e^{-Gx}}{|x|^{1+Y}}dx & \text{if } x < 0, \end{cases} \tag{4.19}$$

Here $C, G, M > 0$ and $Y < 2$. Apparently $Y = 0$ implies the VG model characterized as difference of two Gamma process. This Lévy measure when plugged in (4.13), with $a = b = 0$ provides a closed form characteristic function:

$$\Phi_{X_t}(s) = \exp\left\{Ct\Gamma(-Y)\left((M - is)^Y - M^Y + (G + is)^Y - G^Y\right)\right\} \tag{4.20}$$

Skewness and kurtosis of log-returns over an interval of length one is characterized by:

$$Skew(X_1) = \frac{C(M^{Y-3} - G^{Y-3})\Gamma(3 - Y)}{(C(M^{Y-2} + G^{Y-2})\Gamma(2 - Y))^{\frac{3}{2}}} \tag{4.21}$$

$$Kurt(X_1) = 3 + \frac{C(G^{Y-4} + M^{Y-4})\Gamma(4 - Y)}{(C(M^{Y-2} + G^{Y-2})\Gamma(2 - Y))^2} \tag{4.22}$$

When $G = M$, CGMY provides a symmetric model. For $G < M$ it provides a left skewed model often resembling features observed in market option data. Furthermore if $Y < 0$ the paths have finite jumps in any finite interval , otherwise the paths have infinitely many jumps in any finite interval. For $Y \in [1, 2)$ the process is of infinite variation. Finally the mean-corrected risk-neutral version, required for FFT based Carr Madan pricing, is given by:

$$\begin{aligned}\Phi_{X_t}^{(CGMY;rn)}(s) = \exp\Big\{& i\left(r - C\Gamma(-Y)\left((M-1)^Y - M^Y + (G+1)^Y - G^Y\right)\right)st \\ & + Ct\Gamma(-Y)\left((M - is)^Y - M^Y + (G + is)^Y - G^Y\right)\Big\}\end{aligned} \tag{4.23}$$

Another model of our consideration in pure jump category is Normal Inverse Gaussian(NIG). We saw that VG process can be interpreted as Brownian motion fluctuating not continuously but only at time points controlled by a Gamma Subordinator, so called business times. A similar interpretation holds for the NIG process that can be viewed as Brownian motion fluctuating only at Inverse Gaussian(IG) time. Intuitive interpretation of such time axis modelling by a subordinator is well documented in Geman(2002)[62], Clark(1973)[37].

The NIG process characterizes the random variable X_1 through the parameters (α, β, δ) and the Lévy measure:

$$\nu_{nig}(dx) = \frac{\delta\alpha}{\pi|x|} e^{\beta x} \mathbb{K}_1(\alpha|x|)dx \tag{4.24}$$

where $\mathbb{K}_1$ is a modified Bessel function of third kind with index 1. Like VG, NIG is an infinite activity process with numerous arrival of small jumps. Plugging this Lévy measure into (4.13) with $a = b = 0$; we obtain a closed form characteristic function:

$$\Phi_{X_t}(s) = \exp\left\{-\delta t(\sqrt{\alpha^2 - (\beta + is)^2} - \sqrt{\alpha^2 - \beta^2})\right\} \tag{4.25}$$

For NIG model the skewness and kurtosis, of log returns over an interval of length one are characterized by:

$$skew(X_1) = \frac{3\beta}{\alpha\delta^{-\frac{1}{2}}(\alpha^2 - \beta^2)^{-\frac{1}{4}}} \tag{4.26}$$

$$Kurt(X_1) = 3\left(1 + \frac{\alpha^2 + 4\beta^2}{\delta\alpha^2\sqrt{\alpha^2 - \beta^2}}\right) \tag{4.27}$$

We obtain the risk-neutral form of the characteristic function by mean correction:

$$\begin{aligned} \Phi_{X_t}^{(NIG;rn)}(s) &= \exp\left\{i\left(r + \delta(\sqrt{\alpha^2 - (\beta+1)^2} - \sqrt{\alpha^2 - \beta^2})\right) st \right. \\ &\qquad \left. -\delta(\sqrt{\alpha^2 - (\beta + is)^2} - \sqrt{\alpha^2 - \beta^2})\right\} \end{aligned} \tag{4.28}$$

4.2.5 A jump-diffusion model

We consider a jump diffusion model to examine the market response to diffusion combined with jumps in contrast to those with pure jump models. When jump diffusion models are a Lévy models, they are not pure jump because of the presence of diffusion. The choice of Kou's(2002)[75] double exponential model is motivated by the findings in Ramezani and Zeng(1999)[94], where it is suggested that double exponential jump-diffusion model fits the stock market data better than normal-diffusion model of Merton(1976)[85]. Kou assumes, see Kou(2002)[75], in addition to drifted diffusion the log-returns have occasional jumps following a double exponential distribution $DE(p, \eta_1, \eta_2)$. Here p is the probability of an upward jump and η_1 and η_2 govern the decay of the tails for the distribution of negative and positive jumps respectively. The Lévy measure is given by:

$$\nu(dx) = \left[p\lambda\eta_1 e^{-\eta_1 x}\mathbb{I}_{x<0} + (1-p)\lambda\eta_2 e^{-\eta_2|x|}\mathbb{I}_{x>0}\right] dx \tag{4.29}$$

where $\lambda = \int_{-\infty}^{\infty}\nu(dx) < \infty$, so unlike pure jump processes Kou's jump diffusion model is a finite activity model. The Lévy measure, through (4.13)(this time with non zero a and b), provides a closed form characteristic function:

$$\Phi_{X_t}(s) = \exp\left\{t\left(ias - \frac{1}{2}b^2s^2 + is\lambda\left[\frac{p\eta_1}{\eta_1 + is} + \frac{(1-p)\eta_2}{\eta_2 + is} - 1\right]\right)\right\} \tag{4.30}$$

The skewness in this model is not explicitly characterized. However Kou(2002)[75] suggests that the feature of heavier tails become more pronounced with the increase of either the jump size expectation($\frac{1}{\eta_j}$) or jump rate (λ). The mean-corrected characteristic function is obtained as:

$$\begin{aligned}\Phi_{X_t}^{(JD;rn)}(s) \quad = \quad & \exp\left\{i\left(r - \frac{1}{2}b^2 - \lambda\left[\frac{p\eta_1}{\eta_1+1} + \frac{(1-p)\eta_2}{\eta_2+1} - 1\right]\right) st\right. \\ & \left. -\frac{1}{2}b^2s^2t + is\lambda t\left[\frac{p\eta_1}{\eta_1+is} + \frac{(1-p)\eta_2}{\eta_2+is} - 1\right]\right\}\end{aligned} \tag{4.31}$$

We consider logarithm of the prices, $s_t = log(S_t)$ and $k = log(K)$ where K is the strike price of the option. As in Carr et al(1999)[27] the value of an European call with maturity T can be expressed as a function of k:

$$C_T(k) = \int_k^{\infty} e^{-rT}(e^s - e^k)q_T(s)ds \tag{4.32}$$

Here $q_T(s)$ is the risk-neutral density of the log prices. To ensure square integrability of $C_T(k)$ Carr and Madan, see Carr et al(1999)[27], introduced modified call prices:

$$c_T(k) = e^{\alpha k}C_T(k), \qquad \alpha > 0 \tag{4.33}$$

where α is known as the dampening factor. Following Carr et al(1999)[27] an analytic expression for the pricing formula (4.32) can be obtained as:

$$C_T(k) = \frac{e^{-\alpha k}}{\pi}\int_0^{\infty} e^{-iuk}\psi_T(u)du \tag{4.34}$$

where $\psi_T(u)$ has an analytic expression:

$$\psi_T(u) = \frac{e^{-rT}\Phi_T(u - (\alpha+1)i)}{\alpha^2 + \alpha - u^2 - i(2\alpha+1)u} \tag{4.35}$$

Here Φ is the characteristic function of the model for which prices need to be computed. In our empirical study we will consider Φ under the risk-neutral dynamics for all the considered Lévy models.

Brownian motion, being the simplest and the only continuous member of the Lévy family, can provide a closed form solution for European options:

Theorem 4.2 *Consider a European call option with strike price K and expiration time T. If the underlying option pays no dividends and continuously compounded risk-free rate is r, then the price of the contract at time t is given by:*

$$C(t, S_t) = S_t\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2) \tag{4.36}$$

where $\Phi(x)$ denotes the cumulative distribution function of standard normal random variable evaluated at the point x, $d_1 = \frac{[ln(\frac{S_t}{K})+(r+\frac{\sigma^2}{2})(T-t)]}{\sigma\sqrt{T-t}}$ and $d_2 = \frac{[ln(\frac{S_t}{K})+(r-\frac{\sigma^2}{2})(T-t)]}{\sigma\sqrt{T-t}}$ with $d_2 = d_1 - \sigma\sqrt{T-t}$.

For a proof see Black and Scholes(1973)[19].

4.3 Option Pricing and Delta-Gamma Approximation

For small fluctuations in underlying, option prices can be approximated using options Delta and Gamma. Inconsistency and pitfall in such approximations arise from big fluctuations of underlying, one point Delta and Gamma estimates, failure of Delta and Gamma to reflect true non-linearity in the pay-off of the option portfolio. Though Delta and Gamma are given in closed form only in few cases, Black-Scholes and Gram-Charlier in our case, in most cases we can obtain them upon numerical integration. Finite-difference technique can be applied to estimate all the Greeks reasonably quickly, see Duffy(2006)[48]. For the uniformity in comparison we will apply finite difference approach to all models of our consideration. Finite difference scheme of Greeks computation is extremely sensitive to the choice of amount of perturbation. For comparison this amount should remain same for all models. Perturbation chosen outside a particular range makes the Greek surfaces completely unstable and that range varies for different Greeks as well as models under consideration. To our knowledge there is no working rule to choose perturbation which works for all the models. We basically use trial and error approach to find a perturbation which works for all the models.

Suppose $C^{model}(S_t)$, for a particular model, be the price of an European option, when the price of the underlying is S_t. The Delta of that particular model can be obtained by finite difference method:

$$\delta^{model} = \frac{\partial C}{\partial S} = \frac{C^{model}(S + dS) - C^{model}(S)}{dS} \tag{4.37}$$

where dS is a small perturbation to the price of the underlying. Similarly to obtain Gamma, which measures the sensitivity of Delta, we need to obtain two values of Delta. Let δ_1 be the δ as defined in (4.37) and the δ_2 is:

$$\delta_2 = \frac{C^{model}(S + 2dS) - C^{model}(S + dS)}{dS} \tag{4.38}$$

Then the Gamma of a pricing model can be computed by finite difference:

$$\begin{aligned} \gamma^{model} &= \frac{\partial^2 \delta}{\partial S^2} \\ &= \frac{\delta^2 - \delta^1}{dS} \\ &= \frac{C^{model}(S + 2dS) - 2C^{model}(S + dS) + C^{model}(S)}{(dS)^2}. \end{aligned} \tag{4.39}$$

In figure 4.7 we plot the Delta surfaces for all the models under considerations. Delta changes dramatically when the option is close to ATM. For OTM option the delta converges to zero and for ITM all the Delta surfaces converge to one. Similar surfaces for Gamma are plotted in figure 4.8. Again for a short maturity option the Gamma changes dramatically when the option is close to ATM. However in case of Gamma the surfaces converge to zero for both ITM and OTM options.

The Delta Gamma approximations to model option prices for generic underlying asset price S, close to current price S_t, are given by:

$$C^{model}(S) \approx C^{model}(S_t) + \delta^{model}(S - S_t) \tag{4.40}$$

$$C^{model}(S) \approx C^{model}(S_t) + \delta^{model}(S - S_t) + \frac{1}{2}\gamma^{model}(S - S_t)^2 \tag{4.41}$$

See Dowd(2005)[42] and Christoffersen(2003)[31]. For any generic underlying asset price the option price is approximated using the same Delta and Gamma, which are calculated once(only for the current value of the underlying) no matter how deviated the future underlying asset prices are. The issue is though the continuity assumption of BS model can justify the fact that an ATM option will remain ATM for short maturity, in case of jumpy Lévy models as well as stochastic volatility models this is hardly justified. For models using non-normal distributions, tail events have substantial mass. Thus even in short time underlying can move significantly because of jumps and/or higher level of volatility rendering the approximation inconsistent. Since option portfolio is a linear combination of options, this inconsistency turns into pitfall in portfolio approximation. As mentioned in Christoffersen(2003)[31], in fact there is no alternative to the true valuation of the portfolio even for BS model. We investigate relative extent of pitfall in such approximation for market models of our consideration.

4.4 Comparative Look into Approximation Pitfall

We consider a portfolio similar to one used in Britten and Schaefer(1999)[22] but constructed from our data set. While the call options in the portfolio are traded in the market, the put option is priced using put-call parity. The option portfolio is described in table4.4. The portfolio in Britten and Schaefer(1999)[22] is used in Christoffersen(2003)[31], as well, to investigate the pitfall in approximation but in case of Black-Schole-Merton model only. So this section is an extension of similar analysis for Gram-Charlier, closed form GARCH and various Lev́y option pricing models including Kou's(2002)[75] double exponential jump-diffusion model.

We consider a risk-management horizon of five trading days(seven calender days), which corresponds to the sampling interval for our weekly data. As in Christoffersen(2003)[31] instead of computing the VaR's we will consider the complete pay-off profile of the portfolio–under all considered models–for different future values of the underlying asset prices S_{t+5}. However given that we are dealing with jumpy Lev́y models as well as stochastic volatility models, we consider a wider range of possible future values of the underlying in five trading days. Let P_t and $P(S_{t+5})$ denote the portfolio value today and at the end of five trading days respectively. We have:

$$P_t = m_1 * put + m_2 * Call_1 + m_3 * Call_2 \tag{4.42}$$

$$P^{\delta}(S_{t+5}) = P_t + \delta^p * (S_{t+5} - S_t) \tag{4.43}$$

$$P^{\delta\gamma}(S_{t+5}) = P_t + \delta^p * (S_{t+5} - S_t) + \frac{1}{2} * \gamma^p * (S_{t+5} - S_t)^2 \tag{4.44}$$

Here δ^p and γ^p are model dependent portfolio hedge factors defined as:

$$\delta^p = m_1 * \delta^m_{put} + m_2 * \delta^m_{call_1} + \delta^m_{call_2} \tag{4.45}$$

$$\gamma^p = m_1 * \gamma^m_{put} + m_2 * \gamma^m_{call_1} + \gamma^m_{call_2} \tag{4.46}$$

m indicates the model dependence; i.e. δ^p and γ^p will be different for different models.

True value of the portfolio is obtained through full-valuation of the option portfolio using model option prices:

$$\begin{aligned} P^{exact}(S_{t+5}) &= m_1 * put^m(K = 1200, T = 23 - 7) \\ &\quad + m_2 * Call^m_1(K = 1200, T = 23 - 7) \\ &\quad + m_3 * Call^m_2(K = 1550, T = 23 - 7) \end{aligned} \tag{4.47}$$

Each model m will have its own parameters to be used in pricing the options and evaluating hedging co-efficient. The pattern of non-linearity exhibited in figure4.5 is basically caused by the difference in strikes considered in the portfolio. Though for all models the approximations appear to be almost similar, in fact there are significant differences. Propagation of important and apparently more consistent, compared to the stark non-linearity, portion of the approximation errors are presented in figure4.6.

4.5 Data and Calibration

We consider options on S&P500 index traded on Wednesday 23rd January,2008. These are daily traded options of weekly record. After cleaning the data, see Bakshi et al(1997)[7], we have 178 options on that particular day traded in the market. For calibration we minimize the RMSE defined as:

$$RMSE = \frac{1}{mean\ price}\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(C_i^{market} - C_i^{model}\right)^2} \tag{4.48}$$

For cross-sectional assessment we report APE, in addition to RMSE, which is defined as:

$$APE = \frac{1}{mean\ price} \sum_{i=1}^{n} \frac{|C_i^{market} - C_i^{model}|}{n} \tag{4.49}$$

Table4.1 reports the calibration result for all considered models. After calibration, models pricing performance is investigated for different types of options belonging to different moneyness and maturity criteria. For out-of-sample assessment we consider options traded in the market on immediately next day, 24th of January 2008.

4.6 Empirical Analysis

Option models relative performance order in pricing need not necessarily be preserved in option portfolio approximation based on pricing models Delta and Gamma. This is because the definitions of Delta and Gamma, of a particular model, consider models parameters as constant and generate perturbation in option prices for a small perturbation only in underlying. We separately investigate the empirical observations for pricing and approximation.

4.6.1 Pricing performance with one day information

We observe that for calibration with most recent minimal data, as in single day traded options, Lévy models explicitly characterized by parameters modeling rate of decay on both tails fail to exhibit their true potential. e.g. four parameter CGMY and five parameter DE models are not performing better than three parameter VG and NIG models. Moreover we observe that when jump incorporating Lévy models bring moderate improvements in pricing performance it is stochastic volatility which brings more poignant improvement. Specifically Heston's SV and Heston Nandi CFG models show significant improvement over constant volatility Lévy models with jumps. This empirical comparison of time continuous Lévy and stochastic volatility approaches with discrete time GARCH volatility approach is first, to the best of our knowledge, in the literature.

Another relevant observation is that though GC model shows some improvement over BS model, it clearly falls behind the Lévy, GARCH and stochastic volatility models of our consideration. GC model though explicitly incorporates skewness and kurtosis, thus exhibiting pronounced smile-skew patterns, still embraces the assumption of constant volatility, a characteristic strongly contradicted in the market. On the other hand though stochastic volatility and GARCH volatility models response to smile and skew are less pronounced, they significantly outperform the GC model. The same conclusion holds when we consider Lévy models which exhibit pronounced smile skew patterns, e.g see CGMY and VG impled volatility graphs in figure4.3 and 4.4. This finding is in support of recent focus in the literature where Lévy innovations are blended in GARCH volatility structure. The approach is in-line with Heston-Nandi GARCH volatility structure but it replaces conditional normal innovation by Lévy innovations, or possibly GC innovation. Among other attractions this approach put together a remedy to volatility related imperfections with remedies to cross-maturity and cross-strike related biases.

The empirical features observed are based on entire data set we consider. Table4.2 shows models in-sample pricing performance for various categories of maturity and moneyness. Models relative performance observed in in-sample case has overall satisfactory correspondence in out-of sample pricing as well. In table4.3 we report RMSE and APE for different categories of options for out of sample assessment. In sample(left) and out-of-sample(right) APE for various categories are plotted in figure4.1.

4.6.2 Approximation performance with one day information

The non-linearity in approximation arises from particular choice of portfolio, namely the choice of strike of the options in option portfolio. Theoretical values of the portfolio(solid curves in figures4.5) exhibit little difference under different models. However the Delta and Delta-Gamma approximations to the portfolio exhibit various degree of proximity to the true portfolio valuation. These variation in approximations basically results from the use of Delta and Gamma which are estimated only current value of the underlying. It is partially caused by models response to true non-linearity of the portfolio as well as models sensitivity to ITM and OTM options in the portfolio .

For call options, in the portfolio, with strike 1550 and maturity 16 days, 23 calender days, the Delta and Gamma surfaces are plotted in figures4.7 and 4.8 respectively. As the figure4.7 shows near ATM when Delta changes more dramatically for Lévy and GC models, for stochastic volatility and GARCH models the changes are less dramatic. Consequently any risk-management model for option portfolio(with significant amount of ATM option), which relies on fixed Delta(and\or) Gamma estimates, are susceptible to be more misleading for Lévy and GC pricing models than those of stochastic volatility and GARCH volatility models. Our empirical findings reported in figure4.6 and table4.5 reveal such magnitude of mislead for effective market models.

Non-linearity of option prices is highest when the option is close to ATM. Figure4.8 shows that for Lévy and GC models the response to non-linearity is well captured by gamma, compare to stochastic volatility and GARCH models. We observed this evidence for other calibrations as well e.g. for calibrations not restricted to the use of minimal recent information. Consequently figure4.6 and table4.5 imply that in approximating option portfolio, pricing model's sensitivity to ITM and OTM options is more significant than it's response to non-linearity of the portfolio caused by ATM options in the portfolio.

Finally in figure4.9 we plot the risk-neutral densities of all the pricing models derived by inverting the corresponding characteristic functions. We used the parameters presented in table4.1 which are calibrated from options traded on 23rd January, 2008. The tails and peaks appear separately in figure4.10. Clearly stochastic volatility and GARCH volatility models exhibit distinct features in tails and peaks respectively.

4.7 Conclusion

In this chapter we consider comparative investigation of Gram-Charlier, GARCH and Lévy option pricing models from the perspective of a new investor willing to rely only on most recent minimum market information. Like other studies we found that pure jump Lévy processes with infinite activity and finite variation price the options better than the classical diffusions or jump-diffusion models. However we further observe that their performance are far less appealing when compared with GARCH volatility model as well as diffusion model combined with stochastic volatility. Though pure jump Lévy models can capture pronounced smile-skew patterns we observe that it is stochastic volatility model, even with less pronounced smile-skew patterns, which exhibits superior performance. Furthermore model with less pronounced smile-skew patterns combined with jump and diffusion, instead of stochastic volatility, performs much worse compare to pure jump Lévy models. Thus pure jump approach clearly have a preferential edge over diffusion and jump-diffusion models but not over stochastic volatility and/or GARCH volatility models. Furthermore pronounced smile-skew model combined with diffusion alone(GC model), without jump and stochastic volatility, exhibits worse performance compare to pure jump, stochastic volatility and GARCH volatility models.

In practice an option portfolio composed of options with various strikes lead to an acute non-linearity in portfolio pay-off, so acute that it becomes imperative to rely on full valuation of the portfolio over long period, no matter how skewness and kurtosis are being incorporated into the model. Various models response to that non-linearity is far from practical and reliable. Nonetheless in short period approximation may be found useful. Models performance in approximating option portfolio more importantly rely on their sensitivity to ITM and OTM options in the portfolio than their ability to capture the non-linearity of the portfolio caused by ATM options. All the approaches considered to incorporate skewness and kurtosis provide significant improvement over benchmark model in such approximation.

Our investigation indicates that blending conditional updates of volatility with conditional skewness and kurtosis might reflect the market reality better, which is a recent approach in the literature. In this approach GARCH structure of volatility updates is augmented with conditional Lévy innovations replacing conditional normal one. This will be our main focus in next chapters.

Model	RMSE	Parameters				
VG	0.1398	(σ)	(θ)	(ν)		
		0.1694	−0.6109	0.0343		
		(0.0020)	(0.0234)	(0.0023)		
NIG	0.1392	(α)	(β)	(δ)		
		64.4954	−41.7570	1.1825		
		(0.0262)	(0.0243)	(0.0182)		
DE	0.1464	(σ)	(λ)	(p)	(η_1)	(η_2)
		0.1900	0.3644	0.1183	13.2284	13.6686
		(0.0017)	(0.0499)	(0.0768)	(0.0916)	(0.0679)
CGMY	0.1452	(C)	(G)	(M)	(Y)	
		0.0772	7.1106	29.9656	1.3534	
		(0.0012)	(0.0063)	(0.0165)	(0.0044)	
CFG	0.0919	(α)	(β)	(ω)	(θ)	(σ^2)
		3.3794e-005	0.2500	2.2898e-005	0.500	0.0029
		(3.4e-06)	(1.6e-04)	(1.6e-04)	(3.4e-06)	(1.6e-04)
GC	0.1418	(σ)	(γ_1)	(γ_2)		
		0.2036	−0.3103	0.157		
		(0.0018)	(0.0337)	(0.5562)		
HS	0.0770	(κ)	(θ)	(σ)	(ρ)	(V_0)
		6.5460	0.0393	0.9287	−0.4196	0.1955
		(0.0393)	(0.0010)	(0.0025)	(0.0040)	(0.0196)
BS	0.1472	(σ)				
		0.1974				
		(0.0017)				

Table 4.1: *Model Calibration on 23rd Jan, 2008. The standard error of each parameter appears in brackets. To obtain the standard errors we numerically compute the Jacobian of mean squared error function for each model. Finite difference scheme is adopted for calculating partial derivatives.*

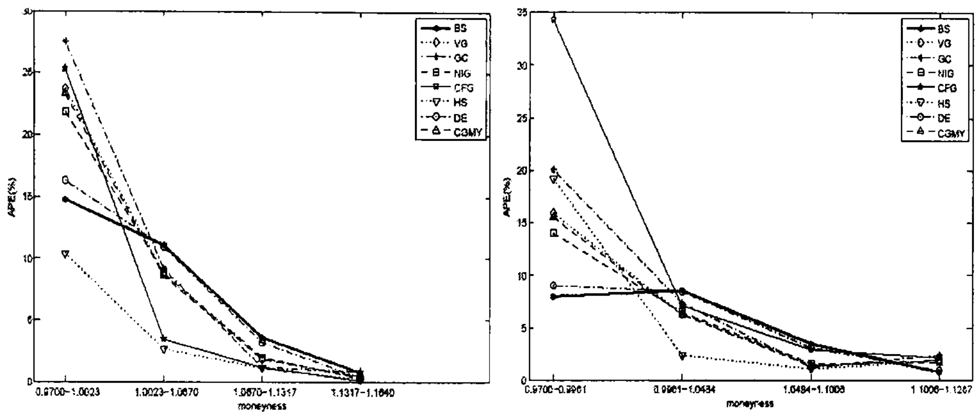
Model/moneyness		$dtm < 60$		$60 \le dtm \le 120$		$dtm > 120$	
		(APE)	(RMSE)	(APE)	(RMSE)	(APE)	(RMSE)
BS	[0.9700,1.0023)	14.7381	0.1761	5.4928	0.0640	8.5710	0.1069
	[1.0023,1.0670)	11.0965	0.1113	7.9715	0.0804	4.8842	0.0547
	[1.0670,1.1317)	3.5747	0.0396	4.3482	0.0450	2.7551	0.0343
	[1.1317,1.1640)	0.7634	0.0076	0.3480	0.0035	0.5749	0.0057
VG	[0.9700,1.0023)	23.6845	0.2508	6.9573	0.0744	9.2028	0.1197
	[1.0023,1.0670)	8.6847	0.0880	5.7726	0.0582	5.4350	0.0687
	[1.0670,1.1317)	1.8637	0.0215	2.3219	0.0258	4.0018	0.0488
	[1.1317,1.1640)	0.4273	0.0043	0.7177	0.0072	1.8879	0.0189
NIG	[0.9700,1.0023)	21.9036	0.2330	6.4282	0.0700	9.2617	0.1199
	[1.0023,1.0670)	8.8586	0.0896	5.8994	0.0595	5.4089	0.0679
	[1.0670,1.1317)	1.9946	0.0232	2.4824	0.0272	3.8766	0.0479
	[1.1317,1.1640)	0.3265	0.0033	0.6318	0.0063	1.7896	0.0179
CGMY	[0.9700,1.0023)	23.3115	0.2472	6.6727	0.0717	9.3591	0.1217
	[1.0023,1.0670)	8.5860	0.0870	5.6628	0.0572	5.4674	0.0696
	[1.0670,1.1317)	1.8451	0.0213	2.2825	0.0254	4.0690	0.0495
	[1.1317,1.1640)	0.4406	0.0044	0.7314	0.0073	1.9116	0.0191
DE	[0.9700,1.0023)	16.2895	0.1887	5.8654	0.0679	8.7540	0.1104
	[1.0023,1.0670)	10.9436	0.1099	7.7095	0.0777	5.1056	0.0585
	[1.0670,1.1317)	3.1877	0.0360	3.8732	0.0406	2.9592	0.0381
	[1.1317,1.1640)	0.3868	0.0039	0.0157	0.0002	1.0035	0.0100
GC	[0.9700,1.0023)	27.5277	0.2911	8.9009	0.0932	8.6363	0.1134
	[1.0023,1.0670)	9.1792	0.0944	6.2943	0.0635	5.4157	0.0670
	[1.0670,1.1317)	1.2034	0.0153	2.0068	0.0238	4.0332	0.0485
	[1.1317,1.1640)	0.8612	0.0086	1.0335	0.0103	2.1388	0.0214
CFG	[0.9700,1.0023)	25.3208	0.2764	3.9098	0.0451	2.2588	0.0266
	[1.0023,1.0670)	3.4560	0.0384	4.5257	0.0471	2.5114	0.0301
	[1.0670,1.1317)	1.1789	0.0134	3.1979	0.0333	1.3835	0.0140
	[1.1317,1.1640)	0.1284	0.0013	0.0177	0.0001	0.2878	0.0029
HS	[0.9700,1.0023)	10.4178	0.1165	5.3381	0.0606	3.4096	0.0399
	[1.0023,1.0670)	2.6392	0.0288	2.6525	0.0285	3.1102	0.0395
	[1.0670,1.1317)	1.1404	0.0133	2.1477	0.0228	4.2525	0.0489
	[1.1317,1.1640)	0.2377	0.0024	0.4194	0.0042	1.5825	0.0158

Table 4.2: *In-sample pricing performance on 23rd Jan,2008.*

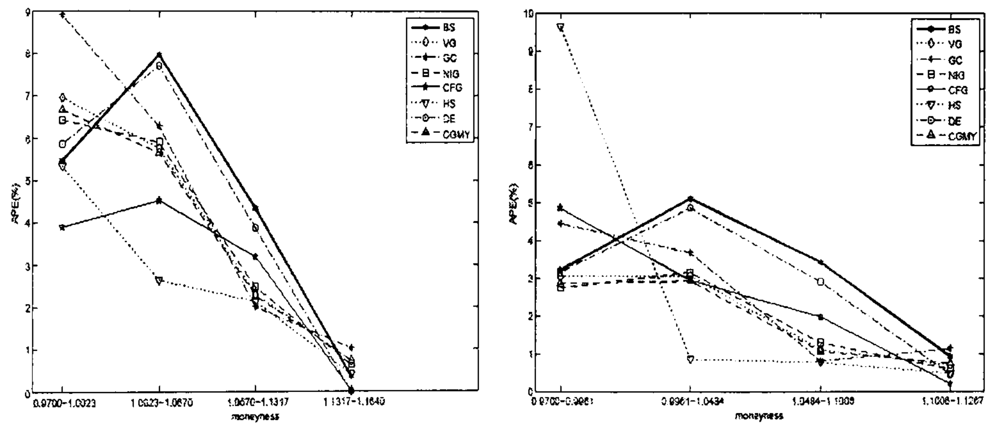
Model/moneyness		$dtm < 60$		$60 \leq dtm \leq 120$		$dtm > 120$	
		(APE)	(RMSE)	(APE)	(RMSE)	(APE)	(RMSE)
BS	[0.9700,0.9961)	7.9965	0.0973	3.2136	0.0367	18.3521	0.3508
	[0.9961,1.0484)	8.4986	0.0856	5.1048	0.0522	5.0215	0.0640
	[1.0484,1.1006)	3.4449	0.0374	3.4104	0.0341	3.3697	0.0419
	[1.1006,1.1267)	0.7420	0.0099	0.9100	0.0091	-	-
VG	[0.9700,0.9961)	15.8791	0.1717	3.0316	0.0363	19.0302	0.3488
	[0.9961,1.0484)	6.3289	0.0657	3.0457	0.0316	6.1418	0.0808
	[1.0484,1.1006)	1.3961	0.0146	1.1089	0.0111	4.7069	0.0575
	[1.1006,1.1267)	1.8330	0.0189	0.7022	0.0070	-	-
NIG	[0.9700,0.9961)	13.9849	0.1526	2.7344	0.0337	19.0829	0.3488
	[0.9961,1.0484)	6.4431	0.0663	3.1385	0.0325	6.0429	0.0800
	[1.0484,1.1006)	1.5502	0.0163	1.2902	0.0129	4.5772	0.0565
	[1.1006,1.1267)	1.7210	0.0178	0.5739	0.0057	-	-
CGMY	[0.9700,0.9961)	15.4815	0.1681	2.8410	0.0345	19.1794	0.3485
	[0.9961,1.0484)	6.2042	0.0644	2.9172	0.0304	6.2677	0.0819
	[1.0484,1.1006)	1.3650	0.0143	1.0605	0.0106	4.7781	0.0582
	[1.1006,1.1267)	1.8456	0.0190	0.7276	0.0073	-	-
DE	[0.9700,0.9961)	9.0330	0.1084	3.1394	0.0366	18.4779	0.3501
	[0.9961,1.0484)	8.4755	0.0854	4.8477	0.0495	5.2692	0.0686
	[1.0484,1.1006)	3.0369	0.0336	2.9081	0.0291	3.6197	0.0462
	[1.1006,1.1267)	0.9516	0.0119	0.4501	0.0045	-	-
GC	[0.9700,0.9961)	20.0068	0.2136	4.4467	0.0525	18.4013	0.3496
	[0.9961,1.0484)	7.3136	0.0781	3.6779	0.0379	5.7994	0.0785
	[1.0484,1.1006)	1.2652	0.0152	0.8125	0.0081	4.7123	0.0573
	[1.1006,1.1267)	2.3124	0.0233	1.1224	0.0112	-	-
CFG	[0.9700,0.9961)	34.2457	0.3618	4.8611	0.0565	14.2848	0.3786
	[0.9961,1.0484)	7.1414	0.0857	2.9455	0.0314	2.2500	0.0230
	[1.0484,1.1006)	2.8828	0.0352	1.9730	0.0197	1.6160	0.0179
	[1.1006,1.1267)	2.1397	0.0228	0.2038	0.0020	-	-
HS	[0.9700,0.9961)	19.1544	0.1956	9.6450	0.0640	16.9233	0.3908
	[0.9961,1.0484)	2.3641	0.0272	0.8576	0.0804	4.3251	0.0481
	[1.0484,1.1006)	1.0421	0.0128	0.7610	0.0450	5.4854	0.0630
	[1.1006,1.1267)	1.8298	0.0189	0.4693	0.0035	-	-

Table 4.3: *Out-of-sample performance for options traded on 24th January, 2008.*

(a) *Options with 0-60 days to maturity (dtm)*



(b)*Options with 60-120 days to maturity (dtm)*



(c)*Options with over 120 days to maturity (dtm)*

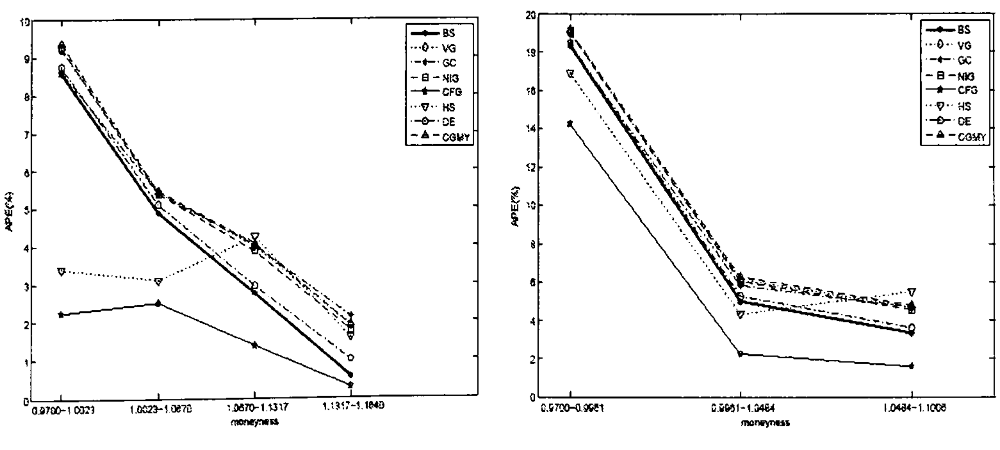


Figure 4.1: *Models Pricing Performance for Options traded on S&P500 Index on January 23 2008(In-sample, left) and January 24 2008(Out-of-sample, right)*

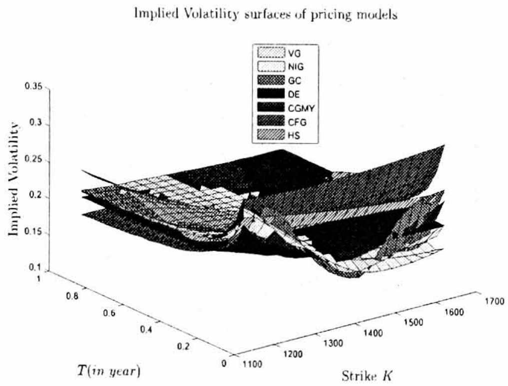


Figure 4.2: *Implied volatility surfaces of pricing models calibrated with the minimum information contained in one day traded options on 23rd of January, 2008.*

Type of Option:	Put	$Call_1$	$Call_2$
Strike(K_j):	1200	1200	1550
Maturity(days):	23	23	23
Option price:	4.2251	146.6	0.1750
Position(m_j):	-1	-1.5	2.5

Table 4.4: *Option portfolio constructed using the option traded on 23rd, January 2008. The current spot is 1338.6.*

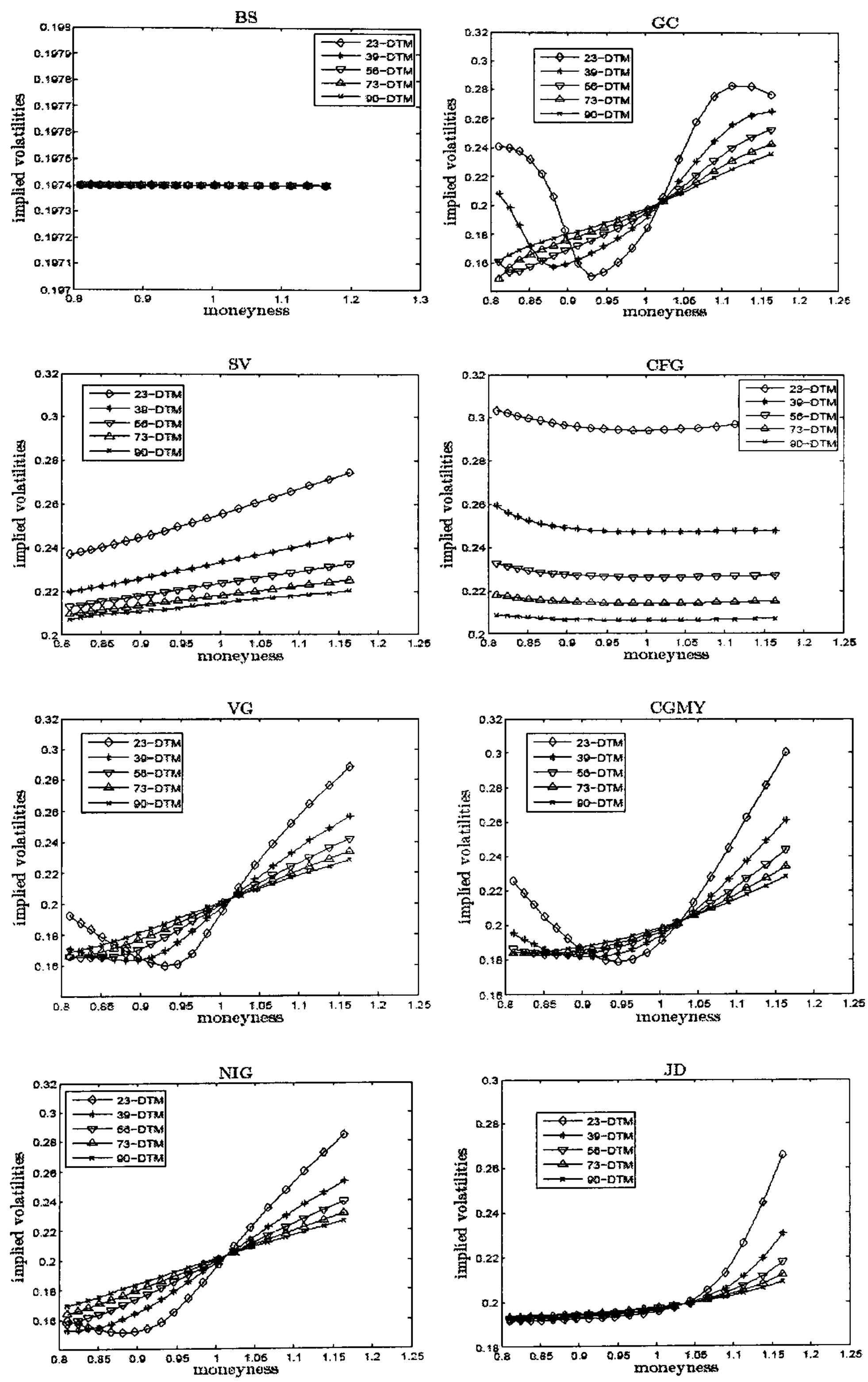


Figure 4.3: *Smile-Skew patterns exhibited by pricing models calibrated to S&P500 index options traded on January 23 2008: Short Maturities.*

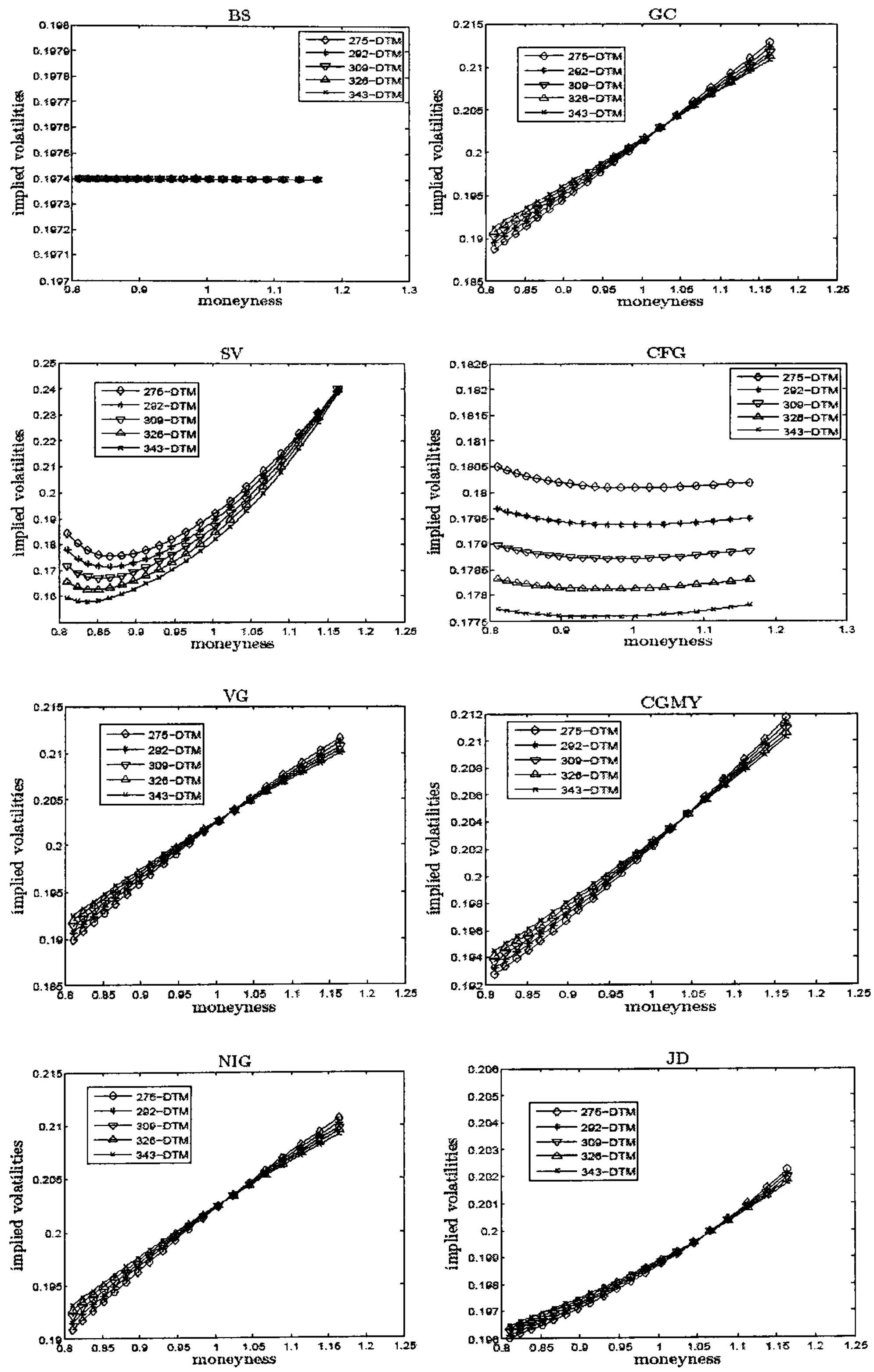


Figure 4.4: *Smile-Skew patterns exhibited by pricing models calibrated to S&P500 index options traded on January 23 2008: Long Maturities.*

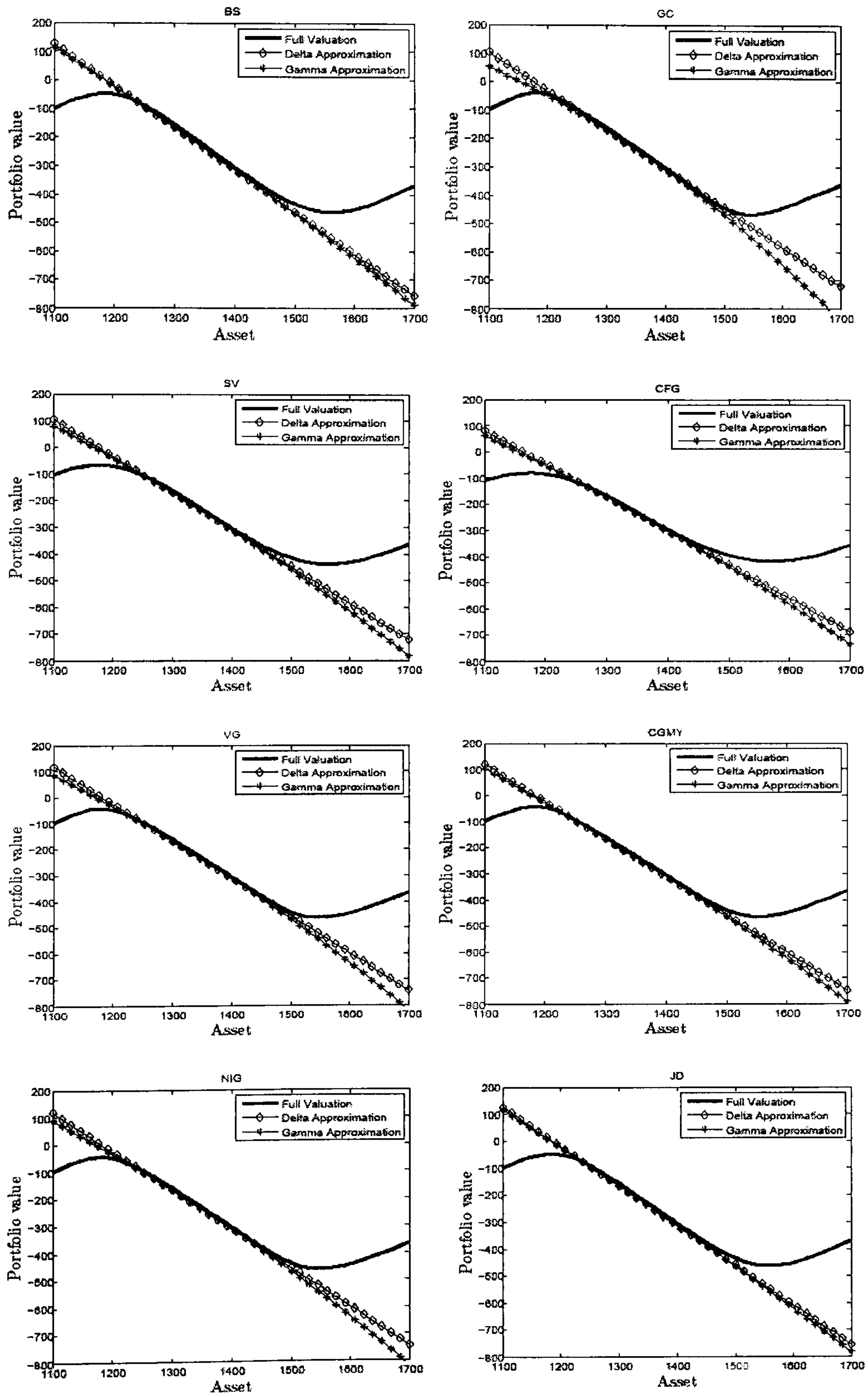


Figure 4.5: *Portfolio valuation: Full valuation vs. Greek approximations.*

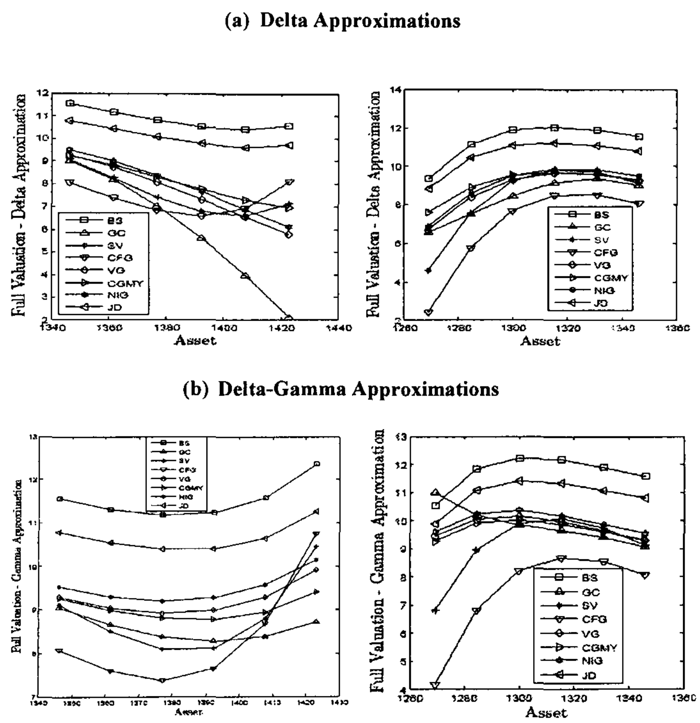


Figure 4.6: *Approximation errors of Greek-based valuations: Increasing(left) and decreasing(right) asset prices.*

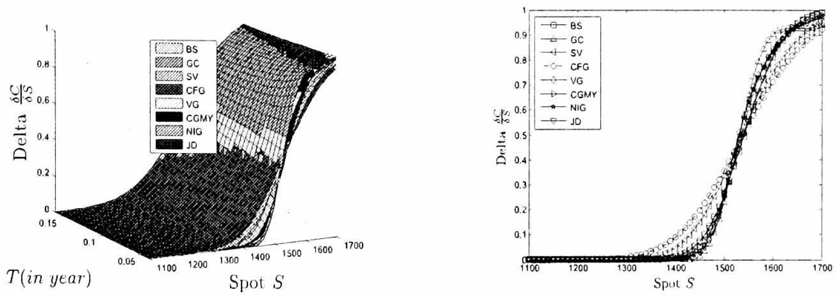


Figure 4.7: *Delta surfaces of different pricing models. On right hand side we take a slice corresponding to an option with 16 DTM and a strike of 1550.*

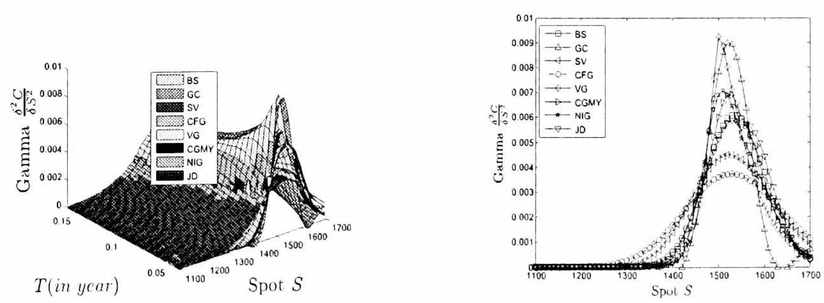


Figure 4.8: *Gamma surfaces of different pricing models. On right hand side we take a slice corresponding to an option with 16 DTM and a strike of 1550.*

Asset	Type	Models						
		(VG)	(NIG)	(CGMY)	(DE)	(GC)	(CFG)	(HS)
1300.0	Delta	21.8	20.0	19.8	6.57	28.8	35.5	22.4
	Gamma	17.1	15.4	18.0	6.75	20.0	33.0	19.1
1315.4	Delta	19.8	18.0	19.5	6.72	24.0	29.5	18.4
	Gamma	18.2	16.4	18.9	6.79	21.0	29.0	17.2
1330.8	Delta	19.1	17.3	19.4	6.73	21.2	28.3	18.3
	Gamma	19.0	17.1	19.3	6.74	21.0	28.2	18.2
1346.2	Delta	19.7	17.7	19.9	6.67	22.0	30.2	21.2
	Gamma	19.5	17.5	19.9	6.68	21.7	30.1	21.0
1361.5	Delta	21.8	19.4	21.1	6.65	26.8	33.7	26.0
	Gamma	20.0	17.5	20.4	6.72	23.2	33.0	25.0
1376.9	Delta	25.3	22.5	23.1	6.74	35.0	36.8	31.3
	Gamma	20.1	17.6	21.1	6.90	25.0	34.0	27.4

Table 4.5: *Percentage(%) reduction of approximation errors with respect to BS model. Current asset is 1338.6.*

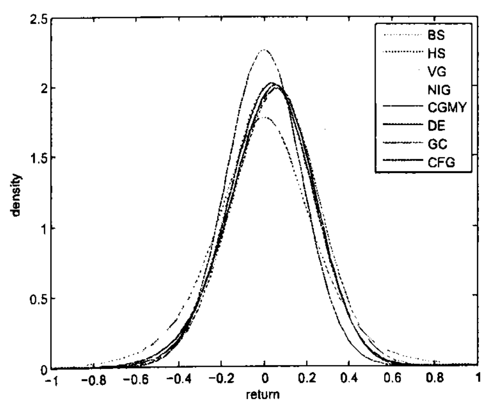


Figure 4.9: *Risk-neutral densities of pricing models calibrated with the minimum information contained in one day traded options on 23rd of January, 2008.*

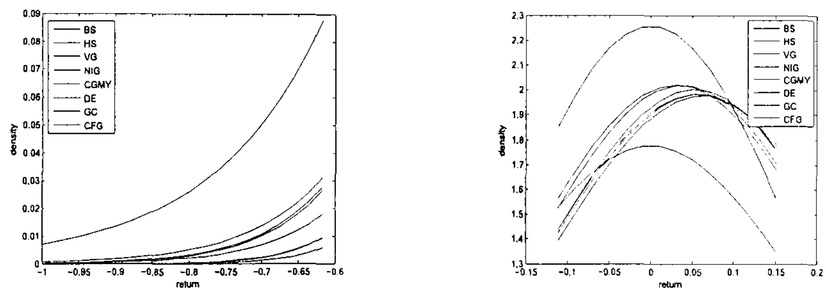


Figure 4.10: *Tails and peaks of risk-neutral densities of pricing models calibrated with the minimum information contained in one day traded options on 23rd of January, 2008.*

Chapter 5

Lévy Innovations to GARCH Model

In chapter two we discussed the main imperfections of Black-Scholes-Merton model. It is not able to capture the stylized facts such as skewness, heavy-tailedness and volatility clustering of real market data. As early as 1963 Mandelbrot(1963)[82] proposed to rely on α-stable distribution for modelling such stylized facts. But the infinite variance of this family of distributions is what makes the model questionable to researchers, as infinite variance yields a much heavier tail than one observed in the market. To circumvent this situation the density of the positively skewed $\alpha-$stable distribution– corresponds to those values of α satisfying $0 < \alpha < 1-$ is multiplied by an exponential function. This adaptation is know in the literature as tempered $\alpha-$stable distribution, see Tweedie (1984)[111]. Such an adaptation makes the tail thinner than the $\alpha-$stable distribution but heavier than the Gaussian one. This tempered stable family is again studied in recent time. It is applied by Kim et al.(2006)[74] to GARCH option pricing. However, as it is common for GARCH models, the price of an option needs to be computed using Monte-Carlo simulation. As discussed in chapter one, GARCH type models can capture the so called stylized facts. In general for derivative pricing what requires is the knowledge of the risk neutral distribution at maturity. But the problem is that for standard GARCH set up only one step ahead distribution is available. Heston and Nandi(2000)[70] proposed a way to overcome this difficulty. They proposed a GARCH-like model with normal innovation where they were able to compute the characteristic function of the underlying using a recursive procedure and then used the Heston(1993)[69] approach to price option using fourier inversion. However their model is not flexible enough to explain some observed option biases, specially when short maturity options are considered. Christoffersen (2006)[33] conjectured that this is due to the fact that single period innovations are Normal in Heston and Nandi(2000)[70]. Recently Badescu and Kulperger(2008)[5] also report the impracticalities with normal innovations in GARCH. In their work they considered semi-parametric pricing. Badescu and Kulperger(2008)[5] proposed a method in which they estimate GARCH parameters from historical stock returns and calculate the historical residuals from these estimates. They then proposed to use a purely non-parametric kernel smooth density, estimated from the historical residuals, to predict future innovations. The problem with their approach is that they are using time consuming simulation approach to price derivatives. Furthermore they proposed to use information contained only in stock prices and are not using market values of options in their calibration. Information contained in stock prices are known to be retrospective and are not forward looking where as information contained in option prices are know to be prospective, forward looking. This way their approach sounds contradictory to the usual approaches in the literature. As mentioned, they rely on simulation to price options; so they didn't find it feasible to apply their model on a long record of options and used only 120 records traded on a single day.

So from the point of view of practical implementation it is crucial to consider analytic approximation while exploring non-normal innovations to GARCH model. The new GARCH-like processes with Lévy innovations, GARCH-Lévy model could be a plausible name, are capable of capturing the conditional skewness and conditional kurtosis. However the promising side is that it is possible to obtain a recursive procedure for the evaluation of the characteristic function multi-period ahead which then yields the closed form prices, up to numerical integration, for European Derivatives.

5.1 Choice of Innovations

The GARCH-Lévy approach, specially with closed form valuation, is a very recent addition to the vast literature of asset pricing and risk-management. Only very recently Christoffersen, Elkamhi and Jacobs(2010) [32] introduce the broad characterizations of these dynamics. However they didn't present any empirical findings in their paper. Furthermore though their characterizations establish the broader perspectives to this very recent approach nonetheless they didn't offer explicit derivations for Lévy innovations which are recently much appreciated in derivative pricing literature. In a sequel Chyawat(2010) [89] considers affine GARCH dynamics with Lévy innovations and found that no matter how sophisticated affine relation is considered, for innovations coming from Lévy processes exhibiting both positive and negative jumps closed form pricing formula is not tractable. In this chapter we, however, provide the detailed mathematical underpinnings to uphold the explicit closed form valuation techniques similar to those of Heston and Nandi(2000)[70]–but we replace the conditional normal innovations by innovations coming from Lévy processes exhibiting both positive and negative jumps e.g. NIG, CGMY etc. For this we consider an approximation of the volatility dynamics. At first we, however, detailed the case with tempered stable(TS) Lévy innovations which is defined only on the positive half of the real line. As mentioned earlier, literature in this approach is not that rich. Some cases with simpler Lévy innovations, defined on positive half of the real line, are [1] available: Christoffersen(2006)[33] introduced Inverse Gaussian innovations, Bellini and Mercuri(2007)[16] considered Gamma innovations.

In chapter one we have discussed the GARCH feature and its pros and cons in derivative pricing. It incorporates time varying volatility and exhibits volatility clustering as observed in real market data. Furthermore it converges to stochastic volatility model of continuous time, see Duan(1995)[47]. This section revisits GARCH feature but not with conditional Gaussian innovations rather with innovations coming from those Lévy processes which have track record of fitting market data pretty well. We start by considering that the stock price

[1] Such processes are known as Subordinator, see chapter two.

process follows:

$$S_t = S_{t-1} e^{X_t}. \tag{5.1}$$

5.1.1 GARCH with Tempered Stable Lévy Innovations

We begin with a characterization of a $\alpha-$stable random variable. In the literature there are several characterizations available for this distribution, which respond to different computational needs. This distribution does not have any closed form density and often simulation based approach is utilized for statistical inference. Some computational facilities are available to generate standard $\alpha-$ stable variates, so we start by linking a general $\alpha-$stable variable with that of a standard $\alpha-$stable variable. For a standard random variable $X \sim S_\alpha(1, \beta, 0)$, the random variable

$$Y = \begin{cases} \sigma X + \mu & \alpha \neq 1 \\ \sigma X + \frac{2}{\pi}\beta\sigma \log\sigma + \mu & \alpha = 1 \end{cases}$$

is $S_\alpha(\sigma, \beta, \mu)$, where σ, μ and β are respectively scaling , location and skewness parameters and α is usually known as index which appears in the characteristic exponent. This rescaling property can be used for simulating general α-stable process. For general properties and different characterizations of stable distributions we refer to Zolotarev(1986)[115] and Zolotarev(1966)[116].

It is known that if $\beta = 1$ then the $\alpha-$stable distribution is positively skewed and when both $\beta = 1$ & $\mu = 0$ the density is defined on $[0, \infty)$, see Zolotarev(1986)[115]. Let us assume $S_\alpha(x; \sigma, \beta, \mu)$ denote the density function of the $\alpha-$stable distribution with $\alpha \in (0, 2]$. Zolotarev(1986)[115] showed that in positively skewed case the Laplace transform is given by:

$$\mathbb{E}(e^{-sX}) = \exp\left\{-s\mu - s^\alpha(\sigma)^\alpha \sec\left(\alpha\frac{\pi}{2}\right)\right\} \tag{5.2}$$

where $\sec(\theta)$ is the ordinary trigonometric function "Secant".

Now with a particular scaling let $S_\alpha\left(x; \frac{\gamma}{2^\alpha \sec\left(\alpha\frac{\pi}{2}\right)}, 1, 0\right)$ be the density of a positively skewed central $\alpha-$stable distribution for $\alpha \in (0, 1)$. Then the density of a tempered stable random variable, $TS_\alpha(\gamma, \eta)$ for $\gamma > 0$ and $\eta \geq 0$ is given by:

$$f(x; \alpha, \gamma, \eta) = e^{\gamma\eta} S_\alpha\left(x; \frac{\gamma}{2^\alpha \sec\left(\alpha\frac{\pi}{2}\right)}, 1, 0\right) e^{-\frac{1}{2}\eta^{\frac{1}{\alpha}} x} \tag{5.3}$$

The parameter η is responsible for a tempered tail making it thinner than $\alpha-$stable but heavier than Gaussian. For $\eta = 0$ tempering collapse and we recover the ordinary $\alpha-$stable distribution, where as for increasing η the tails gradually become thinner than those of the ordinary $\alpha-$stable. Making use of (5.2) with the density (5.3) we obtain the closed form characteristic function of the tempered stable random variable:

$$\mathbb{E}(e^{isX}) = \exp\left\{\gamma\eta - \gamma(\eta^{\frac{1}{\alpha}} - 2is)^{\alpha}\right\} = \exp\left\{\gamma\eta\left(1 - \left\{1 - 2is\eta^{-\frac{1}{\alpha}}\right\}^{\alpha}\right)\right\} \tag{5.4}$$

The unusual scaling of the $\alpha-$stable density, by the factor $\frac{\gamma}{2^{\alpha}\sec\left(\alpha\frac{\pi}{2}\right)}$, is required to obtain this closed from characteristic function for the tempered $\alpha-$ stable random variable.

Special Cases of Equation (5.4):

- Specifying the parameters as $\eta = (2\eta_1)^{\alpha}$, $\gamma = \frac{\gamma_1}{\alpha}$ we obtain $Gamma(\gamma_1, \eta_1)$ in the limit of $\alpha \longrightarrow 0$. To see this we note that:

$$\begin{aligned}
\mathbb{E}(e^{isX}) &= \exp\left\{\gamma\eta\left(1 - \left\{1 - 2is\eta^{-\frac{1}{\alpha}}\right\}^{\alpha}\right)\right\} \\
&= \left[\exp\left\{\eta\left(1 - \left\{1 - 2is\eta^{-\frac{1}{\alpha}}\right\}^{\alpha}\right)\right\}\right]^{\gamma} \\
&= \left[\frac{e^{\eta}}{e^{\eta\left\{1-2is\eta^{-\frac{1}{\alpha}}\right\}^{\alpha}}}\right]^{\gamma} \\
&= \left[\frac{e^{(2\eta_1)^{\alpha}}}{e^{(2\eta_1)^{\alpha}\left(\frac{\eta_1 - is}{\eta_1}\right)^{\alpha}}}\right]^{\gamma} \\
&= \left[\frac{e^{(2\eta_1)^{\alpha}}\left(\frac{\eta_1}{\eta_1 - is}\right)^{\alpha}}{e^{(2\eta_1)^{\alpha}\left(\frac{\eta_1 - is}{\eta_1}\right)^{\alpha}}\left(\frac{\eta_1}{\eta_1 - is}\right)^{\alpha}}\right]^{\gamma} \\
&= \frac{\left[\left\{e^{(2\eta_1)}\right\}^{\alpha}\right]^{\gamma}\left(\frac{\eta_1}{\eta_1 - is}\right)^{\gamma_1}}{\left[e^{\left\{(2\eta_1)\left(\frac{\eta_1 - is}{\eta_1}\right)\right\}^{\alpha}}\right]^{\gamma}\left[\left(\frac{\eta_1}{\eta_1 - is}\right)^{\alpha}\right]^{\gamma}}
\end{aligned} \tag{5.5}$$

Thus letting $\alpha \longrightarrow 0$ in (5.5), we obtain:

$$\mathbb{E}(e^{isX}) = \left(\frac{\eta_1}{\eta_1 - is}\right)^{\gamma_1} \tag{5.6}$$

Comparing (5.6) with (2.1) we see that it is the characteristic function of a $Gamma(\gamma_1, \eta_1)$. Hence Gamma distribution is a special case of tempered stable distribution and so will be the innovations coming from them.

- In (5.4) if we set $\alpha = \frac{1}{2}$ then we obtain:

$$\begin{aligned}
\mathbb{E}(e^{isX}) &= \exp\left\{\gamma\eta\left(1 - \frac{\sqrt{\{\eta^2 - 2is\}}}{\eta}\right)\right\} \\
&= \exp\left\{\gamma\left(\eta - \sqrt{\{\eta^2 - 2is\}}\right)\right\} \\
&= \exp\left\{-\gamma\left(\sqrt{\{\eta^2 - 2is\}} - \eta\right)\right\} \qquad (5.7)
\end{aligned}$$

  Comparing (5.7) with (2.3) we see that Inverse Gaussian distribution, $IG(\gamma, \eta)$, is a special case of tempered stable distribution ans so will be the innovations coming from them.

Let us now start by characterizing the log-return process X_t, appearing in equation (5.1) , with the follows the dynamics[2]:

$$X_t = r + \lambda\sigma_t - \frac{z_t}{2\sqrt{\alpha(1-\alpha)\gamma\eta^{\frac{\alpha-2}{\alpha}}}} \qquad (5.8)$$

Here the dynamic is under the real (physical) measure and the conditional distribution of innovations follows $z_t|\mathfrak{F}_{t-1} \sim TS_\alpha(\gamma\sigma_t, \eta)$ with the volatility processes following the GARCH(1,1) specification:

$$\sigma_t = \beta_0 + \beta_1 \frac{z_{t-1}}{2\sqrt{\alpha(1-\alpha)\gamma\eta^{\frac{\alpha-2}{\alpha}}}} + \alpha_1\sigma_{t-1} \qquad (5.9)$$

The scaling of innovations ensures the unit variance . It can be shown that, see Schouten(2003)[102] and the references there in, for a $X \sim TS_\alpha(\gamma, \eta)$ the moments are given by:

$$\begin{aligned}
\mathbb{E}[X] &= 2\alpha\gamma\eta^{\frac{\alpha-2}{\alpha}} && (5.10) \\
\mathbb{V}[X] &= 4\alpha(1-\alpha)\gamma\eta^{\frac{\alpha-2}{\alpha}} && (5.11) \\
\mathbb{S}kew[X] &= \frac{\alpha-2}{\sqrt{\alpha(1-\alpha)\gamma\eta}} && (5.12) \\
\mathbb{K}urt[X] &= 3 + \frac{4\alpha - 6 - \alpha(1-\alpha)}{\alpha(1-\alpha)\gamma\eta} && (5.13)
\end{aligned}$$

With these moment expressions for TS random variables, given $z_t|\mathfrak{F}_{t-1} \sim TS_\alpha(\gamma\sigma_t, \eta)$ the

[2]The way we scaled the innovation z_t ensures the unit variance of innovations. Requiring the innovations to have zero mean,in addition, will require the market price of risk λ to be adjusted accordingly.

corresponding conditional versions for the log-return process (5.36) can be obtained as:

$$
\begin{aligned}
\mathbb{E}[X_t \mid \mathfrak{F}_{t-1}] &= \mathbb{E}\left[r + \lambda\sigma_t - \frac{z_t}{2\sqrt{\alpha(1-\alpha)\gamma\eta^{\frac{\alpha-2}{\alpha}}}} \mid \mathfrak{F}_{t-1} \right] \\
&= r + \lambda\sigma_t - \frac{2\alpha\gamma\sigma_t\eta^{\frac{\alpha-1}{\alpha}}}{2\sqrt{\alpha(1-\alpha)\gamma\eta^{\frac{\alpha-2}{\alpha}}}} \\
&= r + \left(\lambda - \frac{\sqrt{\alpha\gamma\eta}}{\sqrt{1-\alpha}}\right)\sigma_t \qquad (5.14) \\
\mathbb{V}[X_t \mid \mathfrak{F}_{t-1}] &= \mathbb{V}\left[r + \lambda\sigma_t - \frac{z_t}{2\sqrt{\alpha(1-\alpha)\gamma\eta^{\frac{\alpha-2}{\alpha}}}} \mid \mathfrak{F}_{t-1} \right] \\
&= \frac{4\alpha(1-\alpha)\gamma\sigma_t\eta^{\frac{\alpha-2}{\alpha}}}{4\alpha(1-\alpha)\gamma\eta^{\frac{\alpha-2}{\alpha}}} \\
&= \sigma_t \qquad (5.15)
\end{aligned}
$$

Also since X_t, as in (5.36), is just a scaled and shifted version of $z_t|\mathfrak{F}_{t-1} \sim TS_\alpha(\gamma\sigma_t, \eta)$, it's conditional skewness and kurtosis can be obtained as:

$$
\begin{aligned}
\mathbb{S}kew[X_t|\mathfrak{F}_{t-1}] &= \frac{\alpha - 2}{\sqrt{\alpha(1-\alpha)\gamma\sigma_t\eta}} \qquad (5.16) \\
\mathbb{K}urt[X_t|\mathfrak{F}_{t-1}] &= 3 + \frac{4\alpha - 6 - \alpha(1-\alpha)}{\alpha(1-\alpha)\gamma\sigma_t\eta} \qquad (5.17)
\end{aligned}
$$

Thus the smile-skew patterns in implied volatilities, exhibited by market option prices could be modeled when we consider that the log return dynamics follow a GARCH with tempered stable innovations. Later we would like to explore similar settings for innovations coming from time changed Lévy processes. We can then compare whether market behavior is better captured by time changed Lévy innovations compare to the ordinary Lévy innovations.

Selecting GARCH-TS Equivalent Martingale Measure:

The perennial problem of selecting an appropriate EMM is always a concern when developing new dynamics for return process. For option pricing such a problem is explicitly treated by Gerber and Shiu(1994)[65] and its extension, known as conditional Esscher transform, is proposed by Buhlmann et al(1996) [23]. However in discrete time settings, for incomplete market, Shiu,Tong and Yang(2004)[105] show how to explore Esscher transform to price derivatives when only the conditional Moment Generating Function(MGF) is available. We start by recalling, see Gerber et al(1994)[65] and Shiu et al(2004)[105], that selecting an EMM basically depends on finding a solution of the conditional Esscher equation:

$$\frac{M_{X_t|\mathfrak{F}_{t-1}}(\hat{\theta}_t+1)}{M_{X_t|\mathfrak{F}_{t-1}}(\hat{\theta}_t)} = e^{r_t} \tag{5.18}$$

where $M_{X_t|\mathfrak{F}_{t-1}}(s)$ is the conditional moment generating function having definition:

$$M_{X_t|\mathfrak{F}_{t-1}}(s) = \mathbb{E}[e^{sX_t} \mid \mathfrak{F}_{t-1}] \tag{5.19}$$

For our GARCH dynamics with tempered stable innovations, (5.36), conditional Esscher equation (5.18) becomes:

$$\frac{\mathbb{E}_{t-1}\left[e^{(\hat{\theta}_t+1)\left(r+\lambda\sigma_t-\frac{z_t}{2\sqrt{\alpha(1-\alpha)\gamma\eta^{\frac{\alpha-2}{\alpha}}}}\right)}\right]}{\mathbb{E}_{t-1}\left[e^{\hat{\theta}_t\left(r+\lambda\sigma_t-\frac{z_t}{2\sqrt{\alpha(1-\alpha)\gamma\eta^{\frac{\alpha-2}{\alpha}}}}\right)}\right]} = e^{r_t}$$

$$\Longrightarrow \frac{\mathbb{E}_{t-1}\left[e^{-(\hat{\theta}_t+1)\left(\frac{z_t}{2\sqrt{\alpha(1-\alpha)\gamma\eta^{\frac{\alpha-2}{\alpha}}}}\right)}\right]}{\mathbb{E}_{t-1}\left[e^{-\hat{\theta}_t\left(\frac{z_t}{2\sqrt{\alpha(1-\alpha)\gamma\eta^{\frac{\alpha-2}{\alpha}}}}\right)}\right]} = e^{-\lambda\sigma_t} \tag{5.20}$$

Introducing the constant $c = \frac{-1}{2\sqrt{\alpha(1-\alpha)\gamma\eta^{\frac{\alpha-2}{\alpha}}}}$, (5.20) becomes:

$$\frac{\mathbb{E}_{t-1}\left[e^{(\hat{\theta}_t+1)z_t c}\right]}{\mathbb{E}_{t-1}\left[e^{\hat{\theta}_t z_t c}\right]} = e^{-\lambda\sigma_t} \tag{5.21}$$

Now applying equation (5.4) in (5.21) we obtain:

$$\begin{aligned} e^{-\lambda\sigma_t} &= \frac{\exp\left[\gamma\eta-\gamma(\eta^{\frac{1}{\alpha}}-2(1+\hat{\theta}_t)c)^{\alpha}\right]}{\exp\left[\gamma\eta-\gamma(\eta^{\frac{1}{\alpha}}-2\hat{\theta}_t c)^{\alpha}\right]} \\ &= \exp\left[\gamma((\eta^{\frac{1}{\alpha}}-2\hat{\theta}_t c)^{\alpha}-(\eta^{\frac{1}{\alpha}}-2(1+\hat{\theta}_t)c)^{\alpha})\right] \end{aligned} \tag{5.22}$$

Given a set of values of the parameters of $\alpha-$TS Lévy process, γ and η, and the GARCH volatility estimate σ_t, as in (5.9), the solution $\hat{\theta}_t$ of (5.22) can be used to describe the distribution of log-returns under EMM through (5.18). Unfortunately equation (5.22) needs to be solved numerically.

Now how to use $\hat{\theta}_t$ to describe the distribution of log returns? To answer this question we need to be familiar with the application of Conditional Esscher Transform(CET). The main application of Esscher transform lies in the fact that the moment generating function(conditional) of the log returns under EMM can be derived , using $\hat{\theta}_t$, as:

$$\tilde{M}_{X_s|\mathfrak{F}_{s-1}}(l) = \frac{M_{X_s|\mathfrak{F}_{s-1}}(l+\hat{\theta}_t)}{M_{X_s|\mathfrak{F}_{s-1}}(\hat{\theta}_t)} \tag{5.23}$$

With our assumption of distribution for innovations and volatility structure, equation (5.23) becomes:

$$\begin{aligned}
\tilde{M}_{X_s|\mathfrak{F}_{s-1}}(l) &= \frac{\mathbb{E}_{t-1}\left[e^{(\hat{\theta}_t+l)(r+\lambda\sigma_t+z_tc)}\right]}{\mathbb{E}_{t-1}\left[e^{\hat{\theta}_t(r+\lambda\sigma_t+z_tc)}\right]} \\
&= e^{l(r+\lambda\sigma_t)}\frac{\mathbb{E}_{t-1}\left[e^{(\hat{\theta}_t+l)cz_t}\right]}{\mathbb{E}_{t-1}\left[e^{\hat{\theta}_tcz_t}\right]} \\
&\overset{(5.4)}{=} e^{l(r+\lambda\sigma_t)}\frac{\exp\left[\gamma\eta-\gamma(\eta^{\frac{1}{\alpha}}-2(\hat{\theta}_t+l)c)^\alpha\right]}{\exp\left[\gamma\eta-\gamma(\eta^{\frac{1}{\alpha}}-2\hat{\theta}_tc)^\alpha\right]} \\
&= e^{l(r+\lambda\sigma_t)}\exp\left[\gamma(\eta^{\frac{1}{\alpha}}-2\hat{\theta}_tc)^\alpha-\gamma(\eta^{\frac{1}{\alpha}}-2(\hat{\theta}_t+l)c)^\alpha\right] \\
&= e^{l(r+\lambda\sigma_t)}\exp\left[\gamma(\eta^{\frac{1}{\alpha}}-2\hat{\theta}_tc)^\alpha-\gamma\left[(\eta^{\frac{1}{\alpha}}-2\hat{\theta}_tc)-2lc\right]^\alpha\right] \\
&= e^{l(r+\lambda\sigma_t)}\exp\left[\gamma(\eta^{\frac{1}{\alpha}}-2\hat{\theta}_tc)^\alpha\left\{1-\left[1-\frac{2lc}{(\eta^{\frac{1}{\alpha}}-2\hat{\theta}_tc)}\right]^\alpha\right\}\right] \\
&= e^{l(r+\lambda\sigma_t)}\exp\left[\gamma(\eta^{\frac{1}{\alpha}}-2\hat{\theta}_tc)^\alpha\left\{1-\left[1-2lc\left[(\eta^{\frac{1}{\alpha}}-2\hat{\theta}_tc)^\alpha\right]^{-\frac{1}{\alpha}}\right]^\alpha\right\}\right]
\end{aligned} \tag{5.24}$$

Comparing equation (5.24) with equation (5.4) we recognize that under the EMM the innovations are again TS distributed but with a parameter having new characterization $\eta' = (\eta^{\frac{1}{\alpha}} - 2\hat{\theta}_tc)^\alpha$. Because of this new characterization of a parameter of the model we need to check what other parameters of the entire settings require to be characterized anew.

Lets start with the dynamics of the volatility under the martingale measure:

$$\begin{aligned}
\sigma'_t = \tilde{\mathbb{V}}[X_t \mid \mathfrak{F}_{t-1}] &= \tilde{\mathbb{V}}\left[r + \lambda\sigma_t - \frac{z_t}{2\sqrt{\alpha(1-\alpha)\gamma\eta^{\frac{\alpha-2}{\alpha}}}} \mid \mathfrak{F}_{t-1}\right] \\
&= \frac{4\alpha(1-\alpha)\gamma\sigma_t[(\eta^{\frac{1}{\alpha}} - 2\hat{\theta}_t c)^{\alpha}]^{\left(\frac{\alpha-2}{\alpha}\right)}}{4\alpha(1-\alpha)\gamma\eta^{\frac{\alpha-2}{\alpha}}} \\
&= \left[\frac{\eta'}{\eta}\right]^{\left(\frac{\alpha-2}{\alpha}\right)} \sigma_t \qquad (5.25)
\end{aligned}$$

Now that the market and real measures are related through a new characterization $\eta' = (\eta^{\frac{1}{\alpha}} - 2\hat{\theta}_t c)^{\alpha}$ and market and real volatility processes are related through $\sigma'_t = \left[\frac{\eta'}{\eta}\right]^{\left(\frac{\alpha-2}{\alpha}\right)} \sigma_t$, we need to figure out what other parameters need to be characterized newly keeping the dynamics equivalent. Under real measure we have:

$$\begin{aligned}
X_t &= r + \lambda\sigma_t - \frac{z_t}{2\sqrt{\alpha(1-\alpha)\gamma\eta^{\frac{\alpha-2}{\alpha}}}} \\
\Rightarrow X_t &= r + \lambda\left[\frac{\eta}{\eta'}\right]^{\left(\frac{\alpha-2}{\alpha}\right)}\left[\frac{\eta'}{\eta}\right]^{\left(\frac{\alpha-2}{\alpha}\right)}\sigma_t - \frac{z_t}{2\sqrt{\alpha(1-\alpha)\gamma\eta^{\frac{\alpha-2}{\alpha}}}} \\
\Rightarrow X_t &= r + \lambda\left[\frac{\eta}{\eta'}\right]^{\left(\frac{\alpha-2}{\alpha}\right)}\sigma'_t - \frac{z_t}{2\sqrt{\alpha(1-\alpha)\gamma\eta^{\frac{\alpha-2}{\alpha}}}} \\
\Rightarrow X_t &= r + \lambda'\sigma'_t - \frac{z_t}{2\sqrt{\alpha(1-\alpha)\gamma\eta^{\frac{\alpha-2}{\alpha}}}} \\
&\qquad \left[\text{introducing new characterization} \lambda' = \lambda\left[\frac{\eta}{\eta'}\right]^{\left(\frac{\alpha-2}{\alpha}\right)}\right] \qquad (5.26)
\end{aligned}$$

For the equivalent dynamics to be characterized by martingale measure we need to introduce η' replacing η which can be done equivalently in the following way:

$$\begin{aligned}
\alpha(1-\alpha)\gamma\eta^{\frac{\alpha-2}{\alpha}} &= \alpha(1-\alpha)\gamma\eta'^{\frac{\alpha-2}{\alpha}}\left[\frac{\eta}{\eta'}\right]^{\left(\frac{\alpha-2}{\alpha}\right)} \\
&= \alpha(1-\alpha)\left[\gamma\left[\frac{\eta}{\eta'}\right]^{\left(\frac{\alpha-2}{\alpha}\right)}\right]\eta'^{\frac{\alpha-2}{\alpha}} \\
&= \alpha(1-\alpha)\gamma'\eta'^{\frac{\alpha-2}{\alpha}} \\
&\qquad \left[\text{introducing new characterization} \gamma' = \gamma\left[\frac{\eta}{\eta'}\right]^{\left(\frac{\alpha-2}{\alpha}\right)}\right] \qquad (5.27)
\end{aligned}$$

Thus finally we have the equivalent dynamics of log-returns, from (5.26), under the martingale measure:

$$X_t = r + \lambda'\sigma'_t - \frac{z_t}{2\sqrt{\alpha(1-\alpha)\gamma'\eta'^{\frac{\alpha-2}{\alpha}}}} \qquad \text{with } z_t|\mathfrak{F}_{t-1} \sim TS_\alpha(\gamma'\sigma'_t, \eta') \tag{5.28}$$

where parameters of the martingale dynamics, which ensures equivalence, are related with those of the market dynamics as:

$$\sigma'_t = \left[\frac{\eta'}{\eta}\right]^{\left(\frac{\alpha-2}{\alpha}\right)} \sigma_t \tag{5.29}$$

$$\eta' = (\eta^{\frac{1}{\alpha}} - 2\hat{\theta}_t c)^\alpha \tag{5.30}$$

$$\gamma' = \gamma\left[\frac{\eta}{\eta'}\right]^{\left(\frac{\alpha-2}{\alpha}\right)} \tag{5.31}$$

Now lets see the essential changes in GARCH parameters. We have the GARCH dynamics:

$$\begin{aligned} \sigma_t &= \beta_0 + \beta_1 \frac{z_{t-1}}{2\sqrt{\alpha(1-\alpha)\gamma\eta^{\frac{\alpha-2}{\alpha}}}} + \alpha_1\sigma_{t-1} \\ \Rightarrow \sigma_t\left[\frac{\eta'}{\eta}\right]^{\left(\frac{\alpha-2}{\alpha}\right)} &= \beta_0\left[\frac{\eta'}{\eta}\right]^{\left(\frac{\alpha-2}{\alpha}\right)} + \beta_1\left[\frac{\eta'}{\eta}\right]^{\left(\frac{\alpha-2}{\alpha}\right)} \frac{z_{t-1}}{2\sqrt{\alpha(1-\alpha)\gamma\eta^{\frac{\alpha-2}{\alpha}}}} + \alpha_1\left[\frac{\eta'}{\eta}\right]^{\left(\frac{\alpha-2}{\alpha}\right)}\sigma_{t-1} \\ & \left[\text{multiplying both sides by } \left[\frac{\eta'}{\eta}\right]^{\left(\frac{\alpha-2}{\alpha}\right)}\right] \end{aligned} \tag{5.32}$$

Thus the equivalent GARCH volatility dynamics under the martingale measure can be written as:

$$\sigma'_t = \beta'_0 + \beta'_1 \frac{z_{t-1}}{2\sqrt{\alpha(1-\alpha)\gamma'\eta'^{\frac{\alpha-2}{\alpha}}}} + \alpha_1\sigma'_{t-1} \qquad \text{with } z_t|\mathfrak{F}_{t-1} \sim TS_\alpha(\gamma'\sigma'_t, \eta') \tag{5.33}$$

where:

$$\beta'_0 = \beta_0\left[\frac{\eta'}{\eta}\right]^{\left(\frac{\alpha-2}{\alpha}\right)} \tag{5.34}$$

$$\beta'_1 = \beta_1\left[\frac{\eta'}{\eta}\right]^{\left(\frac{\alpha-2}{\alpha}\right)} \tag{5.35}$$

GARCH-TS risk-neutral characterization through market price of risk:

We have

$$X_{t+1} = r + \lambda\sigma_{t+1} - \frac{z_{t+1}}{2\sqrt{\alpha(1-\alpha)\gamma\eta^{\frac{\alpha-2}{\alpha}}}} \tag{5.36}$$

With $u = \frac{-1}{2\sqrt{\alpha(1-\alpha)\gamma\eta^{\frac{\alpha-2}{\alpha}}}}$, equation (5.36) becomes:

$$X_{t+1} = r + \lambda\sigma_{t+1} + uz_{t+1} \tag{5.37}$$

The m.g.f under risk-neutral dynamics can be obtained from the expression of the characteristic function:

$$\begin{aligned} E_t^Q[e^{uz_{t+1}}] &= E_t^Q[e^{i(-iu)z_{t+1}}] && (5.38)\\ &= \exp\left\{\gamma'\sigma'_{t+1}\eta'\left(1-\left\{1-2i(-iu)\eta'^{-\frac{1}{\alpha}}\right\}^{\alpha}\right)\right\} && (5.39) \end{aligned}$$

Thus:

$$\begin{aligned} E_t^Q[e^{X_{t+1}}] &= e^{r+\lambda'\sigma'_{t+1}}\exp\left\{\gamma'\sigma'_{t+1}\eta'\left(1-\left\{1-2i(-iu)\eta'^{-\frac{1}{\alpha}}\right\}^{\alpha}\right)\right\} && (5.40)\\ &= e^{r}\exp\left\{\sigma'_{t+1}\left[\lambda'-\gamma'\eta'\left(\left\{1-2u\eta'^{-\frac{1}{\alpha}}\right\}^{\alpha}-1\right)\right]\right\} && (5.41)\\ &= e^{r}\exp\left\{\sigma'_{t+1}\left[\lambda'-\gamma'\eta'\left[\left(1+\frac{1}{\sqrt{\gamma'\eta'\alpha(1-\alpha)}}\right)^{\alpha}-1\right]\right]\right\} && (5.42) \end{aligned}$$

the last equality follows by plugging the value of u.

We want to choose λ'(under Q) in terms of other parameters such that:

$$E_t^Q[e^{X_{t+1}}] = e^{(t+1-t)r} \tag{5.43}$$

That is:

$$\begin{aligned} & e^{r}\exp\left\{\sigma'_{t+1}\left[\lambda'-\gamma'\eta'\left[\left(1+\frac{1}{\sqrt{\gamma'\eta'\alpha(1-\alpha)}}\right)^{\alpha}-1\right]\right]\right\} = e^{r}\\ \Rightarrow\ & \sigma'_{t+1}\left[\lambda'-\gamma'\eta'\left[\left(1+\frac{1}{\sqrt{\gamma'\eta'\alpha(1-\alpha)}}\right)^{\alpha}-1\right]\right] = 0 && (5.44) \end{aligned}$$

Since σ'_{t+1} can not be zero, we must have:

$$\left[\lambda'-\gamma'\eta'\left[\left(1+\frac{1}{\sqrt{\gamma'\eta'\alpha(1-\alpha)}}\right)^{\alpha}-1\right]\right] = 0 \tag{5.45}$$

So

$$\lambda' = \gamma'\eta'\left[\left(1+\frac{1}{\sqrt{\gamma'\eta'\alpha(1-\alpha)}}\right)^{\alpha}-1\right] \tag{5.46}$$

5.2 GARCH with Time Changed Lévy Innovations

After revisiting time changed Lévy processes for option pricing, our focus in this section is to incorporate time changed Lévy innovations to GARCH(1,1) dynamics for derivative pricing. We will detail the mathematical underpinnings required to provide analytic GARCH dynamics for option pricing with innovations coming from tempered stable process, NIG process, VG process & CGMY process. Developing analytic valuation techniques for GARCH models which innovations follow Lévy processes with both sided jumps is the main focus of this chapter. However we will focus on the implementation of one of these models; namely one with GARCH-NIG dynamics.

5.2.1 GARCH with NIG Lévy Innovation

We start with the essential tool for our modeling ,the characteristic function. The characteristic function of a $NIG(\alpha, \beta, \delta)$ random variable is given by:

$$\mathbb{E}[e^{isX}] = \exp\left(-\delta\left\{\sqrt{\alpha^2 - (\beta + is)^2} - \sqrt{\alpha^2 - \beta^2}\right\}\right) \tag{5.47}$$

Using the intuitive development in section2.1.1 it follows that $NIG(\alpha, \beta, \delta)$ is infinitely divisible and the associated Lévy process has the distribution of increments over $[s, t + s]$ characterized by $NIG(\alpha, \beta, t\delta)$. Important moments of the $X \sim NIG(\alpha, \beta, \delta)$ random variable are given by:

$$\mathbb{E}[X] = \frac{\delta\beta}{\sqrt{\alpha^2 - \beta^2}} \tag{5.48}$$

$$\mathbb{V}[X] = \alpha^2\delta(\alpha^2 - \beta^2)^{\frac{-3}{2}} \tag{5.49}$$

$$\mathbb{S}kew[X] = 3\beta\alpha^{-1}\delta^{\frac{-1}{2}}(\alpha^2 - \beta^2)^{\frac{-1}{4}} \tag{5.50}$$

$$\tag{5.51}$$

$$\mathbb{K}urt[X] = 3\left(1 + \frac{\alpha^2 + 4\beta^2}{\delta\alpha^2\sqrt{\alpha^2 - \beta^2}}\right) \tag{5.52}$$

We refer to Schouten(2003)[102]. As usual we assume the stock price follows the dynamics (5.1), where , as in GARCH settings, the log return process now follows:

$$X_t = r + \lambda\sigma_t - \frac{z_t}{\sqrt{\alpha^2\delta(\alpha^2 - \beta^2)^{\frac{-3}{2}}}} \tag{5.53}$$

Here $z_t \mid \mathfrak{F}_{t-1} \sim NIG(\alpha, \beta, \delta\sigma_t)$ with the volatility processes,σ_t, following the GARCH(1,1) specification:

$$\sigma_t = \beta_0 + \beta_1 \frac{z_{t-1}}{\sqrt{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}}} + \alpha_1\sigma_{t-1} \tag{5.54}$$

Again the scaling ensures unit variance for innovations. With the moments of the NIG random variable as in (5.48)-(5.52), the conditional moments of the log-returns become:

$$\begin{aligned}
\mathbb{E}[X_t \mid \mathfrak{F}_{t-1}] &= \mathbb{E}\left[r + \lambda\sigma_t - \frac{z_t}{\sqrt{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}}} \mid \mathfrak{F}_{t-1}\right] \\
&= r + \lambda\sigma_t - \frac{\frac{\delta\sigma_t\beta}{\sqrt{\alpha^2-\beta^2}}}{\sqrt{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}}} \\
&= r + \left(\lambda - \frac{\beta\sqrt{\delta}}{\alpha(\alpha^2-\beta^2)^{\frac{-1}{4}}}\right)\sigma_t \qquad (5.55) \\
\mathbb{V}[X_t \mid \mathfrak{F}_{t-1}] &= \mathbb{V}\left[r + \lambda\sigma_t - \frac{z_t}{\sqrt{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}}} \mid \mathfrak{F}_{t-1}\right] \\
&= \frac{\alpha^2\delta\sigma_t(\alpha^2-\beta^2)^{\frac{-3}{2}}}{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}} \\
&= \sigma_t \qquad (5.56)
\end{aligned}$$

Similar to the case with GARCH-TS dynamics, since X_t, as in (5.53), is just a scaled and shifted version of $z_t \mid \mathfrak{F}_{t-1} \sim NIG(\alpha, \beta, \delta\sigma_t)$, it's conditional skewness and kurtosis can be obtained as:

$$\begin{aligned}
\mathbb{S}kew[X_t|\mathfrak{F}_{t-1}] &= 3\beta\alpha^{-1}(\delta\sigma_t)^{\frac{-1}{2}}(\alpha^2-\beta^2)^{\frac{-1}{4}} \qquad (5.57) \\
\mathbb{K}urt[X_t|\mathfrak{F}_{t-1}] &= 3\left(1 + \frac{\alpha^2+4\beta^2}{\delta\sigma_t\alpha^2\sqrt{\alpha^2-\beta^2}}\right) \qquad (5.58)
\end{aligned}$$

Existence of conditional skewness and conditional kurtosis ensures that smile-skew patterns could be modeled when we consider log-return dynamics following a GARCH with NIG-Lévy innovations, see Cfristoffersen(2003)[34].

Selecting a GARCH-NIG Equivalent Martingale Measure:

We follow exactly the same approach as explained in section 5.1.1 for TS Lévy innovations. That is we are interested in finding a solution, $\hat{\theta}_t$, from the conditional Esscher equation:

$$\frac{M_{X_t|\mathfrak{F}_{t-1}}(\hat{\theta}_t+1)}{M_{X_t|\mathfrak{F}_{t-1}}(\hat{\theta}_t)} = e^{rt} \tag{5.59}$$

where $M_{X_t|\mathfrak{F}_{t-1}}(s)$ is the conditional moment generating function defined as:

$$M_{X_t|\mathfrak{F}_{t-1}}(s) = \mathbb{E}[e^{sX_t} \mid \mathfrak{F}_{t-1}] \tag{5.60}$$

In case of GARCH dynamics with NIG -Lévy innovations (5.53), conditional Esscher equation (5.59) becomes:

$$\frac{\mathbb{E}_{t-1}\left[e^{(\hat{\theta}_t+1)\left(r+\lambda\sigma_t-\frac{z_t}{\sqrt{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}}}\right)}\right]}{\mathbb{E}_{t-1}\left[e^{\hat{\theta}_t\left(r+\lambda\sigma_t-\frac{z_t}{\sqrt{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}}}\right)}\right]} = e^{rt}$$

$$\Longrightarrow \frac{\mathbb{E}_{t-1}\left[e^{-(\hat{\theta}_t+1)\left(\frac{z_t}{\sqrt{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}}}\right)}\right]}{\mathbb{E}_{t-1}\left[e^{-\hat{\theta}_t\left(\frac{z_t}{\sqrt{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}}}\right)}\right]} = e^{-\lambda\sigma_t} \tag{5.61}$$

Introducing the constant $c = \frac{-1}{\sqrt{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}}}$, (5.61) becomes:

$$\frac{\mathbb{E}_{t-1}\left[e^{(\hat{\theta}_t+1)z_t c}\right]}{\mathbb{E}_{t-1}\left[e^{\hat{\theta}_t z_t c}\right]} = e^{-\lambda\sigma_t} \tag{5.62}$$

Using equation(5.47) in (5.62) we obtain:

$$\begin{aligned} e^{-\lambda\sigma_t} &= \frac{\exp\left(-\delta\left\{\sqrt{\alpha^2-(\beta+(\hat{\theta}_t+1)c)^2}-\sqrt{\alpha^2-\beta^2}\right\}\right)}{\exp\left(-\delta\left\{\sqrt{\alpha^2-(\beta+\hat{\theta}_t c)^2}-\sqrt{\alpha^2-\beta^2}\right\}\right)} \\ &= \exp\left(-\delta\left\{\sqrt{\alpha^2-(\beta+(\hat{\theta}_t+1)c)^2}-\sqrt{\alpha^2-(\beta+\hat{\theta}_t c)^2}\right\}\right) \end{aligned} \tag{5.63}$$

Again we need a numerical solution. Given the parameters of the NIG Lévy process, α, β, δ and the GARCH volatility estimate σ_t, as in (5.54), the solution ,$\hat{\theta}_t$, of (5.63) can be used to describe the distribution of log-returns as:

$$\tilde{M}_{X_s|\mathfrak{F}_{s-1}}(l) = \frac{M_{X_s|\mathfrak{F}_{s-1}}(l+\hat{\theta}_t)}{M_{X_s|\mathfrak{F}_{s-1}}(\hat{\theta}_t)} \tag{5.64}$$

With our assumption of distribution for innovations and the fact that one period ahead volatility is known in GARCH set up, equation (5.64) becomes:

$$\begin{aligned}
& \tilde{M}_{X_s|\mathfrak{F}_{s-1}}(l) \\
= \quad & \frac{\mathbb{E}_{t-1}\left[e^{(\hat{\theta}_t+l)(r+\lambda\sigma_t+z_t c)}\right]}{\mathbb{E}_{t-1}\left[e^{\hat{\theta}_t(r+\lambda\sigma_t+z_t c)}\right]}; \qquad z_t \mid \mathfrak{F}_{t-1} \sim NIG(\alpha,\beta,\delta\sigma_t) \\
= \quad & e^{l(r+\lambda\sigma_t)}\frac{\mathbb{E}_{t-1}\left[e^{(\hat{\theta}_t+l)cz_t}\right]}{\mathbb{E}_{t-1}\left[e^{\hat{\theta}_t cz_t}\right]} \\
\overset{(5.62)\&(5.63)}{=} \quad & e^{l(r+\lambda\sigma_t)}\exp\left(-\delta\left\{\sqrt{\alpha^2-(\beta+(\hat{\theta}_t+l)c)^2}-\sqrt{\alpha^2-(\beta+\hat{\theta}_t c)^2}\right\}\right) \\
= \quad & e^{l(r+\lambda\sigma_t)}\exp\left(-\delta\left\{\sqrt{\alpha^2-((\beta+\hat{\theta}_t c)+lc)^2}-\sqrt{\alpha^2-(\beta+\hat{\theta}_t c)^2}\right\}\right)
\end{aligned} \tag{5.65}$$

Comparing equations (5.47) and (5.65) we recognize that under EMM innovations are again NIG-distributed with a new characterization $\beta' = \beta + \hat{\theta}_t c$.

As in TS-innovations, we would like to see what other parameters of the entire settings are influenced by this new characterization. Let us start with the dynamics of the volatility under the martingale measure:

$$\begin{aligned}
\sigma_t' = \tilde{\mathbb{V}}[X_t \mid \mathfrak{F}_{t-1}] \quad &= \quad \tilde{\mathbb{V}}\left[r+\lambda\sigma_t - \frac{z_t}{\sqrt{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}}} \mid \mathfrak{F}_{t-1}\right] \\
&= \quad \frac{\alpha^2\delta\sigma_t(\alpha^2-\beta'^2)^{\frac{-3}{2}}}{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}} \\
&= \quad \left[\frac{\alpha^2-\beta'^2}{\alpha^2-\beta^2}\right]^{\frac{-3}{2}}\sigma_t
\end{aligned} \tag{5.66}$$

So the market and real measures are related through a new characterization $\beta' = \beta + \hat{\theta}_t c$ and market and real volatility processes are related through $\sigma_t' = \left[\frac{\alpha^2-\beta'^2}{\alpha^2-\beta^2}\right]^{\frac{-3}{2}}\sigma_t$. We need to figure out what other parameters require new characterization in order to keep the return dynamics equivalent. Under real measure we have:

$$\begin{aligned}
X_t &= r + \lambda\sigma_t - \frac{z_t}{\sqrt{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}}} \\
\Rightarrow X_t &= r + \lambda\left[\frac{\alpha^2-\beta^2}{\alpha^2-\beta'^2}\right]^{\frac{-3}{2}}\left[\frac{\alpha^2-\beta'^2}{\alpha^2-\beta^2}\right]^{\frac{-3}{2}}\sigma_t - \frac{z_t}{\sqrt{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}}} \\
\Rightarrow X_t &= r + \lambda\left[\frac{\alpha^2-\beta^2}{\alpha^2-\beta'^2}\right]^{\frac{-3}{2}}\sigma'_t - \frac{z_t}{\sqrt{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}}} \\
\Rightarrow X_t &= r + \lambda'\sigma'_t - \frac{z_t}{\sqrt{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}}} \\
&\qquad\left[\text{introducing new characterization}\lambda' = \lambda\left[\frac{\alpha^2-\beta^2}{\alpha^2-\beta'^2}\right]^{\frac{-3}{2}}\right]
\end{aligned} \tag{5.67}$$

For the dynamics to be characterized by martingale measure we need to introduce α', β', δ' replacing α, β, δ. Maintaining the equivalence this can be accomplished in the following way:

$$\begin{aligned}
\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}} &= \alpha^2\delta\left[\frac{\beta^2}{\beta'^2}\right]^{\frac{-3}{2}}\left(\frac{\beta'^2}{\beta^2}\alpha^2-\beta'^2\right)^{\frac{-3}{2}} \\
&= \left(\alpha^2\frac{\beta'^2}{\beta^2}\right)\left(\delta\frac{\beta'}{\beta}\right)\left(\frac{\beta'^2}{\beta^2}\alpha^2-\beta'^2\right)^{\frac{-3}{2}} \\
&= \alpha'^2\delta'(\alpha'^2-\beta'^2)^{\frac{-3}{2}} \\
&\qquad\left[\text{introducing new characterization}\alpha' = \alpha\frac{\beta'}{\beta}\&\delta' = \delta\frac{\beta'}{\beta}\right]
\end{aligned} \tag{5.68}$$

Thus finally we have the equivalent dynamics for log-returns, from (5.67), under the martingale measure:

$$X_t = r + \lambda'\sigma'_t - \frac{z_t}{\sqrt{\alpha'^2\delta'(\alpha'^2-\beta'^2)^{\frac{-3}{2}}}} \qquad \text{with } z_t \mid \mathfrak{F}_{t-1} \sim NIG(\alpha',\beta',\delta'\sigma'_t) \tag{5.69}$$

where the parameters of the martingale dynamics, maintaining equivalence, are related with those of the market dynamics through:

$$\sigma'_t = \left[\frac{\alpha^2-\beta'^2}{\alpha^2-\beta^2}\right]^{\frac{-3}{2}} \sigma_t \tag{5.70}$$

$$\beta' = \beta + \hat{\theta}_t c \tag{5.71}$$

$$\alpha' = \alpha\frac{\beta'}{\beta} \tag{5.72}$$

$$\delta' = \delta\frac{\beta'}{\beta} \tag{5.73}$$

As last step we need to figure out the essential changes in GARCH parameters. We have the GARCH dynamics:

$$\begin{aligned}
\sigma_t &= \beta_0 + \beta_1 \frac{z_{t-1}}{\sqrt{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}}} + \alpha_1\sigma_{t-1} \\
\Rightarrow \sigma_t \left[\frac{\alpha^2-\beta'^2}{\alpha^2-\beta^2}\right]^{\frac{-3}{2}} &= \beta_0 \left[\frac{\alpha^2-\beta'^2}{\alpha^2-\beta^2}\right]^{\frac{-3}{2}} + \beta_1 \left[\frac{\alpha^2-\beta'^2}{\alpha^2-\beta^2}\right]^{\frac{-3}{2}} \frac{z_{t-1}}{\sqrt{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}}} \\
&\quad + \alpha_1 \left[\frac{\alpha^2-\beta'^2}{\alpha^2-\beta^2}\right]^{\frac{-3}{2}} \sigma_{t-1}
\end{aligned} \tag{5.74}$$

$$\left[\text{multiplying both sides by } \left[\frac{\alpha^2-\beta'^2}{\alpha^2-\beta^2}\right]^{\frac{-3}{2}}\right]$$

Thus the equivalent GARCH volatility dynamics under the martingale measure can be written as:

$$\sigma'_t = \beta'_0 + \beta'_1 \frac{z_{t-1}}{\sqrt{\alpha'^2\delta'(\alpha'^2-\beta'^2)^{\frac{-3}{2}}}} + \alpha_1\sigma'_{t-1} \qquad \text{with } z_t|\mathfrak{F}_{t-1} \sim NIG(\alpha', \beta', \delta'\sigma'_t) \tag{5.75}$$

where:

$$\beta'_0 = \beta_0 \left[\frac{\alpha^2-\beta'^2}{\alpha^2-\beta^2}\right]^{\frac{-3}{2}} \tag{5.76}$$

$$\beta'_1 = \beta_1 \left[\frac{\alpha^2-\beta'^2}{\alpha^2-\beta^2}\right]^{\frac{-3}{2}} \tag{5.77}$$

GARCH-NIG risk-neutral characterization through market price of risk:

With the scale factor $u = -\frac{1}{\sqrt{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}}}$ equation (5.53) becomes:

$$X_t = r + \lambda\sigma_t + uz_t. \tag{5.78}$$

The m.g.f under risk-neutral measure can be obtained from the expression of the NIG-characteristic function of NIG distribution:

$$\begin{aligned}
E_t^Q[e^{uz_{t+1}}] &= E_t^Q[e^{i(-iu)z_{t+1}}] \\
&= \exp\left(-\delta'\sigma'_{t+1}\left\{\sqrt{\alpha'^2-(\beta'+i(-iu))^2}-\sqrt{\alpha'^2-\beta'^2}\right\}\right) \\
&= \exp\left(-\delta'\sigma'_{t+1}\left\{\sqrt{\alpha'^2-(\beta'+u)^2}-\sqrt{\alpha'^2-\beta'^2}\right\}\right) \qquad (5.79)
\end{aligned}$$

Thus:

$$\begin{aligned}
E_t^Q[e^{X_{t+1}}] &= e^{(r+\lambda'\sigma'_{t+1})}\exp\left\{-\delta'\sigma'_{t+1}\left\{\sqrt{\alpha'^2-(\beta'+u)^2}-\sqrt{\alpha'^2-\beta'^2}\right\}\right\} \\
&= e^{(r+\lambda'\sigma'_{t+1})}\exp\left\{-\delta'\sigma'_{t+1}\left\{\sqrt{\alpha'^2-\left(\beta'-\frac{1}{\sqrt{\alpha'^2\delta'(\alpha'^2-\beta'^2)^{\frac{-3}{2}}}}\right)^2}-\sqrt{\alpha'^2-\beta'^2}\right\}\right\}
\end{aligned} \qquad (5.80)$$

We want to choose λ', under Q, in terms of other risk-neutral parameters such that:

$$\begin{aligned}
& E_t^Q[e^{X_{t+1}}] = e^{(t+1-t)r} \\
\Rightarrow \quad & e^{(r+\lambda'\sigma'_{t+1})}\exp\left\{-\delta'\sigma'_{t+1}\left\{\sqrt{\alpha'^2-\left(\beta'-\frac{1}{\sqrt{\alpha'^2\delta'(\alpha'^2-\beta'^2)^{\frac{-3}{2}}}}\right)^2}-\sqrt{\alpha'^2-\beta'^2}\right\}\right\} = e^r \\
\Rightarrow \quad & e^r\exp\left\{\sigma'_{t+1}\left(\lambda'-\delta'\left\{\sqrt{\alpha'^2-\left(\beta'-\frac{1}{\sqrt{\alpha'^2\delta'(\alpha'^2-\beta'^2)^{\frac{-3}{2}}}}\right)^2}-\sqrt{\alpha'^2-\beta'^2}\right\}\right)\right\} = e^r
\end{aligned} \qquad (5.81)$$

Since $\sigma'_{t+1} \neq 0$, we must have:

$$\lambda' = \delta'\left\{\sqrt{\alpha'^2-\left(\beta'-\frac{1}{\sqrt{\alpha'^2\delta'(\alpha'^2-\beta'^2)^{\frac{-3}{2}}}}\right)^2}-\sqrt{\alpha'^2-\beta'^2}\right\} \qquad (5.82)$$

This is the final characterization which we will be using in the expression of mgf, which in turn will be used in our pricing(and hence in calibration).

5.2.2 GARCH with CGMY innovation

As in the case with NIG innovation. We start with characteristic function of the *CGMY(C,G,M,Y)* process:

$$\mathbb{E}[e^{iuX_1}] = \exp\left\{C\Gamma(-Y)((M-iu)^Y - M^Y + (G+iu)^Y - G^Y)\right\} \qquad (5.83)$$

We recall the intuitive development in section2.1.1 which shows that $CGMY(C,G,M,Y)$ is infinitely divisible and the associated Lévy process has the distribution of increments over $[s,t+s]$ characterized by $CGMY(tC,G,M,Y)$:

$$\begin{aligned}\mathbb{E}[e^{iuX_t}] &= \left[\exp\left\{C\Gamma(-Y)((M-iu)^Y - M^Y + (G+iu)^Y - G^Y)\right\}\right]^t \\ &= \exp\left\{tC\Gamma(-Y)((M-iu)^Y - M^Y + (G+iu)^Y - G^Y)\right\} \qquad (5.84)\end{aligned}$$

For future reference we report the important moments of $X \sim CGMY(C,G,M,Y)$ random variable, see Schouten(2003)[102]:

$$\mathbb{E}[X] = C(M^{Y-1} - G^{Y-1})\Gamma(1-Y) \qquad (5.85)$$

$$\mathbb{V}[X] = C(M^{Y-2} + G^{Y-2})\Gamma(2-Y) \qquad (5.86)$$

$$\mathbb{S}kew[X] = \frac{C(M^{Y-3} - G^{Y-3})\Gamma(3-Y)}{\left\{C(M^{Y-2} + G^{Y-2})\Gamma(2-Y)\right\}^{\frac{3}{2}}} \qquad (5.87)$$

$$\mathbb{K}urt[X] = 3 + \frac{C(M^{Y-4} - G^{Y-4})\Gamma(4-Y)}{\left\{C(M^{Y-2} + G^{Y-2})\Gamma(2-Y)\right\}^2} \qquad (5.88)$$

Similar to GARCH-TS or GARCH-NIG cases, we assume the stock price follows the dynamics (5.1), where according to the GARCH settings, the log return process now follows:

$$X_t = r + \lambda\sigma_t - \frac{z_t}{\sqrt{C(M^{Y-2} + G^{Y-2})\Gamma(2-Y)}} \qquad (5.89)$$

Here $z_t \mid \mathfrak{F}_{t-1} \sim CGMY(C\sigma_t, G, M, Y)$ with the volatility processes,σ_t, following the GARCH(1,1) specification:

$$\sigma_t = \beta_0 + \beta_1\frac{z_{t-1}}{\sqrt{C(M^{Y-2} + G^{Y-2})\Gamma(2-Y)}} + \alpha_1\sigma_{t-1} \qquad (5.90)$$

Similar to the volatility dynamics with GARCH-TS, GARCH-NIG cases here the scaling ensures unit variance for innovations. With the moments of the CGMY random variable as in (5.85)-(5.88), the conditional moments of the log-returns become:

$$\begin{aligned}
\mathbb{E}[X_t \mid \mathfrak{F}_{t-1}] &= \mathbb{E}\left[r + \lambda\sigma_t - \frac{z_t}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}} \mid \mathfrak{F}_{t-1}\right] \\
&= r + \lambda\sigma_t - \frac{C\sigma_t(M^{Y-1}-G^{Y-1})\Gamma(1-Y)}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}} \\
&= r + \left(\lambda - \frac{C(M^{Y-1}-G^{Y-1})\Gamma(1-Y)}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}}\right)\sigma_t \qquad (5.91)\\
\mathbb{V}[X_t \mid \mathfrak{F}_{t-1}] &= \mathbb{V}\left[r + \lambda\sigma_t - \frac{z_t}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}} \mid \mathfrak{F}_{t-1}\right] \\
&= \frac{C\sigma_t(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)} \\
&= \sigma_t \qquad (5.92)
\end{aligned}$$

Similar to the dynamics with other innovations since X_t– as in (5.89)– is just a scaled and shifted version of $z_t \mid \mathfrak{F}_{t-1} \sim CGMY(C\sigma_t, G, M, Y)$, it's conditional skewness and kurtosis can be obtained as:

$$\begin{aligned}
\mathbb{S}kew[X_t|\mathfrak{F}_{t-1}] &= \frac{C\sigma_t(M^{Y-3}-G^{Y-3})\Gamma(3-Y)}{\left\{C\sigma_t(M^{Y-2}+G^{Y-2})\Gamma(2-Y)\right\}^{\frac{3}{2}}} \\
&= \frac{C(M^{Y-3}-G^{Y-3})\Gamma(3-Y)}{\sqrt{\sigma_t}\left\{C\sigma_t(M^{Y-2}+G^{Y-2})\Gamma(2-Y)\right\}^{\frac{3}{2}}} \qquad (5.93)\\
\mathbb{K}urt[X_t|\mathfrak{F}_{t-1}] &= 3 + \frac{C\sigma_t(M^{Y-4}-G^{Y-4})\Gamma(4-Y)}{\left\{C\sigma_t(M^{Y-2}+G^{Y-2})\Gamma(2-Y)\right\}^{2}} \\
&= 3 + \frac{C(M^{Y-4}-G^{Y-4})\Gamma(4-Y)}{\sigma_t\left\{C\sigma_t(M^{Y-2}+G^{Y-2})\Gamma(2-Y)\right\}^{2}} \qquad (5.94)
\end{aligned}$$

These conditional skewness and conditional kurtosis provide the essential tools necessary to model smile-skew patterns when log-returns are modeled with GARCH adapted to CGMY Lévy innovations.

Selecting a GARCH-CGMY Equivalent Martingale Measure:

Similar to the cases with other Lévy innovations, we are interested in finding a solution, $\hat{\theta}_t$, of the conditional Esscher equation:

$$\frac{M_{X_t|\mathfrak{F}_{t-1}}(\hat{\theta}_t + 1)}{M_{X_t|\mathfrak{F}_{t-1}}(\hat{\theta}_t)} = e^{r_t} \qquad (5.95)$$

In case of GARCH dynamics (5.89) with CGMY -Lévy innovations, conditional Esscher equation (5.95) becomes:

$$\frac{\mathbb{E}_{t-1}\left[e^{(\hat{\theta}_t+1)\left(r+\lambda\sigma_t-\frac{z_t}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}}\right)}\right]}{\mathbb{E}_{t-1}\left[e^{\hat{\theta}_t\left(r+\lambda\sigma_t-\frac{z_t}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}}\right)}\right]}=e^{r_t}$$

$$\Longrightarrow\frac{\mathbb{E}_{t-1}\left[e^{-(\hat{\theta}_t+1)\left(\frac{z_t}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}}\right)}\right]}{\mathbb{E}_{t-1}\left[e^{-\hat{\theta}_t\left(\frac{z_t}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}}\right)}\right]}=e^{-\lambda\sigma_t} \tag{5.96}$$

Introducing the constant $c=\frac{-1}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}}$, (5.96) becomes:

$$\frac{\mathbb{E}_{t-1}\left[e^{(\hat{\theta}_t+1)z_tc}\right]}{\mathbb{E}_{t-1}\left[e^{\hat{\theta}_t z_t c}\right]}=e^{-\lambda\sigma_t} \tag{5.97}$$

Using equation(5.83) in (5.97) we obtain:

$$\begin{aligned}e^{-\lambda\sigma_t} &= \frac{\exp\left\{C\Gamma(-Y)((M-(\hat{\theta}_t+1)c)^Y-M^Y+(G+(\hat{\theta}_t+1)c)^Y-G^Y)\right\}}{\exp\left\{C\Gamma(-Y)((M-\hat{\theta}_tc)^Y-M^Y+(G+\hat{\theta}_tc)^Y-G^Y)\right\}}\\ &= \exp\left\{C\Gamma(-Y)\Big(((M-\hat{\theta}_tc)-c)^Y-(M-\hat{\theta}_tc)^Y\right.\\ &\qquad\left.+((G+\hat{\theta}_tc)+c)^Y-(G+\hat{\theta}_tc)^Y\Big)\right\}\end{aligned} \tag{5.98}$$

A numerical scheme is required to obtain the solution. Given the parameters of the CGMY Lévy process, C, G, M, Y and the GARCH volatility estimate σ_t, as in (5.90), the solution ,$\hat{\theta}_t$, of (5.98) can be used to describe the distribution of log-returns as:

$$\tilde{M}_{X_s|\mathfrak{F}_{s-1}}(l)=\frac{M_{X_s|\mathfrak{F}_{s-1}}(l+\hat{\theta}_t)}{M_{X_s|\mathfrak{F}_{s-1}}(\hat{\theta}_t)} \tag{5.99}$$

With our assumption of distribution for innovations and volatility structure, equation

(5.99) becomes:

$$\begin{aligned}
& \tilde{M}_{X_s|\mathfrak{F}_{s-1}}(l) \\
= \quad & \frac{\mathbb{E}_{t-1}\left[e^{(\hat{\theta}_t+l)(r+\lambda\sigma_t+z_t c)}\right]}{\mathbb{E}_{t-1}\left[e^{\hat{\theta}_t(r+\lambda\sigma_t+z_t c)}\right]}; \qquad z_t \mid \mathfrak{F}_{t-1} \sim CGMY(C\sigma_t, G, M, Y) \\
= \quad & e^{l(r+\lambda\sigma_t)}\frac{\mathbb{E}_{t-1}\left[e^{(\hat{\theta}_t+l)cz_t}\right]}{\mathbb{E}_{t-1}\left[e^{\hat{\theta}_t cz_t}\right]} \\
\overset{(5.97)\&(5.98)}{=} \quad & e^{l(r+\lambda\sigma_t)}\exp\Big\{C\Gamma(-Y)\Big(((M-\hat{\theta}_t c)-lc)^Y-(M-\hat{\theta}_t c)^Y \\
& +((G+\hat{\theta}_t c)+lc)^Y-(G+\hat{\theta}_t c)^Y\Big)\Big\} \qquad (5.100)
\end{aligned}$$

Comparing equations (5.83) and (5.100) we recognize that under EMM innovations are again CGMY-distributed with a new characterization $M' = (M - \hat{\theta}_t c)$ and $G' = (G + \hat{\theta}_t c)$.

Similar to other GARCH-Lévy dynamics we studied, we would like to see what other parameters of the entire settings are influenced by this new characterizations. Lets start with the dynamics of the volatility under the martingale measure:

$$\begin{aligned}
\sigma'_t = \tilde{\mathbb{V}}[X_t \mid \mathfrak{F}_{t-1}] &= \tilde{\mathbb{V}}\left[r + \lambda\sigma_t - \frac{z_t}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}} \mid \mathfrak{F}_{t-1}\right] \\
& \qquad z_t \mid \mathfrak{F}_{t-1} \sim CGMY(C\sigma_t, G', M', Y) \\
&= \frac{C\sigma_t(M'^{Y-2}+G'^{Y-2})\Gamma(2-Y)}{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)} \\
&= \frac{(M'^{Y-2}+G'^{Y-2})}{(M^{Y-2}+G^{Y-2})}\sigma_t \qquad (5.101)
\end{aligned}$$

So the market and real measures are related through new characterizations $M' = (M - \hat{\theta}_t c)$, $G' = (G + \hat{\theta}_t c)$ and market and real volatility processes are related through $\sigma'_t = \frac{(M'^{Y-2}+G'^{Y-2})}{(M^{Y-2}+G^{Y-2})}\sigma_t$. We need to figure out what other parameters need to be characterized newly keeping the dynamics equivalent. Under real measure we have:

$$\begin{aligned}
X_t &= r + \lambda\sigma_t - \frac{z_t}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}} \\
\Rightarrow X_t &= r + \lambda\left[\frac{(M^{Y-2}+G^{Y-2})}{(M'^{Y-2}+G'^{Y-2})}\right]\left[\frac{(M'^{Y-2}+G'^{Y-2})}{(M^{Y-2}+G^{Y-2})}\right]\sigma_t - \frac{z_t}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}} \\
\Rightarrow X_t &= r + \lambda\left[\frac{(M^{Y-2}+G^{Y-2})}{(M'^{Y-2}+G'^{Y-2})}\right]\sigma'_t - \frac{z_t}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}} \\
\Rightarrow X_t &= r + \lambda'\sigma'_t - \frac{z_t}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}} \\
&\qquad \left[\text{introducing new characterization}\lambda' = \lambda\left[\frac{(M^{Y-2}+G^{Y-2})}{(M'^{Y-2}+G'^{Y-2})}\right]\right]
\end{aligned} \tag{5.102}$$

Furthermore we can write:

$$\begin{aligned}
C((M^{Y-2}+G^{Y-2}))\Gamma(2-Y) &= C\frac{(M^{Y-2}+G^{Y-2})}{(M'^{Y-2}+G'^{Y-2})}(M'^{Y-2}+G'^{Y-2})\Gamma(2-Y) \\
&= C'(M'^{Y-2}+G'^{Y-2})\Gamma(2-Y) \\
&\qquad \left[\text{introducing new characterization } C' = C\frac{(M^{Y-2}+G^{Y-2})}{(M'^{Y-2}+G'^{Y-2})}\right]
\end{aligned} \tag{5.103}$$

Thus finally we have the equivalent dynamics for log-returns, from (5.102), under the martingale measure:

$$X_t = r+\lambda'\sigma'_t - \frac{z_t}{\sqrt{C'(M'^{Y-2}+G'^{Y-2})\Gamma(2-Y)}} \qquad \text{with } z_t \mid \mathfrak{F}_{t-1} \sim CGMY(C'\sigma'_t, G', M', Y) \tag{5.104}$$

where parameters of the martingale dynamics, maintaining equivalence, are related with those of the market dynamics through:

$$M' = (M - \hat{\theta}_t c) \tag{5.105}$$

$$G' = (G + \hat{\theta}_t c) \tag{5.106}$$

$$C' = C\frac{(M^{Y-2}+G^{Y-2})}{(M'^{Y-2}+G'^{Y-2})} \tag{5.107}$$

$$\sigma'_t = \frac{(M'^{Y-2}+G'^{Y-2})}{(M^{Y-2}+G^{Y-2})}\sigma_t \tag{5.108}$$

As a last step we need to figure out the essential changes in GARCH parameters. We have the GARCH-CGMY dynamics:

$$\begin{aligned}
\sigma_t &= \beta_0 + \beta_1 \frac{z_{t-1}}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}} + \alpha_1 \sigma_{t-1} \\
\Rightarrow \sigma_t \left[\frac{(M'^{Y-2}+G'^{Y-2})}{(M^{Y-2}+G^{Y-2})}\right] &= \beta_0 \left[\frac{(M'^{Y-2}+G'^{Y-2})}{(M^{Y-2}+G^{Y-2})}\right] \\
&\quad + \beta_1 \left[\frac{(M'^{Y-2}+G'^{Y-2})}{(M^{Y-2}+G^{Y-2})}\right] \frac{z_{t-1}}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}} \\
&\quad + \alpha_1 \left[\frac{(M'^{Y-2}+G'^{Y-2})}{(M^{Y-2}+G^{Y-2})}\right] \sigma_{t-1} \qquad (5.109) \\
&\quad \left[\text{multiplying both sides by } \left[\frac{(M'^{Y-2}+G'^{Y-2})}{(M^{Y-2}+G^{Y-2})}\right]\right]
\end{aligned}$$

Thus considering (5.103) the equivalent GARCH volatility dynamics under the martingale measure can be written as:

$$\sigma'_t = \beta'_0 + \beta'_1 \frac{z_{t-1}}{\sqrt{C'(M'^{Y-2}+G'^{Y-2})\Gamma(2-Y)}} + \alpha_1 \sigma'_{t-1} \qquad \text{with } z_t|\mathfrak{F}_{t-1} \sim CGMY(C'\sigma'_t, G', M', Y) \qquad (5.110)$$

where:

$$\beta'_0 = \beta_0 \left[\frac{(M'^{Y-2}+G'^{Y-2})}{(M^{Y-2}+G^{Y-2})}\right] \qquad (5.111)$$

$$\beta'_1 = \beta_1 \left[\frac{(M'^{Y-2}+G'^{Y-2})}{(M^{Y-2}+G^{Y-2})}\right] \qquad (5.112)$$

GARCH-CGMY risk-neutral characterization through market price of risk:

We recall that our scale factor is $c = \frac{-1}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}}$. Thus from equation (5.89) we have:

$$X_t = r + \lambda\sigma_t + cz_t. \qquad (5.113)$$

The mgf under risk-neutral measure can be obtained from the expression of the CGMY-characteristic function:

$$\begin{aligned}
E_t^Q[e^{cz_{t+1}}] &= E_t^Q[e^{i(-ic)z_{t+1}}] \\
&= \exp\left\{C'\sigma'_{t+1}\Gamma(-Y)((M'-i(-ic))^Y - M'^Y + (G'+i(-ic))^Y - G'^Y)\right\} \\
&= \exp\left\{C'\sigma'_{t+1}\Gamma(-Y)((M'-c)^Y - M'^Y + (G'+c)^Y - G'^Y)\right\} \qquad (5.114)
\end{aligned}$$

Thus:

$$E_t^Q[e^{X_{t+1}}] = e^{(r+\lambda'\sigma'_{t+1})}\exp\left\{C'\sigma'_{t+1}\Gamma(-Y)((M'-c)^Y - M'^Y + (G'+c)^Y - G'^Y)\right\} \tag{5.115}$$

We want to choose λ', under Q, in terms of other risk-neutral parameters such that:

$$\begin{aligned} & E_t^Q[e^{X_{t+1}}] = e^{(t+1-t)r} \\ \Rightarrow \quad & e^{(r+\lambda'\sigma'_{t+1})}\exp\left\{C'\sigma'_{t+1}\Gamma(-Y)((M'-c)^Y - M'^Y + (G'+c)^Y - G'^Y)\right\} = e^r \\ \Rightarrow \quad & e^r \exp\left\{\sigma'_{t+1}\left(\lambda' - C'\Gamma(-Y)((M'-c)^Y - M'^Y + (G'+c)^Y - G'^Y)\right)\right\} = e^r \end{aligned} \tag{5.116}$$

Since $\sigma'_{t+1} \neq 0$, we must have:

$$\begin{aligned} \lambda' &= C'\Gamma(-Y)((M'-c)^Y - M'^Y + (G'+c)^Y - G'^Y) \\ &= C'\Gamma(-Y)\Big\{\left(M' + \frac{1}{\sqrt{C'(M'^{Y-2}+G'^{Y-2})\Gamma(2-Y)}}\right)^Y - M'^Y \\ &\quad + \left(G' - \frac{1}{\sqrt{C'(M'^{Y-2}+G'^{Y-2})\Gamma(2-Y)}}\right)^Y - G'^Y\Big\} \end{aligned} \tag{5.117}$$

This is the final characterization which we will be using in the expression of m.g.f which in turn will be used in our pricing(and hence in calibration).

5.2.3 GARCH with VG Lévy Innovation

Similar to other GARCH-Lévy dynamics we start with the essential tool for modeling ,the characteristic function. The characteristic function of a $VG(\sigma,\theta,\nu)$ random variable is given by:

$$\mathbb{E}[e^{isX}] = \left(1 - is\theta\nu + \frac{1}{2}\sigma^2\nu s^2\right)^{\frac{-1}{\nu}}. \tag{5.118}$$

Considering the intuitive development in section2.1.1 it follows that $VG(\sigma,\theta,\nu)$ is infinitely divisible and the associated Lévy process has the distribution of increments over $[s, t+s]$ characterized by $VG(\sigma\sqrt{t}, \theta t, \frac{\nu}{t})$. Important moments of the $VG(\sigma,\theta,\nu)$ random variable are given by:

$$\begin{aligned}
\mathbb{E}[X] &= \theta && (5.119)\\
\mathbb{V}[X] &= \sigma^2 + \nu\theta^2 && (5.120)\\
\mathbb{S}kew[X] &= \frac{\theta\nu(3\sigma^2 + 2\nu\theta^2)}{(\sigma^2 + \nu\theta^2)^{\frac{3}{2}}} && (5.121)\\
\mathbb{K}urt[X] &= 3\left[1 + 2\nu - \frac{\nu\sigma^4}{(\sigma^2 + \nu\theta^2)^2}\right] && (5.122)
\end{aligned}$$

See e.g. Schouten(2003)[102]. Similar to other GARCH-Lévy dynamics we assume the stock price follows the dynamics (5.1), where , as in GARCH settings, the log return process now follows:

$$X_t = r + \lambda\sigma_t - \frac{z_t}{\sqrt{\sigma^2 + \nu\theta^2}} \tag{5.123}$$

Here $z_t \mid \mathfrak{F}_{t-1} \sim VG(\sigma\sqrt{\sigma_t}, \theta\sigma_t, \frac{\nu}{\sigma_t})$, with the volatility processes,σ_t, following the non-linear GARCH(1,1) specification:

$$\begin{aligned}
\sigma_t &= \beta_0 + \beta_1\left[\frac{VG(\sigma\sqrt{\sigma_{t-1}}, \theta\sigma_{t-1}, \frac{\nu}{\sigma_{t-1}})}{\sqrt{(\sigma\sqrt{\sigma_{t-1}})^2 + (\frac{\nu}{\sigma_{t-1}})(\theta\sigma_{t-1})^2}} - \gamma\sqrt{\sigma_{t-1}}\right]^2 + \alpha_1\sigma_{t-1}\\
&= \beta_0 + \beta_1\left[\frac{VG(\sigma\sqrt{\sigma_{t-1}}, \theta\sigma_{t-1}, \frac{\nu}{\sigma_{t-1}}) - \mu + \mu}{\sqrt{(\sigma\sqrt{\sigma_{t-1}})^2 + (\frac{\nu}{\sigma_{t-1}})(\theta\sigma_{t-1})^2}} - \gamma\sqrt{\sigma_{t-1}}\right]^2 + \alpha_1\sigma_{t-1}\\
&= \beta_0 + \beta_1\left[stdVG + \frac{\mu}{\sqrt{(\sigma\sqrt{\sigma_{t-1}})^2 + (\frac{\nu}{\sigma_{t-1}})(\theta\sigma_{t-1})^2}} - \gamma\sqrt{\sigma_{t-1}}\right]^2 + \alpha_1\sigma_{t-1}\\
&= \beta_0 + \beta_1\left[stdVG + \left\{\frac{\theta}{\sqrt{\sigma^2 + \nu\theta^2}} - \gamma\right\}\sqrt{\sigma_{t-1}}\right]^2 + \alpha_1\sigma_{t-1} && (5.124)
\end{aligned}$$

Again the scaling ensures unit variance for innovations. With the moments of the NIG random variable, as in (5.119)-(5.122), the conditional moments of the log-returns become:

$$
\begin{aligned}
\mathbb{E}[X_t \mid \mathfrak{F}_{t-1}] &= \mathbb{E}\left[r + \lambda\sigma_t - \frac{z_t}{\sqrt{\sigma^2 + \nu\theta^2}} \mid \mathfrak{F}_{t-1}\right] \\
&= r + \lambda\sigma_t - \frac{\theta\sigma_t}{\sqrt{\sigma^2 + \nu\theta^2}} \\
&= r + \left(\lambda - \frac{\theta}{\sqrt{\sigma^2 + \nu\theta^2}}\right)\sigma_t && (5.125) \\
\mathbb{V}[X_t \mid \mathfrak{F}_{t-1}] &= \mathbb{V}\left[r + \lambda\sigma_t - \frac{z_t}{\sqrt{\sigma^2 + \nu\theta^2}} \mid \mathfrak{F}_{t-1}\right] \\
&= \frac{(\sigma\sqrt{\sigma_t})^2 + (\frac{\nu}{\sigma_t})(\theta\sigma_t)^2}{\sigma^2 + \nu\theta^2} \\
&= \sigma_t && (5.126)
\end{aligned}
$$

As in other GARCH-Lévy dynamics since X_t, as in (5.123), is just a scaled and shifted version of $z_t \mid \mathfrak{F}_{t-1} \sim VG(\sigma\sqrt{t}, \theta t, \frac{\nu}{t})$, it's conditional skewness and kurtosis can be obtained as:

$$
\begin{aligned}
\mathbb{S}kew[X_t|\mathfrak{F}_{t-1}] &= \frac{\theta\nu(3\sigma^2 + 2\nu\theta^2)}{\sqrt{\sigma_t}(\sigma^2 + \nu\theta^2)^{\frac{3}{2}}} && (5.127) \\
\mathbb{K}urt[X_t|\mathfrak{F}_{t-1}] &= 3\left[1 + 2\frac{\nu}{\sigma_t} - \frac{\nu\sigma^4}{\sigma_t(\sigma^2 + \nu\theta^2)^2}\right] && (5.128)
\end{aligned}
$$

Thus the smile-skew features can be incorporated when we consider log-return dynamics following a GARCH with VG-Lévy innovations. Existence of conditional skewness and conditional kurtosis provides such feasibility, see Christoffersen(2003)[34].

Selecting a GARCH-VG Equivalent Martingale Measure:

We follow exactly the same approach as explained in previous sections for TS, NIG and CGMY Lévy innovations. That is we are interested in finding a solution, $\hat{\theta}_t$, of the conditional Esscher equation:

$$
\frac{M_{X_t|\mathfrak{F}_{t-1}}(\hat{\theta}_t + 1)}{M_{X_t|\mathfrak{F}_{t-1}}(\hat{\theta}_t)} = e^{rt} \qquad (5.129)
$$

where $M_{X_t|\mathfrak{F}_{t-1}}(s)$ is the conditional moment generating function defined as:

$$
M_{X_t|\mathfrak{F}_{t-1}}(s) = \mathbb{E}[e^{sX_t} \mid \mathfrak{F}_{t-1}] \qquad (5.130)
$$

In case of GARCH dynamics with VG -Lévy innovations, as given by equation (5.123), conditional Esscher equation (5.129) becomes:

$$\frac{\mathbb{E}_{t-1}\left[e^{(\hat{\theta}_t+1)\left(r+\lambda\sigma_t-\frac{z_t}{\sqrt{\sigma^2+\nu\theta^2}}\right)}\right]}{\mathbb{E}_{t-1}\left[e^{\hat{\theta}_t\left(r+\lambda\sigma_t-\frac{z_t}{\sqrt{\sigma^2+\nu\theta^2}}\right)}\right]} = e^{r_t}$$

$$\Longrightarrow \frac{\mathbb{E}_{t-1}\left[e^{-(\hat{\theta}_t+1)\left(\frac{z_t}{\sqrt{\sigma^2+\nu\theta^2}}\right)}\right]}{\mathbb{E}_{t-1}\left[e^{-\hat{\theta}_t\left(\frac{z_t}{\sqrt{\sigma^2+\nu\theta^2}}\right)}\right]} = e^{-\lambda\sigma_t} \tag{5.131}$$

Introducing the constant $c = \frac{-1}{\sqrt{\sigma^2+\nu\theta^2}}$, (5.131) becomes:

$$\frac{\mathbb{E}_{t-1}\left[e^{(\hat{\theta}_t+1)z_t c}\right]}{\mathbb{E}_{t-1}\left[e^{\hat{\theta}_t z_t c}\right]} = e^{-\lambda\sigma_t} \tag{5.132}$$

Using equation(5.118) in (5.132) we obtain:

$$e^{-\lambda\sigma_t} = \frac{\left\{1-((\hat{\theta}_t+1)c)\theta\nu-\frac{1}{2}\sigma^2\nu((\hat{\theta}_t+1)c)^2\right\}^{\frac{-1}{\nu/\sigma_t}}}{\left\{1-(\hat{\theta}_t c)\theta\nu-\frac{1}{2}\sigma^2\nu(\hat{\theta}_t c)^2\right\}^{\frac{-1}{\nu/\sigma_t}}} \tag{5.133}$$

Similar to the case of GARCH-Lévy dynamics with other Lévy innovations we need a numerical solution. Given the parameters of the VG Lévy process, σ, θ, ν and the GARCH volatility estimate σ_t, as in (5.124), the solution ,$\hat{\theta}_t$, of (5.133) can be used to describe the distribution of log-returns as:

$$\tilde{M}_{X_s|\mathfrak{F}_{s-1}}(l) = \frac{M_{X_s|\mathfrak{F}_{s-1}}(l+\hat{\theta}_t)}{M_{X_s|\mathfrak{F}_{s-1}}(\hat{\theta}_t)} \tag{5.134}$$

With our assumption of distribution for innovations and volatility structure, equation

(5.134) becomes:

$$\begin{aligned}
& \tilde{M}_{X_s|\mathfrak{F}_{s-1}}(l) \\
= \quad & \frac{\mathbb{E}_{t-1}\left[e^{(\hat{\theta}_t+l)(r+\lambda\sigma_t+z_t c)}\right]}{\mathbb{E}_{t-1}\left[e^{\hat{\theta}_t(r+\lambda\sigma_t+z_t c)}\right]}; \qquad z_t \mid \mathfrak{F}_{t-1} \sim VG(\sigma\sqrt{\sigma_t}, \theta\sigma_t, \frac{\nu}{\sigma_t}) \\
= \quad & e^{l(r+\lambda\sigma_t)}\frac{\mathbb{E}_{t-1}\left[e^{(\hat{\theta}_t+l)cz_t}\right]}{\mathbb{E}_{t-1}\left[e^{\hat{\theta}_t cz_t}\right]} \\
\overset{(5.132)\&(5.133)}{=} \quad & e^{l(r+\lambda\sigma_t)}\frac{\left\{1-((\hat{\theta}_t+l)c)\theta\nu-\frac{1}{2}\sigma^2\nu((\hat{\theta}_t+l)c)^2\right\}^{\frac{-1}{\nu/\sigma_t}}}{\left\{1-(\hat{\theta}_t c)\theta\nu-\frac{1}{2}\sigma^2\nu(\hat{\theta}_t c)^2\right\}^{\frac{-1}{\nu/\sigma_t}}} \\
= \quad & e^{l(r+\lambda\sigma_t)}\left[\frac{\left\{1-(\hat{\theta}_t c)\theta\nu-\frac{1}{2}\sigma^2\nu(\hat{\theta}_t c)^2\right\}-l\left\{c\theta+\sigma^2\hat{\theta}_t c^2\right\}\nu-\frac{1}{2}\sigma^2\nu l^2c^2}{1-(\hat{\theta}_t c)\theta\nu-\frac{1}{2}\sigma^2\nu(\hat{\theta}_t c)^2}\right]^{\frac{-1}{\nu/\sigma_t}} \\
= \quad & e^{l(r+\lambda\sigma_t)}\left(1-l\theta'\nu-\frac{1}{2}\sigma'^2\nu l^2\right)^{\frac{-1}{\nu/\sigma_t}} \qquad (5.135)
\end{aligned}$$

where

$$\theta' = \frac{c\theta+\sigma^2\hat{\theta}_t c^2}{k} \qquad (5.136)$$

$$\sigma' = \frac{\sigma c}{\sqrt{k}} \qquad (5.137)$$

$$k = 1-(\hat{\theta}_t c)\theta\nu-\frac{1}{2}\sigma^2\nu(\hat{\theta}_t c)^2$$

Comparing equations (5.118) and (5.135) we recognize that under EMM innovations are again VG-distributed with new parameterizations, namely $VG(\sigma'\sqrt{\sigma_t}, \theta'\sigma_t, \frac{\nu}{\sigma_t})$.

The new parameterizations influence the dynamics of the volatility under the martingale measure:

$$\begin{aligned}
\sigma'_t = \tilde{\mathbb{V}}[X_t \mid \mathfrak{F}_{t-1}] &= \tilde{\mathbb{V}}\left[r+\lambda\sigma_t-\frac{z_t}{\sqrt{\sigma^2+\nu\theta^2}} \mid \mathfrak{F}_{t-1}\right] \\
&= \frac{\sigma^2+\nu\theta^2}{\sigma^2+\nu\theta^2} \\
&= \left[\frac{\sigma'^2+\nu\theta'^2}{\sigma^2+\nu\theta^2}\right]\sigma_t \qquad (5.138)
\end{aligned}$$

So the market and real volatility processes are related through $\sigma'_t = \left[\frac{\sigma'^2+\nu\theta'^2}{\sigma^2+\nu\theta^2}\right]\sigma_t$.

GARCH-VG risk-neutral characterization through market price of risk:

With the scale factor $u = \frac{-1}{\sqrt{\sigma^2+\nu\theta^2}}$ we can write equation (5.123) as:

$$X_t = r + \lambda\sigma_t + uz_t. \tag{5.139}$$

The mgf under risk-neutral measure can be obtained from the expression of the VG-characteristic function:

$$\begin{aligned} E_t^Q[e^{uz_{t+1}}] &= E_t^Q[e^{i(-iu)z_{t+1}}] \\ &= \left(1 - u\theta\nu - \frac{1}{2}\sigma^2\nu u^2\right)^{\frac{-1}{\nu/\sigma_{t+1}}} \end{aligned} \tag{5.140}$$

Thus:

$$\begin{aligned} E_t^Q[e^{X_{t+1}}] &= e^{(r+\lambda\sigma_{t+1})}\exp\left\{\frac{-1}{\nu/\sigma_{t+1}} log\left(1 - u\theta\nu - \frac{1}{2}\sigma^2\nu u^2\right)\right\} \\ &= e^{(r+\lambda\sigma_{t+1})}\exp\left\{\frac{-\sigma_{t+1}}{\nu} log\left(1 - u\theta\nu - \frac{1}{2}\sigma^2\nu u^2\right)\right\} \end{aligned} \tag{5.141}$$

We want to choose λ, under Q, in terms of other risk-neutral parameters such that:

$$\begin{aligned} & E_t^Q[e^{X_{t+1}}] = e^{(t+1-t)r} \\ \Rightarrow \quad & e^{(r+\lambda\sigma_{t+1})}\exp\left\{\frac{-\sigma_{t+1}}{\nu} log\left(1 + \frac{1}{\sqrt{\sigma^2+\nu\theta^2}}\theta\nu - \frac{1}{2}\frac{\sigma^2\nu}{\sigma^2+\nu\theta^2}\right)\right\} = e^r \end{aligned} \tag{5.142}$$

Since $\sigma_{t+1} \neq 0$, we must have:

$$\lambda = \frac{1}{\nu} log\left(1 + \frac{1}{\sqrt{\sigma^2+\nu\theta^2}}\theta\nu - \frac{1}{2}\frac{\sigma^2\nu}{\sigma^2+\nu\theta^2}\right) \tag{5.143}$$

This is the final characterization which we will be using in the expression of mgf which in turn will be used in our pricing(and hence in calibration).

5.3 Closed form GARCH Option Pricing with Different Lévy Innovations

At this stage we have the dynamics of log-returns and GARCH volatility with different Lévy innovations under risk neutral martingale measure. Now we need to consider fourier inversion techniques, derived in details in chapter2, for pricing European options, see e.g. Heston(1993)[69]. The tool which is required for this is the characteristic function of the model. In previous sections we present the mathematical underpinnings to derive the fact that for GARCH-Lévy dynamics one period ahead conditional distribution of the underlying asset again follow the GARCH dynamics with innovation coming from same Lévy process. This fact leads to the characterization of the risk neutral martingale measure. But to price option the conditional distribution of underlying asset at multi period ahead maturity, say at 'T', is required. In this section we discuss and derive such a distribution following the recursive method developed by Heston and Nandi(2000)[70]. The idea of their recursive procedure lies in the fact that the general conditional moment generating function can be expressed as:

$$\mathbb{E}[e^{u\log(S_T)} \mid \mathfrak{F}_t] = S_t^u \exp\left[A(t,T,u) + B(t,T,u)\sigma_{t+1}\right] \tag{5.144}$$

The goal is to solve for $A(t,T,u)$ and $B(t,T,u)$ for different Lévy innovations characterizing different σ_t. In this section we will provide the solutions for four Lévy innovations, one from a much cited Lévy process namely tempered stable Lévy innovations(TS) and others from time changed Lévy process namely Normal inverse Gaussian(NIG) ,Variance Gamma(VG) as well as CGMY processes.

Following equation (5.144) we can write:

$$\mathbb{E}[e^{u\log(S_T)} \mid \mathfrak{F}_{t+1}] = S_{t+1}^u \exp\left[A(t+1,T,u) + B(t+1,T,u)\sigma_{t+2}\right] \tag{5.145}$$

That is we assume that the general form of conditional mgf holds for time $t+1$, and now using iterative property of conditional expectation(which is the central feature Heston Nandi(2000)[70] recursive approach) we can write the expression for the conditional mgf at time t :

$$\begin{aligned} \mathbb{E}[e^{u\log(S_T)} \mid \mathfrak{F}_t] &= \mathbb{E}\left[\mathbb{E}[e^{u\log(S_T)} \mid \mathfrak{F}_{t+1}] \mid \mathfrak{F}_t\right] \\ &\overset{(5.145)}{=} \mathbb{E}\left[S_{t+1}^u \exp\left[A(t+1,T,u) + B(t+1,T,u)\sigma_{t+2}\right] \mid \mathfrak{F}_t\right] \end{aligned} \tag{5.146}$$

Following equation (5.1) we have $S_{t+1}^u = S_t^u e^{uX_{t+1}}$. Plugging this in equation (5.146) we obtain:

$$\begin{aligned}\mathbb{E}[e^{u\log(S_T)} \mid \mathfrak{F}_t] &= \mathbb{E}\left[S_t^u e^{uX_{t+1}} \exp\left[A(t+1,T,u) + B(t+1,T,u)\sigma_{t+2}\right] \mid \mathfrak{F}_t\right] \\ &= S_t^u \mathbb{E}\left[e^{uX_{t+1}} \exp\left[A(t+1,T,u) + B(t+1,T,u)\sigma_{t+2}\right] \mid \mathfrak{F}_t\right]\end{aligned} \tag{5.147}$$

This development is the central tool in the discrete time GARCH modeling. It was first put forward in Heston and Nandi(2000)[70] and subsequently used by Christoffersen et al(2006)[32]; which shows that this conditional MGF characterization can have distributional assumption other than normal one. This simply tells us that given the information available up until today the evolution of returns can be characterized , in discrete fashion, for any number of future period, through two well defined recursive relations of "A" and "B". Thus no matter how many steps are between t and T, one could use equations(5.145) and (5.146) recursively to derive the conditional mgf at any maturity T given the information available up to t. Comparing (5.144) and (5.147) we can set up the recursive relations for the co-efficients $A(t,T,u)$ and $B(t,T,u)$. Corresponding to the choices of dynamics for X_t and σ_t for various Lévy innovations we will have different expressions for such recursive relations. In the following sections we will see four such relations, one for TS-Lévy innovation and others are for NIG, VG and CGMY Lévy innovations[3]

[3]One point worth mentioning here is that Heston and Nandi(2000)[70] didn't consider one day ahead volatility, σ_{t+1}, as parameter of the model. They estimated this volatility exogenously and supplied it as constant while running the calibration. However they updated this constant at different time points at which their considered options were recorded. Considering one period ahead volatility as parameter is consistent when calibration considers only few days record; as each day adds one more parameter to be estimated.

5.3.1 The case of TS Lévy innovations

We replace X_{t+1} and σ_{t+2} in equation (5.147), from equations (5.36) and (5.9), respectively, which are developed for the GARCH- TS dynamics[4]:

$$
\begin{aligned}
& \mathbb{E}[e^{u\log(S_T)} \mid \mathfrak{F}_t] \\
= \;& S_t^u \mathbb{E}\Big[e^{u\left(r+\lambda\sigma_{t+1}-\frac{z_{t+1}}{2\sqrt{\alpha(1-\alpha)\gamma\eta^{\frac{\alpha-2}{\alpha}}}}\right)} \exp\big[A(t+1,T,u) \\
& +B(t+1,T,u)\Big(\beta_0+\beta_1\frac{z_{t+1}}{2\sqrt{\alpha(1-\alpha)\gamma\eta^{\frac{\alpha-2}{\alpha}}}}+\alpha_1\sigma_{t+1}\Big)\big] \mid \mathfrak{F}_t\Big] \\
= \;& S_t^u \mathbb{E}\Big[\exp\Big\{ur+\lambda u\sigma_{t+1}+A(t+1,T,u)+\beta_0 B(t+1,T,u) \\
& +\alpha_1\sigma_{t+1}B(t+1,T,u)+\frac{[\beta_1 B(t+1,T,u)-u]\,z_{t+1}}{2\sqrt{\alpha(1-\alpha)\gamma\eta^{\frac{\alpha-2}{\alpha}}}}\Big\} \mid \mathfrak{F}_t\Big] \\
\overset{(5.4)}{=} \;& S_t^u \exp\Big\{ur+\lambda u\sigma_{t+1}+A(t+1,T,u)+\beta_0 B(t+1,T,u) \\
& +\alpha_1\sigma_{t+1}B(t+1,T,u)+\gamma\sigma_{t+1}\eta\left[1-\left(1-\frac{2.\frac{\beta_1 B(t+1,T,u)-u}{2\sqrt{\alpha(1-\alpha)\gamma\eta^{\frac{\alpha-2}{\alpha}}}}}{\eta^{\frac{1}{\alpha}}}\right)^{\alpha}\right]\Big\} \\
& \qquad\qquad \text{since } \; z_{t+1}|\mathfrak{F}_t \sim TS_\alpha(\gamma\sigma_{t+1},\eta) \\
= \;& S_t^u \exp\Big\{ur+\lambda u\sigma_{t+1}+A(t+1,T,u)+\beta_0 B(t+1,T,u) \\
& +\alpha_1\sigma_{t+1}B(t+1,T,u)+\gamma\sigma_{t+1}\eta\left[1-\left(1-\frac{\beta_1 B(t+1,T,u)-u}{\sqrt{\alpha(1-\alpha)\gamma\eta}}\right)^{\alpha}\right]\Big\} \\
= \;& S_t^u \exp\Big\{ur+A(t+1,T,u)+\beta_0 B(t+1,T,u) \\
& +\left\{\lambda u+\alpha_1 B(t+1,T,u)+\gamma\eta\left[1-\left(1-\frac{\beta_1 B(t+1,T,u)-u}{\sqrt{\alpha(1-\alpha)\gamma\eta}}\right)^{\alpha}\right]\right\}\sigma_{t+1}\Big\}
\end{aligned}
\tag{5.148}
$$

[4]However one subtle issue is worth noting with TS innovations. This characterization is not suitable for risk-management because TS distribution is defined on positive half of the real line. In risk management we are particularly interested about downside of the distribution to estimate the risk measures e.g. VaR. So negative innovations are of particular interest. Off-course this can be circumvented by considering negative transformation of the TS distribution but that essentially complicates the model. Furthermore though VaR can be obtained by considering negative transformation of TS distribution, coherent risk measure like Spectral Risk Measure(SRM) can not be estimated for one sided distribution like TS. SRM considers entire spectrum of returns(both profit and losses). Thus naturally for risk management innovation from more flexible Lévy process defined on the entire real line will be desired, e.g. NIG, CGMY etc.

Comparing equation (5.148) with equation (5.144) we obtain the following recursive relations:

$$\begin{aligned} A(t,T,u) &= ur + A(t+1,T,u) + \beta_0 B(t+1,T,u) \\ B(t,T,u) &= \lambda u + \alpha_1 B(t+1,T,u) + \gamma\eta \left[1 - \left(1 - \frac{\beta_1 B(t+1,T,u) - u}{\sqrt{\alpha(1-\alpha)\gamma\eta}} \right)^{\alpha} \right] \end{aligned} \tag{5.149}$$

5.3.2 The case of NIG time changed Lévy innovations

In this case we replace X_{t+1} and σ_{t+2} in equation (5.147), from equations (5.53) and (5.54) respectively, which are developed for GARCH-NIG Lévy dynamics:

$$\begin{aligned} & \mathbb{E}[e^{u\log(S_T)} \mid \mathfrak{F}_t] \\ = \;& S_t^u \mathbb{E}\Big[e^{u\left(r + \lambda\sigma_{t+1} - \frac{z_{t+1}}{\sqrt{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}}}\right)} \exp\big[A(t+1,T,u) \qquad (5.150) \\ & + B(t+1,T,u)\Big(\beta_0 + \beta_1 \frac{z_{t+1}}{\sqrt{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}}} + \alpha_1\sigma_{t+1}\Big)\big] \mid \mathfrak{F}_t\Big] \\ = \;& S_t^u \mathbb{E}\Big[\exp\Big\{ ur + \lambda u\sigma_{t+1} + A(t+1,T,u) + \beta_0 B(t+1,T,u) \\ & + \alpha_1\sigma_{t+1} B(t+1,T,u) + \frac{[\beta_1 B(t+1,T,u) - u]\, z_{t+1}}{\sqrt{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}}} \Big\} \mid \mathfrak{F}_t\Big] \\ \overset{(5.47)}{=} \;& S_t^u \exp\Big\{ ur + \lambda u\sigma_{t+1} + A(t+1,T,u) + \beta_0 B(t+1,T,u) \\ & + \alpha_1\sigma_{t+1} B(t+1,T,u) \\ & - \delta\sigma_{t+1}\left[\sqrt{\alpha^2 - \left(\beta + \frac{[\beta_1 B(t+1,T,u) - u]}{\sqrt{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}}}\right)^2} - \sqrt{\alpha^2-\beta^2} \right] \Big\} \\ & \text{since } \; z_{t+1}|\mathfrak{F}_t \sim NIG(\alpha,\beta,\delta\sigma_{t+1}) \\ = \;& S_t^u \exp\Big\{ ur + A(t+1,T,u) + \beta_0 B(t+1,T,u) \\ & + \Big\{ \lambda u + \alpha_1 B(t+1,T,u) \\ & - \delta\left[\sqrt{\alpha^2 - \left(\beta + \frac{[\beta_1 B(t+1,T,u) - u]}{\sqrt{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}}}\right)^2} - \sqrt{\alpha^2-\beta^2} \right] \Big\}\sigma_{t+1} \Big\} \end{aligned} \tag{5.151}$$

Comparing equation (5.150) with equation (5.144) we obtain the following recursive relations:

$$\begin{aligned}
A(t,T,u) &= ur + A(t+1,T,u) + \beta_0 B(t+1,T,u) \\
B(t,T,u) &= \lambda u + \alpha_1 B(t+1,T,u) \\
&\quad -\delta\left[\sqrt{\alpha^2 - \left(\beta + \frac{[\beta_1 B(t+1,T,u) - u]}{\sqrt{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}}}\right)^2} - \sqrt{\alpha^2-\beta^2}\right]
\end{aligned} \tag{5.152}$$

Unlike tempered stable(TS) Lévy innovation, the dynamics with NIG Lévy innovation has serious problem associated with it. Namely when TS Lévy process is a subordinator, NIG process is not. Consequently the support of NIG distribution is the entire real line that is why NIG process can exhibit both positive as well as negative jumps. Thus from our calibration we realize that no matter how we restrict the parameters, the volatility often becomes negative, as $z_t \sim NIG(\alpha, \beta, \delta\sigma_t)$ often assumes both positive as well as negative values.

Hence for Lévy innovations coming from Lévy processes exhibiting both sided jumps we need to consider non-linear dynamics of Heston-Nandi type to ensure positivity. Let us assume the non-linear dynamics under risk neutral parameters as:

$$\begin{aligned}
\sigma_t &= \beta_0 + \beta_1\left[\frac{NIG(\alpha,\beta,\delta\sigma_{t-1})}{\sqrt{\alpha^2\delta\sigma_t(\alpha^2-\beta^2)^{\frac{-3}{2}}}} - \gamma\sqrt{\sigma_{t-1}}\right]^2 + \alpha_1\sigma_{t-1} \\
&= \beta_0 + \beta_1\left[\frac{NIG(\alpha,\beta,\delta\sigma_{t-1}) - \mu + \mu}{\sqrt{\alpha^2\delta\sigma_{t-1}(\alpha^2-\beta^2)^{\frac{-3}{2}}}} - \gamma\sqrt{\sigma_{t-1}}\right]^2 + \alpha_1\sigma_{t-1} \\
&= \beta_0 + \beta_1\left[stdNIG + \frac{\mu}{\sqrt{\alpha^2\delta\sigma_{t-1}(\alpha^2-\beta^2)^{\frac{-3}{2}}}} - \gamma\sqrt{\sigma_{t-1}}\right]^2 + \alpha_1\sigma_{t-1}
\end{aligned} \tag{5.153}$$

Here μ is the expected value of a $NIG(\alpha, \beta, \delta\sigma_{t-1})$ random variable and is given by equation (5.48). Hence we can further simplify the above equation to:

$$\sigma_t = \beta_0 + \beta_1\left[stdNIG + \left\{\frac{\beta\sqrt{\delta}}{\alpha(\alpha^2-\beta^2)^{\frac{-1}{4}}} - \gamma\right\}\sqrt{\sigma_{t-1}}\right]^2 + \alpha_1\sigma_{t-1} \tag{5.154}$$

The problem with this characterization of volatility is that when σ_t , as in equation (5.154), is plugged into equation (5.147), it does not yield explicit recursive relations for $A(t, T, u)$ and $B(t, T, u)$. Hence no closed form valuation of European option is possible without further approximation. Recently similar problem is encountered by Chayawat Ornthanalai [5](2010)[89] where component affine transformations are considered to incorporate Lévy innovations to GARCH volatility dynamics. The conclusion in Chayawat Ornthanalai(2010)[89] is that no matter how sophisticated affine relation is considered, for truly potential Lévy innovations–capable of exhibiting both positive and negative jumps–there is no alternative to Monte-Carlo valuation of derivatives. Though compare to the affine transformation of Chayawat Ornthanalai(2010)[89] our volatility dynamics, described in equation (5.154), is apparently simpler, nonetheless like Chayawat Ornthanalai(2010)[89] we found that the simulated prices of European options are much appreciable compare to other available models. But the problem is that it requires long time to price even a single option. To appreciate such pricing we need to consider large number of simulations[6] by the expense of huge computational time and that renders quick calibration practically infeasible.

We apply approximation to the dynamics (5.154) to uphold the closed form valuation techniques similar to those of Heston and Nandi. When the dynamics (5.154) is characterized for NIG innovations we propose an approximation which preserves the characterization of dynamics but replaces the standard NIG innovations by standard Normal innovations[7]:

$$\sigma_t \simeq \beta_0 + \beta_1 \left[stdNormal + \left\{ \frac{\beta\sqrt{\delta}}{\alpha(\alpha^2 - \beta^2)^{\frac{-1}{4}}} - \gamma \right\} \sqrt{\sigma_{t-1}} \right]^2 + \alpha_1 \sigma_{t-1} \qquad (5.155)$$

The idea behind such approximation is driven by the fact that Heston and Nandi's

[5]Chayawat Ornthanalai is developing GARCH-Lévy dynamics for asset pricing. In his work he considered Monte Carlo simulation for pricing European options when Lévy innovations have entire real line as its support, thus the dynamics can demonstrate negative as well as positive jumps.

[6]Appreciation increases with the increase in number of paths. Less than 5000 paths are not usually that much appreciable.

[7]We understand that it is a rough approximation. But we realize that the benefit achieved is enormous, as this offers a characterization yielding to analytic valuation. Also by plotting densities of Normal and Normal Inverse Gaussian(NIG) random numbers with same mean and variance we realize that stochastically we loose very little.

closed form pricing was possible solely because of the following relation involving a standard Normal random variable:

$$\mathbb{E}[\exp\{a(z+b)^2\}] = \exp\left\{-\frac{1}{2}ln(1-2a) + \frac{ab^2}{1-2a}\right\} \tag{5.156}$$

where $z \sim N(0,1)$.

We now apply the volatility dynamics (5.155) (which is characterized for NIG innovation and approximated through the replacement of standard NIG variate by standard Normal variate) in general recursive relation (5.147) where scaled NIG innovations drive the returns:

$$\begin{aligned}
&\mathbb{E}[e^{u\log(S_T)} \mid \mathfrak{F}_t^1] \\
= \;& S_t^u \mathbb{E}\Big[e^{u\left(r+\lambda\sigma_{t+1}-\frac{z_{t+1}}{\sqrt{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}}}\right)} \exp\big[A(t+1,T,u) \\
&+B(t+1,T,u)\Big(\beta_0+\beta_1\left[z+\left\{\frac{\beta\sqrt{\delta}}{\alpha(\alpha^2-\beta^2)^{\frac{-1}{4}}}-\gamma\right\}\sqrt{\sigma_{t+1}}\right]^2+\alpha_1\sigma_{t+1}\Big)\big] \mid \mathfrak{F}_t^1\Big]
\end{aligned} \tag{5.157}$$

We apply the relation (5.156) to (5.157) and simplify it further:

$$\begin{aligned}
&\mathbb{E}[e^{u\log(S_T)} \mid \mathfrak{F}_t^1] \\
\overset{(5.47)}{=} \;& S_t^u \exp\left\{u\Big(r+\lambda\sigma_{t+1}\Big) - \delta\sigma_{t+1}\left[\sqrt{\alpha^2-\Big(\beta+\frac{(-u)}{\sqrt{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}}}\Big)^2}-\sqrt{\alpha^2-\beta^2}\right]\right\} \\
&\exp\Big\{A(t+1,T,u)+B(t+1,T,u)\beta_0-\frac{1}{2}log\Big[1-2B(t+1,T,u)\beta_1\Big] \\
&+B(t+1,T,u)\beta_1\left(\frac{\beta\sqrt{\delta}}{\alpha(\alpha^2-\beta^2)^{\frac{-1}{4}}}-\gamma\right)^2\sigma_{t+1}\frac{1}{[1-2B(t+1,T,u)\beta_1]} \\
&\qquad +\alpha_1 B(t+1,T,u)\sigma_{t+1}\Big\} \\
&\qquad\qquad \text{since } z_{t+1}|\mathfrak{F}_t^1 \sim NIG(\alpha,\beta,\delta\sigma_{t+1}) \\
= \;& S_t^u\exp\Big\{ur+A(t+1,T,u)+\beta_0 B(t+1,T,u)-\frac{1}{2}log\Big[1-2B(t+1,T,u)\beta_1\Big] \\
&+\Big[\lambda u+\alpha_1 B(t+1,T,u) \\
&-\delta\left[\sqrt{\alpha^2-\Big(\beta+\frac{(-u)}{\sqrt{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}}}\Big)^2}-\sqrt{\alpha^2-\beta^2}\right] \\
&+B(t+1,T,u)\beta_1\Big(\frac{\beta\sqrt{\delta}}{\alpha(\alpha^2-\beta^2)^{\frac{-1}{4}}}-\gamma\Big)^2\frac{1}{[1-2B(t+1,T,u)\beta_1]}\Big]\sigma_{t+1}\Big\}
\end{aligned} \tag{5.158}$$

Comparing equation (5.158) with equation (5.144) we obtain the following recursive relations:

$$\begin{aligned} A(t,T,u) &= ur + A(t+1,T,u) + \beta_0 B(t+1,T,u) - \frac{1}{2}log\Big[1 - 2B(t+1,T,u)\beta_1\Big] \\ B(t,T,u) &= \lambda u + \alpha_1 B(t+1,T,u) - \delta\left[\sqrt{\alpha^2 - \left(\beta + \frac{(-u)}{\sqrt{\alpha^2\delta(\alpha^2-\beta^2)^{\frac{-3}{2}}}}\right)^2} - \sqrt{\alpha^2-\beta^2}\right] \\ &\quad + B(t+1,T,u)\beta_1\left(\frac{\beta\sqrt{\delta}}{\alpha(\alpha^2-\beta^2)^{\frac{-1}{4}}} - \gamma\right)^2 \frac{1}{[1-2B(t+1,T,u)\beta_1]} \end{aligned} \tag{5.159}$$

We will obtain the option prices through Fourier Inversion as in Heston(1993)[69] and Heston and Nandi(2000)[70]. For closed form (up to numerical integration) GARCH model with NIG innovations let us denote the model price by c_{cfgnig}. This model has seven parameters to be estimated: [8] $[\beta_0, \beta_1, \alpha_1, \gamma, \alpha, \beta, \delta]$. c_{cfgnig} is obtained as in (4.12) but replacing $A(t,T,u)$ and $B(t,T,u)$ recursive relations in (4.11) by those in (5.159).

Given the constraints on several parameters, from our empirical observation we realize that it is more effective to consider the calibration as a constrained optimization problem rather than a simple non-linear least square one. The constraints we need to consider arise from basic GARCH structure as well as usual NIG parametrization. Namely we need to ensure that $\beta_0 \geq 0, \beta_1 \geq 0, \alpha_1 \geq 0, \alpha_1 + \beta_1 < 1$; $\alpha > 0$ and $\mid \beta \mid \leq \alpha$ i.e. $-\alpha - \beta < 0$, $\delta > 0$.

[8] One day ahead GARCH variance σ^2_{t+1}, as is required in (4.10), can also be treated as parameter. Considerable variation in σ^2_{t+1} is usually observed when the calibration is carried out in day-by-day dynamic fashion. However treating one-day-ahead volatility as parameter the calibration may deem manageable only for market data of few days. This is because each new day increases a new parameter to be estimated. So for calibrations using option records over a long period we need to directly feed the one-day ahead volatilities in dynamic fashion. Heston and Nandi(2000)[70] feeded these values in calibration by estimating them through GARCH process. This force the calibration to heavily rely on long time series of asset returns, in addition to the market price of options. We implement the same approach for GARCH-NIG volatility dynamics. But we realize that though with GARCH-NIG volatility, our calibration often provides better fit than Heston Nandi's model, nonetheless if we proxy daily GARCH volatility by the average daily implied volatilities, the calibration systematically outperforms Heston Nandi's model. Thus we report the calibration results with the average implied volatilities. Another positive side with average implied volatilities, perhaps a very important one, is that it renders past asset returns redundant.

Thus to calibrate the model we consider the following optimization problem [9]:

$$\begin{aligned} &Minimize \left[\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(C^i_{market} - c^i_{cfgnig}[\beta_0,\beta_1,\alpha_1,\gamma,\alpha,\beta,\delta]\right)^2} \right] \\ &s.t. \qquad A.[\beta_0,\beta_1,\alpha_1,\gamma,\alpha,\beta,\delta]' \leq b. \end{aligned} \tag{5.160}$$

Here

$$A = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}$$

and

$$b = [0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0]'.$$

[9] We use Matlab function "fmincon" to implement such constrained optimization.

5.3.3 The case of CGMY time changed Lévy Innovations:

This time we replace X_{t+1} and σ_{t+2} in equation (5.147) from equations (5.89) and (5.90) respectively, for the GARCH with CGMY Lévy dynamics :

$$
\begin{aligned}
&\mathbb{E}[e^{u\log(S_T)} \mid \mathfrak{F}_t] \\
= \quad & S_t^u \mathbb{E}\Big[e^{u\left(r+\lambda\sigma_{t+1}-\frac{z_{t+1}}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}}\right)} \exp\big[A(t+1,T,u) \\
& +B(t+1,T,u)\Big(\beta_0+\beta_1\frac{z_{t+1}}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}}+\alpha_1\sigma_{t+1}\Big)\big] \mid \mathfrak{F}_t\Big] \\
= \quad & S_t^u \mathbb{E}\Big[\exp\Big\{ur+\lambda u\sigma_{t+1}+A(t+1,T,u)+\beta_0 B(t+1,T,u) \\
& +\alpha_1\sigma_{t+1}B(t+1,T,u)+\frac{[\beta_1 B(t+1,T,u)-u]\,z_{t+1}}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}}\Big\} \mid \mathfrak{F}_t\Big] \\
\overset{(5.83)}{=} \quad & S_t^u \exp\Big\{ur+\lambda u\sigma_{t+1}+A(t+1,T,u)+\beta_0 B(t+1,T,u) \\
& \qquad +\alpha_1\sigma_{t+1}B(t+1,T,u) \\
& \qquad +C\sigma_{t+1}\Gamma(-Y)\Big[\Big(\Big\{M-\frac{[\beta_1 B(t+1,T,u)-u]}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}}\Big\}^Y-M^Y \\
& \qquad +\Big\{G+\frac{[\beta_1 B(t+1,T,u)-u]}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}}\Big\}^Y-G^Y\Big)\Big]\Big\} \\
& \qquad\qquad\qquad \text{since } z_{t+1}|\mathfrak{F}_t \sim CGMY(C\sigma_{t+1},G,M,Y) \\
= \quad & S_t^u \exp\Big\{ur+A(t+1,T,u)+\beta_0 B(t+1,T,u) \\
& +\Big\{\lambda u+\alpha_1 B(t+1,T,u) \\
& \quad +C\Gamma(-Y)\Big[\Big(\Big\{M-\frac{[\beta_1 B(t+1,T,u)-u]}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}}\Big\}^Y-M^Y \\
& \quad +\Big\{G+\frac{[\beta_1 B(t+1,T,u)-u]}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}}\Big\}^Y-G^Y\Big)\Big]\Big\}\sigma_{t+1}\Big\}
\end{aligned}
\tag{5.161}
$$

Comparing equation (5.161) with equation (5.144) we obtain the following recursive relations:

$$\begin{aligned}
A(t,T,u) &= ur + A(t+1,T,u) + \beta_0 B(t+1,T,u) \\
B(t,T,u) &= \lambda u + \alpha_1 B(t+1,T,u) \\
&\quad + C\Gamma(-Y)\Big[\Big(\left\{M - \frac{[\beta_1 B(t+1,T,u) - u]}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}}\right\}^Y - M^Y \\
&\quad + \left\{G + \frac{[\beta_1 B(t+1,T,u) - u]}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}}\right\}^Y - G^Y\Big)\Big]
\end{aligned} \tag{5.162}$$

Similar to the case with NIG Lévy innovations this dynamics with CGMY Lévy innovations has serious problem associated with it. Since CGMY process is not a subordinator, the volatility often becomes negative, as $z_t \sim CGMY(C\sigma_{t-1}, G, M, Y)$ often assumes both positive as well as negative values.

Hence like NIG case we need to consider non-linear dynamics of Heston-Nandi type to ensure positivity. Let us assume the non-linear dynamics under risk neutral parameters as:

$$\begin{aligned}
\sigma_t &= \beta_0 + \beta_1 \left[\frac{CGMY(C\sigma_{t-1}, G, M, Y)}{\sqrt{C\sigma_{t-1}(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}} - \gamma\sqrt{\sigma_{t-1}}\right]^2 + \alpha_1\sigma_{t-1} \\
&= \beta_0 + \beta_1 \left[\frac{CGMY(C\sigma_{t-1}, G, M, Y) - \mu + \mu}{\sqrt{C\sigma_{t-1}(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}} - \gamma\sqrt{\sigma_{t-1}}\right]^2 + \alpha_1\sigma_{t-1} \\
&= \beta_0 + \beta_1 \left[stdCGMY + \frac{\mu}{\sqrt{C\sigma_{t-1}(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}} - \gamma\sqrt{\sigma_{t-1}}\right]^2 + \alpha_1\sigma_{t-1}
\end{aligned} \tag{5.163}$$

Here μ is the expected value of a $CGMY(C\sigma_{t-1}, G, M, Y)$ random variable and is given by equation (5.85). Hence we can further simplify the above equation to:

$$\sigma_t = \beta_0 + \beta_1 \left[stdCGMY + \left\{\frac{C(M^{Y-1}-G^{Y-1})\Gamma(1-Y)}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}} - \gamma\right\}\sqrt{\sigma_{t-1}}\right]^2 + \alpha_1\sigma_{t-1} \tag{5.164}$$

The problem with this characterization of volatility is that when σ_t , as into equation (5.164), is plugged in equation (5.147) it doesn't yield explicit recursive relations for

$A(t,T,u)$ and $B(t,T,u)$. Hence no closed form valuation of European option is possible without further approximation. Again we refer the findings in Chayawat Ornthanalai(2010)[89] which in essence implies that no matter how sophisticated affine relation is considered, for truly potential Lévy innovations– capable of exhibiting both positive and negative jumps– there is no alternative to Monte-Carlo valuation of options[10]. Our aim is to get rid off simulation to make the model practically implementable.

We apply approximation to the dynamics (5.164) to uphold the analytic valuation techniques similar to those of Heston and Nandi. When the dynamics (5.164) is characterized for CGMY innovations we propose an approximation which preserves the characterization of the dynamics but replaces the standard CGMY innovations by standard Normal innovations:

$$\sigma_t \simeq \beta_0 + \beta_1 \left[stdNormal + \left\{ \frac{C(M^{Y-1} - G^{Y-1})\Gamma(1-Y)}{\sqrt{C(M^{Y-2} + G^{Y-2})\Gamma(2-Y)}} - \gamma \right\} \sqrt{\sigma_{t-1}} \right]^2 + \alpha_1 \sigma_{t-1} \tag{5.165}$$

The idea behind such approximation is similar to that explained in case of NIG innovations.

We now apply the volatility dynamics (5.165) (which is characterized for CGMY innovation and approximated though the replacement of standard CGMY variate by standard Normal variate.) in general recursive relation (5.147) where scaled CGMY innovation drives the return:

$$\begin{aligned}
&\mathbb{E}[e^{u\log(S_T)} \mid \mathfrak{F}_t^1] \\
= \;& S_t^u \mathbb{E}\Big[e^{u\left(r + \lambda\sigma_{t+1} - \frac{z_{t+1}}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}}\right)} \exp\big[A(t+1,T,u) \\
&\quad + B(t+1,T,u)\Big(\beta_0 + \beta_1 \Big[stdNormal \\
&\quad + \left\{ \frac{C(M^{Y-1} - G^{Y-1})\Gamma(1-Y)}{\sqrt{C(M^{Y-2} + G^{Y-2})\Gamma(2-Y)}} - \gamma \right\} \sqrt{\sigma_{t-1}} \Big]^2 + \alpha_1\sigma_{t-1} \Big)\big] \mid \mathfrak{F}_t^1 \Big]
\end{aligned} \tag{5.166}$$

We apply the relation (5.156) to (5.166) and simplify it further:

[10] Chayawat Ornthanalai(2010)[89] considers a particular form of affine GARCH which in a way resembles jump-diffusion type approaches. Namely they considered two different components for innovations: one is normal and the other coming from some pure jump Lévy process.

$$
\begin{aligned}
&\mathbb{E}[e^{u\log(S_T)} \mid \mathfrak{F}_t^1] \\
\overset{(5.47)}{=}\;& S_t^u \exp\Big\{u\Big(r+\lambda\sigma_{t+1}\Big) \\
&\quad +C\sigma_{t+1}\Gamma(-Y)\Big\{\Big(M-\frac{(-u)}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}}\Big)^{Y}-M^{Y} \\
&\quad +\Big(G+\frac{(-u)}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}}\Big)^{Y}-G^{Y}\Big\}\Big\} \\
&\exp\Big\{A(t+1,T,u)+B(t+1,T,u)\beta_0-\frac{1}{2}log\Big[1-2B(t+1,T,u)\beta_1\Big] \\
&\quad +B(t+1,T,u)\beta_1\left(\frac{C(M^{Y-1}-G^{Y-1})\Gamma(1-Y)}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}}-\gamma\right)^2\sigma_{t+1}\frac{1}{[1-2B(t+1,T,u)\beta_1]} \\
&\quad +\alpha_1 B(t+1,T,u)\sigma_{t+1}\Big\} \\
&\qquad\qquad \text{since } z_{t+1}|\mathfrak{F}_t^1\sim CGMY(C\sigma_{t+1},G,M,Y) \\
=\;& S_t^u\exp\Big\{ur+A(t+1,T,u)+\beta_0 B(t+1,T,u)-\frac{1}{2}log\Big[1-2B(t+1,T,u)\beta_1\Big] \\
&\quad +\Big[\lambda u+\alpha_1 B(t+1,T,u)+C\Gamma(-Y)\Big\{\Big(M-\frac{(-u)}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}}\Big)^{Y}-M^{Y} \\
&\quad +\Big(G+\frac{(-u)}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}}\Big)^{Y}-G^{Y}\Big\} \\
&\quad +B(t+1,T,u)\beta_1\Big(\frac{C(M^{Y-1}-G^{Y-1})\Gamma(1-Y)}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}}-\gamma\Big)^2\frac{1}{[1-2B(t+1,T,u)\beta_1]}\Big]\sigma_{t+1}\Big\}
\end{aligned}
\tag{5.167}
$$

Comparing equation (5.158) with equation (5.144) we obtain the following recursive relations:

$$
\begin{aligned}
A(t,T,u) &= ur+A(t+1,T,u)+\beta_0 B(t+1,T,u)-\frac{1}{2}log\Big[1-2B(t+1,T,u)\beta_1\Big] \\
B(t,T,u) &= \lambda u+\alpha_1 B(t+1,T,u)+C\Gamma(-Y)\Big\{\Big(M-\frac{(-u)}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}}\Big)^{Y}-M^{Y} \\
&\quad +\Big(G+\frac{(-u)}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}}\Big)^{Y}-G^{Y}\Big\} \\
&\quad +B(t+1,T,u)\beta_1\Big(\frac{C(M^{Y-1}-G^{Y-1})\Gamma(1-Y)}{\sqrt{C(M^{Y-2}+G^{Y-2})\Gamma(2-Y)}}-\gamma\Big)^2\frac{1}{[1-2B(t+1,T,u)\beta_1]}
\end{aligned}
\tag{5.168}
$$

We will obtain the option prices through Fourier Inversion as in Heston(1993)[69] and Heston and Nandi(2000)[70]. For closed form (up to numerical integration) GARCH price with CGMY innovations let us denote the model price by $c_{cfgcgmy}$. This model has eight parameters to be estimated: $[\beta_0, \beta_1, \alpha_1, \gamma, C, G, M, Y]$. $c_{cfgcgmy}$ is obtained as in (4.12) but replacing $A(t, T, u)$ and $B(t, T, u)$ recursive relations in (4.11) by those in (5.168).

Given the constraints on several parameters, from our empirical observation we realize that it is more effective to treat the calibration as a constrained optimization problem rather than a simple non-linear least square one. The constraints we need to consider are coming from basic GARCH structure as well as usual CGMY parametrization. Namely we need to ensure that $\beta_0 \geq 0, \beta_1 \geq 0, \alpha_1 \geq 0, \alpha_1 + \beta_1 < 1$; $C, G, M > 0$ and $Y < 2$. Thus to calibrate the model we consider the following optimization problem on each day[11]:

$$\begin{aligned} &Minimize \left[\sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(C^i_{market} - c^i_{cfgcgmy}[\beta_0, \beta_1, \alpha_1, \gamma, C, G, M, Y] \right)^2} \right] \\ &s.t. \qquad A.[\beta_0, \beta_1, \alpha_1, \gamma, C, G, M, Y]' \leq b. \end{aligned} \tag{5.169}$$

Here

$$A = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

and

$$b = [0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 2 - \epsilon]'; \qquad \text{with } \epsilon > 0.$$

5.3.4 The case of VG time changed Lévy Innovations:

We apply approximation to the dynamics (5.124) to uphold the closed form valuation techniques similar to those of Heston and Nandi. When the dynamics (5.124) is characterized for VG innovations we propose an approximation which preserves the characterization of

[11] Again Matlab function "fmincon" can be used to carry out such constrained optimization.

dynamics but replaces the standard VG innovations by standard Normal innovations:

$$\sigma_t \simeq \beta_0 + \beta_1 \left[stdNormal + \left\{\frac{\theta}{\sqrt{\sigma^2+\nu\theta^2}} - \gamma\right\}\sqrt{\sigma_{t-1}}\right]^2 + \alpha_1\sigma_{t-1} \tag{5.170}$$

Similar to other characterizations the idea behind such approximation is motivated by using the useful relation (5.156) yielding analytic valuation.

We now apply the volatility dynamics (5.170) (which is characterized for VG innovation and approximated by standard Normal) in general recursive relation (5.147) where scaled VG innovations drive the returns

$$\begin{aligned}
&\mathbb{E}[e^{u\log(S_T)} \mid \mathfrak{F}_t^1] \\
= \quad & S_t^u \mathbb{E}\Big[e^{u\left(r+\lambda\sigma_t - \frac{z_t}{\sqrt{\sigma^2+\nu\theta^2}}\right)} \exp\big[A(t+1,T,u) \\
&+B(t+1,T,u)\Big(\beta_0 + \beta_1\left[stdNormal + \left\{\frac{\theta}{\sqrt{\sigma^2+\nu\theta^2}} - \gamma\right\}\sqrt{\sigma_{t-1}}\right]^2 + \alpha_1\sigma_{t-1}\Big)\big] \mid \mathfrak{F}_t^1\Big]
\end{aligned} \tag{5.171}$$

We apply the relation (5.156) to (5.171) and simplify it further:

$$\begin{aligned}
&\mathbb{E}[e^{u\log(S_T)} \mid \mathfrak{F}_t^1] \\
\overset{(5.118)}{=} \quad & S_t^u \exp\left\{u\Big(r+\lambda\sigma_{t+1}\Big) - \frac{\sigma_{t+1}}{\nu}log\left(1-u\theta\nu-\frac{1}{2}\sigma^2\nu u^2\right)\right\} \\
& \exp\Big\{A(t+1,T,u) + B(t+1,T,u)\beta_0 - \frac{1}{2}log\Big[1-2B(t+1,T,u)\beta_1\Big] \\
& +B(t+1,T,u)\beta_1\left(\frac{\theta}{\sqrt{\sigma^2+\nu\theta^2}} - \gamma\right)^2 \sigma_{t+1}\frac{1}{[1-2B(t+1,T,u)\beta_1]} \\
& \qquad +\alpha_1 B(t+1,T,u)\sigma_{t+1}\Big\} \\
& \qquad\qquad \text{since } z_{t+1}|\mathfrak{F}_t^1 \sim VG(\sigma\sqrt{\sigma_{t+1}}, \theta\sigma_{t+1}, \frac{\nu}{\sigma_{t+1}}) \\
= \quad & S_t^u \exp\Big\{ur + A(t+1,T,u) + \beta_0 B(t+1,T,u) - \frac{1}{2}log\Big[1-2B(t+1,T,u)\beta_1\Big] \\
& +\Big[\lambda u + \alpha_1 B(t+1,T,u) - \frac{1}{\nu}log\left(1-u\theta\nu-\frac{1}{2}\sigma^2\nu u^2\right) \\
& +B(t+1,T,u)\beta_1\Big(\frac{\theta}{\sqrt{\sigma^2+\nu\theta^2}} - \gamma\Big)^2\frac{1}{[1-2B(t+1,T,u)\beta_1]}\Big]\sigma_{t+1}\Big\}
\end{aligned} \tag{5.172}$$

Comparing equation (5.172) with equation (5.144) we obtain the following recursive relations:

$$
\begin{aligned}
A(t,T,u) &= ur + A(t+1,T,u) + \beta_0 B(t+1,T,u) - \frac{1}{2}log\Big[1 - 2B(t+1,T,u)\beta_1\Big] \\
B(t,T,u) &= \lambda u + \alpha_1 B(t+1,T,u) - \frac{1}{\nu}log\left(1 - u\theta\nu - \frac{1}{2}\sigma^2\nu u^2\right) \\
&\quad + B(t+1,T,u)\beta_1\Big(\frac{\theta}{\sqrt{\sigma^2+\nu\theta^2}} - \gamma\Big)^2 \frac{1}{[1-2B(t+1,T,u)\beta_1]}
\end{aligned}
\tag{5.173}
$$

We will obtain the option prices through Fourier Inversion as in Heston(1993)[69] and Heston and Nandi(2000)[70]. For closed form (up to numerical integration) GARCH price with VG innovation let us denote the model price by c_{cfgvg}. This model has seven parameters to be estimated: $[\beta_0, \beta_1, \alpha_1, \gamma, \sigma, \theta, \nu]$. c_{cfgvg} is obtained as in (4.12) but replacing $A(t,T,u)$ and $B(t,T,u)$ recursive relations in (4.11) by those in (5.173).

As is the case with other innovations we realize that it is more effective to treat the calibration as a constrained optimization problem rather than a simple non-linear least square one. The constraints we need to consider are coming from basic GARCH structure as well as usual VG parametrization. Namely we need to ensure that $\beta_0 \geq 0, \beta_1 \geq 0, \alpha_1 \geq 0, \alpha_1 + \beta_1 < 1$; $\sigma > 0$ and $\theta \in \mathbb{R}$, $\nu > 0$. Thus to calibrate the model we consider the following optimization problem on each day[12]:

$$
\begin{aligned}
&Minimize \left[\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(C^i_{market} - c^i_{cfgvg}[\beta_0, \beta_1, \alpha_1, \gamma, \sigma, \theta, \nu]\right)^2}\right] \\
&s.t. \quad A.[\beta_0, \beta_1, \alpha_1, \gamma, \sigma, \theta, \nu]' \leq b.
\end{aligned}
\tag{5.174}
$$

Here

$$
A = \begin{pmatrix}
-1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -1 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -1
\end{pmatrix}
$$

and

$$
b = [0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0]'.
$$

[12]We use Matlab function "fmincon" to implement such constrained optimization.

5.4 Empirical Results

We use intraday records of options written on S&P500 index and traded at Chicago Board Options Exchange(CBOE). For this empirical part we only implement one of our closed form GARCH-Lévy dynamics, namely GARCH with NIG innovation. We will study other dynamics in future.

After Rubinstein's(1994)[95] suggestions to S&P500 data to test European option pricing models, most of the studies in option pricing literature consider Options traded on this index. In terms of open interest in Options, S&P 500 is the most active index options market and in general it is the second most active index options market in United States.For this index the minimum tick is 1/16 for those series which trades below $3 and the tick is 1/8 for all other series. Strike price spacing are 5 points for near months and 25 points for far away months. The options expire in three near terms months in addition to the months from the quarterly cycle of March, June, September and December.

The intraday data is sampled on every Wednesday. We consider assessing our model's performance with most of the otherwise sophisticated models which we studied in earlier chapters. We accomplish this investigation using various cross-sections of option records. First of all we carry out a pilot survey using the options recorded on the last day in our data set, which is 29th October 2008. Since the discrete time models take considerable time in calibration, Heston and Nandi(2000)[70] didn't consider more than six months at a time. They carry out year-by-year calibrations, considering first six months as in-sample period and the second six months as out-of-sample. So to have similar views of pricing performance as Heston and Nandi(2000)[70], we carry out some year by year calibrations. Like Heston and Nandi(2000)[70] we will also consider first six months records for in-sample calibration and will use the second six months records to assess the models out-of-sample performance in case of year-by-year calibrations. We further consider two years aggregation of option records. Finally we consider long three years recent option records, all together, to calibrate the models. We mention here that such a calibration is possible only because we adapt FRFT approach to pricing.

5.4.1 Data Cleaning Issues

To clean the data we use the same rules as applied by Heston and Nandi(2000)[70]:

- We do not represent an option, with a particular moneyness or maturity criteria, more than once in our sample. This removes quite a good number of options. When records get repeated with same moneyness and maturity corresponding to same or different index levels, we just consider the first record.
- We exclude very deep out-of-the-money and deep in-the-money options. We do that because these options are either infrequently traded and/or have low enough prices as for the bid-ask spread to constitute a major portion of the price. Only the records having index to strike ratio somewhere between 0.9 and 1.1 are included in our sample.
- Heston and Nandi(2000)[70] applied the maturity filtering based on the criteria that options should have days to expiration somewhere between 6 and 100. Their argument goes with the fact that very long term options are not actively traded and are prone to be mispriced. Similarly they argued that very short term options have substantial time decay and create trouble in isolating volatility parameters. However we notice that Bakshi, Cao and Chen(1997)[7] used options with all available maturities in the sample and perhaps it helps them recognize the models performance rather distinctly. However for this research we will consider only options with 6-100 days to maturity.

5.4.2 Practical Issues in Implementation

We calibrate all the models–except Gram-Charlier–that we studied for one day traded options in chapter four. There are many practical issues, around the calibration of different models, which are worth reporting. A number of those issues are reported here:

- For any continuous (in time not in path) model for a given set of parameters the pricing involves one evaluation of characteristic function. e.g. if the option has maturity t=150 days then for continuous time models the prices come through Carr-Madan formula which uses the characteristic function values only for t=150/252. However if we consider a discrete time GARCH model ,say Heston-Nandi model, prices still use only the characteristic function values corresponding to t=150. However to obtain these characteristic function values at t=150, we need to encounter similar evaluation of characteristic function on every previous day and update the CF values as suggested by model dynamics. Thus when a continuous time model involves just one evaluation, the discrete time model involves 150 evaluations. This clarifies why discrete time models calibration is so time consuming compare to continuous time models. Furthermore underlying this fact there is another subtle issue. Calibration on a longer horizon naturally includes options of various maturities thus can update the true characteristics of the discrete time models. In other words to realize the true potential of discrete time models, models parameters should be calibrated on options recorded on a longer horizon. In case discrete time models need to be calibrated on a short horizon e.g. few options traded on a single day, nothing will be surprising if we see discrete time models are performing worse compare to continuous time models, though on a reasonably longer time record it is typically the opposite. e.g. we typically observe that for calibration with one day traded options Heston93 model outperforms Heston-Nandi model where as for calibration with options traded over a horizon of a year or so it is typically the opposite. This can be particularly so if significant number of single day options do not have higher days to maturity.

- Another relevant observation is that it is possible to encounter situations where four parameter CGMY model performs better than three parameter VG and NIG models but it is not necessarily always the case.

- We observe that calibration of CGMY model using option prices is very tricky indeed. In particular the calibration is heavily dependent on the initial choice of 'Y' and its range of variability. We found it is worth trying this range as : (-N, $-\epsilon$), (-N, 2-ϵ), (ϵ ,1-ϵ),(ϵ,2-ϵ),(1+ϵ,2-ϵ) or even (1.5+ϵ,2-ϵ); where $N > 2$. Allowing 'Y' to vary on any of these ranges could yield significantly different calibration result. So for a particular set of data in hand it is more like applying trial and error method to decide which range provides best calibration. We recall that 'Y' alone characterizes the nature of the underlying CGMY process, i.e. whether the process is of finite variation, infinite variation, finite activity, or infinite activity.

- The reason for DE models RMSE sometimes getting even slightly worse than BS models RMSE goes with the fact that BS model is calibrated using its standard formula(with no numerical approximation) where as DE model is calibrated by approximating the pricing integrals using FRFT. We typically observed that shall we use FRFT for BS as well, DE RMSE gets slightly better. However we further typically observed that DE jump diffusion model hardly ever performed significantly better than BS model.

- We must note one subtle issue about computation of standard errors. We numerically calculate the Fishers information matrices to obtain the standard errors for various models. The information matrices are obtained for mean square error (MSE) functions. Finite difference scheme is utilized to obtain the derivatives. However even with our best efforts we failed to apply the same perturbation for all the models when we apply the finite difference to MSE's. This is because for our GARCH type models some parameters are very small, often to the magnitude of 10e-5 or 10e-6. So when we consider a perturbation say 0.0005 or even 0.00005 this amounts are of magnitude much higher than the parameter themselves. Thus the total dynamics of the model collapse and often we obtain something bizarre. On the other hand if we apply a perturbation much smaller say 0.00000005 or even 0.00000005 then it works for GARCH type models but generates some sort of instability in other models for which a perturbation of 0.0005 or 0.00005 works well. Though theoretically we should apply same perturbation to all models to have the SE's comparable across models, we just couldn't do that at this stage.

- There is an important point to remember here. There is no established literature which confirms that a particular model will perform better than others across all cross-sections of data. In fact, as mentioned in Wim Schoutens(2003)[102], a model performing best on a particular data set may perform worst once the data set is changed. Though Wim Schoutens made this comment while investigating the relative performance of Levy models, we empirically observed that this is true for all models of our investigation. Saying otherwise, there is no guarantee that models with more parameters will always perform better than those with fewer parameters. e.g. restricted non-updated version of Heston and Nandi's model has five parameters but it was shown that one parameter Black-Scholes model systematically outperformed this particular version of Heston and Nandi's model, see Heston and Nnadi(2000)[70].

- Another important point– perhaps a vital issue in terms of empirical comparison– is that though under statistical measure the innovations are normal in Heston-Nandis(2000)[70] CFG model , it turns into non-normal under the risk-neutral distribution. More precisely any value of λ other than -0.5 will lead to the case where innovations under risk-neutral dynamics in Heston-Nandi's 2000 model could potentially follow some kind of non-normal distribution. In this case it doesn't have any explicit distributional characterization: the best we can say is that for $\lambda \neq -\frac{1}{2}$ the distribution of innovations under risk-neutral dynamics is not guaranteed to be normal. See Heston and Nandi(2000)[70]. Now the relative performance of non-normal innovations induced through the measure change in Heston and Nandis(2000)[70] and the non-normal innovations explicitly incorporated explicitly by Lévy processes is a nice empirical work left for future. For NIG Lévy innovations, only, the following sections will carry out rigorous investigation compare to Heston and Nandi's(2000)[70] model.

- Heston and Nandi considered four specifications in their paper. First in terms of volatility updates, they used the terms "non-updated" and "updated". By "non-updated" they meant that the parameters used in predicting volatility are calibrated once only, using previous one years daily returns. However by "updated" they meant that the task of estimation is performed on every week, considering a rolling window of one year daily returns. For each of these specifications they further considered "restricted" and "unrestricted" versions of the model. In restricted version they didn't allow the risk neutral distribution to have skewness, so it yields a symmetric distribution. However in unrestricted version the risk neutral distribution could possibly exhibit skewness. Given the empirical findings in their paper, see Heston and Nnadi(2000)[70], we consider restricted and unrestricted versions only for 'updated' specification. Heston and Nandi(2000)[70] demonstrated that non-updated restricted version often performs worse than Black-Scholes.

- The fact that some specifications of Heston and Nandi's(2000)[70] model perform worse than Black Scholes model–when innovations under risk neutral dynamics are again restricted to follow some symmetric distribution–is very much uncomfortable from the modelling point of view. Recently Byan and Min(2010)[25] conjectured that this is because Heston and Nandi(2000)[70] used the volatility which is estimated under statistical measure and then feed this volatility directly into the pricing formula which requires the underlying to follow a risk-neutral dynamics. Despite the criticism it encounters we implement Heston and Nandi's(2000)[70] model as in their original paper. For our CFG-NIG model, however, we avoid this controversy and completely ignore estimation under statistical measure. We observe that the recursive version of the characteristic function in Heston and Nandi(2000)[70], and in other similar works, is a mathematically rich tool which implicitly accommodates the GARCH structure of the volatility and an specification of the return dynamics incorporating the effects of that volatility. In other words heteroskedasticity is embedded into the characteristic function itself. So it may not be that significant to ensure that the initial one period ahead volatility itself is predicted from a GARCH process under statistical measure. We further observe that this will help us reduce any adverse effect which could possibly arise from approximation of volatility dynamics. The approximation is required to uphold analytic pricing. In effect this initial one period ahead volatility is not a parameter in the calibration. It is a parameter which we directly feed into the model. So the more realistic this feeded value is, the better will be the performance of the model. Thus we realistically extract the volatilities from forward looking market option prices. In other words we completely ignore historical stock returns and proxy the one period ahead volatilities by the average implied volatilities of the options used in calibration on a given day. In addition to rendering the historical return series completely redundant this ensures that we are using only forward looking option information in calibration. In essence this approach is similar to ad-hoc Black-Scholes model. However in ad-hoc Black-Scholes model at first a volatility function of maturity and strike is fitted using a cross section of option prices, then that functional relationship is used to obtain implied volatility for each particular option characterized by strike and maturity pair. The focus is similar : using only forward looking option information in calibration. However we leave this volatility function fitting issue, on top of implementation of our model, for a future work and simply use the average of daily implied volatilities to feed directly into the model as one-period ahead volatility. [13]

- Though we didn't report, we carried out similar calibration for Heston and Nandi's model as in CFG-NIG model. i.e. ignoring the time series of returns and using the average daily implied volatilities as a proxy to one period ahead volatility. We observed striking improvement in the performance of Heston Nandi(2000)[70] model. e.g. even the symmetric case of this model now performs robustly well compare to Back-Scholes model, which Heston and Nandi reported as always performing worse than Black-Scholes when volatilities are estimated from historical returns. See Heston and Nandi(2000)[70]. As reported in Heston and Nandi(2000)[70] the asymmetric case is obviously expected to perform even better and we found that it is the case. However we reported the case of Heston and Nandi exactly as it is reported in their paper, i.e. using the volatility which is filtered from the time series of historical returns.

- When assessing out-of-sample performance we restrict to models which explicitly incorporates stochastic volatility. It is our observations that all otherwise sophisticated characterizations though could possibly perform robustly well for instantaneous fitting (e.g. one day or couple of days observations), for long time series of option records these models perform way poorer compare to models which consider explicit stochastic volatility dynamics. For example models of pure-jump, Gram-Charlier type could fit the data well on a single day but for a long time series of option records they fail to outperform any model which considers explicit stochastic volatility dynamics. What

[13]Though we report the results with average implied volatility as a proxy to one-step ahead GARCH volatility we, however, separately calibrate the models with one-day-ahead initial volatilities obtained from GARCH-NIG dynamics. We found that the results are overall promising even when the volatilities are estimated from historical asset returns. By promising we mean that overall the improvement achieved over Heston and Nandi(2000)[70] model is significant.

left seeing is how these models with explicit volatility dynamics fare with the so called Lévy stochastic volatility models and will be considered in some future work.

- For assessing in-sample goodness of fit we use the universal measure RMSE, as defined in equation(4.48). We apply two other measures to assess the out-of-sample goodness of fit. One of them is the most naive measure known as average absolute error(AAE) and is defined as:

$$AAE = \sum_{i=1}^{N} \frac{|model\ \ price_i - market\ \ price_i|}{N} \tag{5.175}$$

  The other measure we use in out-of-sample assessment is a special measure which is used to get the idea whether on an average the model exhibits overpricing or underpricing tendency. It is known as Mean-Outside-Error(MOE)[14]:

$$MOE = \frac{1}{N}\sum_{i=1}^{N}\Big[(model\ price_i - ask_i)\mathbb{I}_{\{model\ price_i > ask_i\}} + (model\ price_i - bid_i)\mathbb{I}_{\{model\ price_i < bid_i\}}\Big] \tag{5.176}$$

5.4.3 A Pre-calibration Pilot Survey

We consider the pilot calibration to empirically reinforce the necessity of models with stochastically richer innovations. Saying otherwise this pilot calibration focuses how disastrous the performance of otherwise sophisticated models could be. For this we consider the options traded on 29/10/2008, the last day records we have in our data set. We clean the data following the set rules mentioned earlier. The pilot calibration results are reported in table 5.1 [15] and for three competing models the ARPE's for this pilot study are presented in figure 5.1.

[14] RMSE considers quadratic deviations between model and market prices, AAE considers linear deviations and MOE is a special indicator of models pricing behavior.

[15] For this as well as all other calibration we used FRFT with same parameter values, including the dampening factor. This is essential to compare models performance considering calibration using FRFT. The consistent way of choosing FRFT parameters are discussed in Lee(2004)[78]. In particular a uniformly suitable value of dampening factor could be one around two, see e.g. Chourdakis(2008)[30].

5.4.4 Calibration Using Data from Janury'2005 to December'2007

Our pilot survey exemplifies what can happen on a very rough day: available models can simply collapse. We need models with more sophisticated stochastic properties. In this section we carry out similar calibrations using different cross-sections of options recorded on a wider time frame : from January'2005 to December'2007. After applying the filtering rules, as described above, we have 8931 options to consider on this time window. We consider calibrating models under three information aggregation schemes. The first scheme corresponds to calibrating models using options traded on first six months of each year and then assessing models out-of-sample performance using options traded on the remaining six months of the year. This is exactly what Heston and Nandi(2000)[70] did in their work. However Heston and Nandi(2000)[70] considered the years: 1992, 1993 and 1994. We consider the years: 2005,2006 and 2007.

Table5.2 and figure5.3 report the in-sample and out-of-sample performance, respectively, for 2005 calibration. Similar results for 2006 contracts are reported in table5.3 and figure5.4 where as table5.4 and figure5.5 report the results for 2007 contracts. Our second scheme considers calibration of models using information contained in option contracts traded over two year periods. More precisely, under this scheme, we calibrate the models using options traded on 2005-2006 and 2006-2007. In first case we use first six months contracts of 2007 to assess models out-of-sample fitting performance and in second case the contracts of first three months of 2008 (these are the most recent option contracts recorded in our data set) are used for out-of-sample assessment. Table5.5 and figure5.6 report the in-sample and out-of-sample results, respectively, for 2005-2006 calibration; where as table5.6 and figure5.7 report the similar results for 2006-2007 calibration. Our third scheme considers all three years contracts at a time. Table5.7 reports the calibration results corresponding to this scheme, where as fugure5.8 presents the out-of-sample assessment for this scheme.

A remaining puzzle in empirical option pricing literature, see Bates(2003)[11], is to address the issue which will help us quantify the degree to which cross-sectional option pricing patterns are quantitatively consistent with the time series patterns of the underlying asset prices. An approach–though counterfactual to the standard martingale hypothesis of asset price dynamics–could be investigated in this regard. The risk neutral characteristic functions which are often considered to price options and to reveal cross-sectional option pricing patterns, could be used to reveal the time series properties of the underlying asset as well. As a tool empirical characteristic function method could be utilized. In particular under this setting cross-sectional patterns and time series properties should emerge to be similar. But in practice this issue is not that conforming. As noted by Bakshi et al(2000)[8], instantaneous option price evolution is not fully captured by underlying asset price movements. Furthermore time series properties premised on stationary Markov assumption are presumably under regular bombardment through heteroscedasticity of GARCH model. The degree to which conformity could possibly be achieved should in principle rely on the characteristic functions of the model which could possibly be employed to both asset based time series estimation as well as option based cross-sectional estimation. Thus models risk-neutral characteristic functions, and the parameters there of, should partially explain the degree of conformity between the time series properties and cross-sectional patterns. In future we will investigate such conformity in relative sense, involving most of the apparently successful and otherwise sophisticated alternatives to Black-Scholes model. With our schemes, however, we notice that cross-sectional calibration with gradually increasing amount of information aggregation systematically prefer the GARCH-NIG model: though for six month aggregation we can figure out an instance, at least, when unrestricted version of Heston and Nandi(2000)[70] model outperformed GARCH-NIG model, for two and three years information aggregation GARCH-NIG model distinctly outperforms all otherwise sophisticated models.[16]

There are several reasons for CFG-NIG model to perform robustly well. First of all given that CFG-NIG model has separate characterization to describe conditional evolution of skewness(equation (5.57)) and conditional evolution of kurtosis(equation (5.58)), the prices can accommodate the cross–strike and cross-maturity features better than other models. The main difference with Heston and Nandi's(2000)[70] model arises from the fact

[16]As we mentioned earlier, we implement GARCH-NIG model separately with one-day ahead volatility predicted from asset prices–exactly as Heston and Nandi(2000)[70] did for GARCH-Normal model–and found that overall the improvement achieved is substantial compare to Heston and Nandi's model. However, we did not report these results.

that in their model skewness and kurtosis are captured by structural parameters of GARCH model[17] where as our characterization of CFG-NIG model captures the skewness and kurtosis in time varying fashion, with the variation generated by time varying volatility. We must mention that this richness is solely a feature of non-normal innovation. This reminds a conjecture made in Bates(2003)[11] regarding continuous time SVJ[18](or SVJJ) model: "having jump components addresses moneyness biases, while having stochastic laten variables allow distributions to evolve stochastically overtime". Though this conjecture is made in continuous time SVJ(or SVJJ) models the features are consistent with our CFG-NIG model as well. However SVJ(or SVJJ) model faces the criticism of Markovian structure where as our CFG-NIG dynamics incorporates non-Markovian time series properties through heteroscedasticity. This added feature is expected to give CFG-NIG type models a preferential edge over SVJ(or SVJJ) model. An empirical work should be of interest to clarify the evidence and will be considered in future.

[17]e.g. in equation(4.9) θ determines skewness and α determines the kurtosis, see Heston and Nandi(2000)[70].

[18]It is standard in the literature to write SVJ for dynamics which include jumps only in return and SVJJ for dynamics which include jumps in both return and volatility.

Model	RMSE	Parameters						
BS	13.5886	(σ)						
		0.5172						
		(0.0048)						
VG	12.7437	(σ)	(θ)	(ν)				
		0.0899	-3.2635	0.0279				
		(0.0008)	(0.0009)	(0.0006)				
NIG	12.8452	(α)	(β)	(δ)				
		418.3021	-409.7441	1.0347				
		(0.0251)	(0.0250)	(0.0180)				
JD-DE	12.606	(σ)	(λ)	(p)	(η_1)	(η_2)		
		3.1230e-4	33.4824	0.4741	14.8251	14.6700		
		(0.0577)	(0.2407)	(0.2540)	(0.2047)	(0.1888)		
CGMY	12.209	(C)	(G)	(M)	(Y)			
		1.7432e8	3.5792e1	9.6529e5	-6.0337			
		(0.089)	(0.087)	(0.088)	(0.006)			
HS	11.069	(κ)	(θ)	(σ)	(ρ)	(V_0)		
		0.2175	1.1198	0.6980	-0.9900	0.2992		
		(0.0200)	(0.0486)	(0.0460)	(0.0471)	(0.0058)		
HN(R)	7.3703	(α_1)	(β_1)	(β_0)	(γ)	(λ)		
		0.5876	4.150e-6	9.318e-6	299.993	-300.493		
		(0.151e-23)	(0.082e-23)	(0.152e-23)	(0.181e-23)	(0.292e-23)		
HN(U)	7.2563	(α_1)	(β_1)	(β_0)	(γ)	(λ)		
		0.957	2.567e-9	2.567e-9	3.6154	7.3960		
		(0.001)	(1.404e-7)	(1.395e-7)	(0.004)	(0.005)		
CFGNIG	7.0439	(α_1)	(β_1)	(β_0)	(γ)	(α)	(β)	(δ)
		0.9772	2.56e-9	2.56e-9	-4.3189	32.0734	30.3958	21.4260
		(8.579e-4)	(7.804e-7)	(1.124e-6)	(0.0348)	(0.0249)	(0.0191)	(0.0194)

Table 5.1: *Pilot calibration with Options written on S&P500 index traded at CBOE. We consider options traded on 29/10/2008. After all cleaning we have 69 records to consider, on that particular day, with mean option price of 77.4942. The mean annual implied volatility on 29/10/2008 was 0.5599. Standard errors are obtained by numerically computing the Fisher's information matrix for mean squared error(MSE) function. Calibration is carried out by applying FRFT approach to price options.*

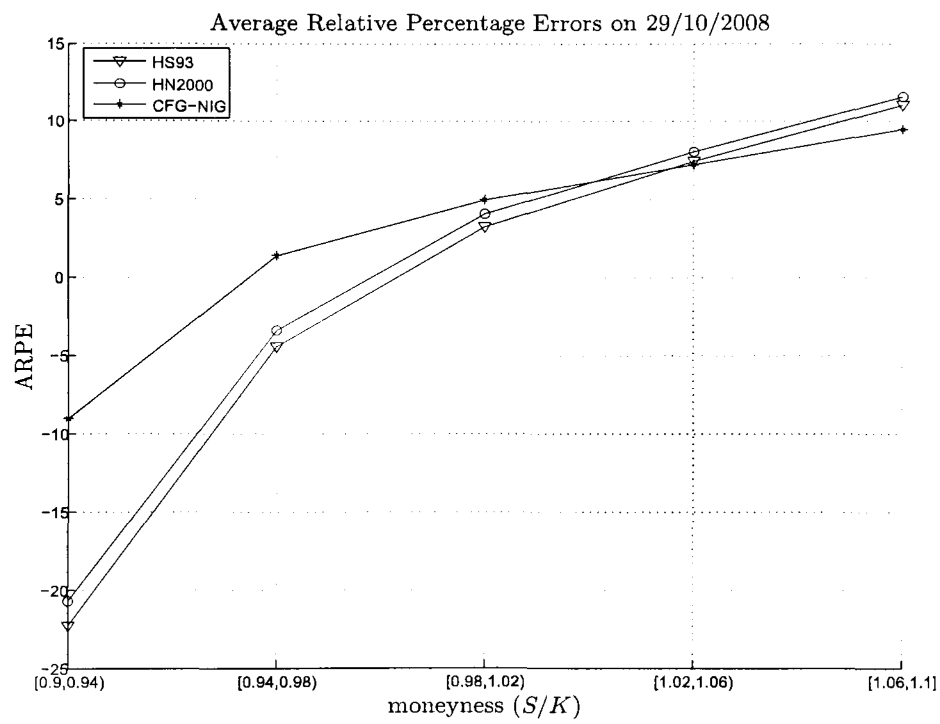


Figure 5.1: *Average Relative Percentage Errors of continuous time Heston's 93 stochastic volatility model, discrete time Heston-Nandi 2000 GARCH model and CFG-NIG model(closed form GARCH with NIG innovations) on 29/10/2008. Both Heston's 93 and Heston-Nandi 2000 models have stochastic properties governed by normal distribution. CFG-NIG replaces conditional normal innovations by conditional Normal Inverse Gaussian(NIG) Lévy innovations.*

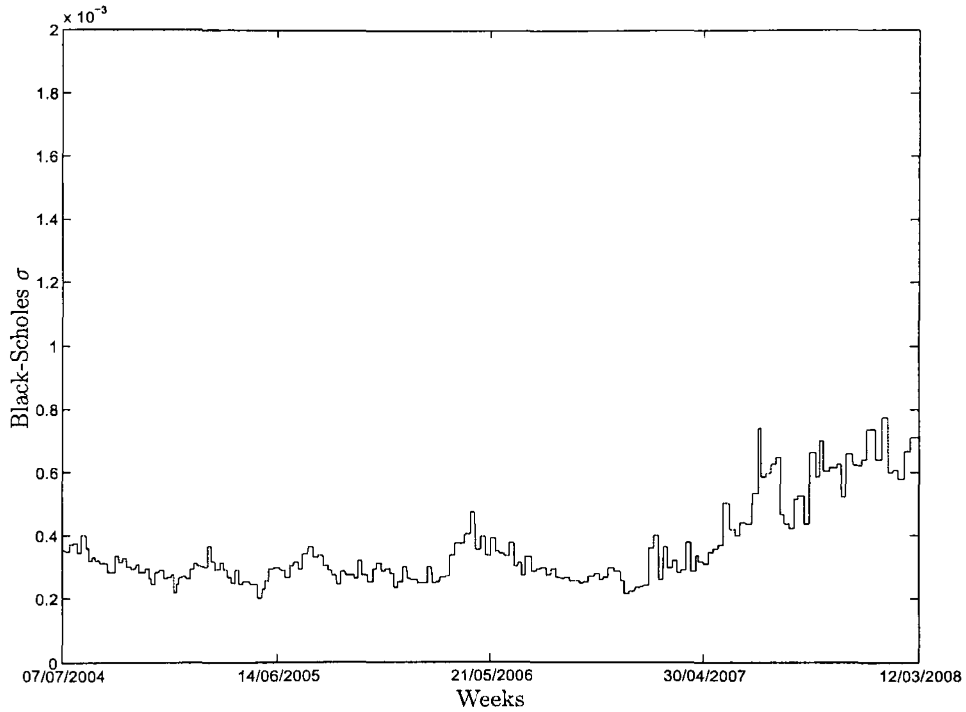


Figure 5.2: *Weekly variability of Black-Scholes σ (annual). This goes against one of the fundamental assumptions of the benchmark model: the volatility remains constant over time. In other words this gives the idea of how turbulent the market is with respect to the Black-Scholes model. However for this observation period the variability is rather mild, specially until 2007.*

Model	RMSE	Parameters						
BS	1.9993	(σ) 0.0678 (0.0059)						
VG	1.8104	(σ) 0.0144 (0.0148)	(θ) 0.2319 (0.0132)	(ν) 0.0828 (0.0119)				
NIG	1.8157	(α) 1621.6 (0.0178)	(β) 1584.1 (0.0177)	(δ) 0.0743 (0.0130)				
JD-DE	1.9995	(σ) 0.0678 (0.0063)	(λ) 0.0813 (0.0190)	(p) 0.9998 (0.0166)	(η_1) 186.1660 (0.0178)	(η_2) 215.3270 (0.0702)		
CGMY	1.8088	(C) 43.5396 (0.7759)	(G) 558.8773 (0.7200)	(M) 57.5013 (0.7389)	(Y) -0.2960 (0.0493)			
HS	1.8386	(κ) 0.1247 (0.0897)	(θ) 0.0427 (0.0482)	(σ) 0.1032 (0.0823)	(ρ) 0.9900 (0.1212)	(V_0) 0.0041 (0.0010)		
HN(R)	1.9483	(α_1) 2.2204e-16 (2.5719e-6)	(β_1) 2.5670e-9 (5.2326e-7)	(β_0) 1.8374e-5 (1.9018e-6)	(γ) 419.0836 (1.0037e-5)	(λ) -419.5836 (1.0037e-5)		
HN(U)	1.4275	(α_1) 0.4583 (0.0058)	(β_1) 2.5670e-9 (5.0629e-7)	(β_0) 1.4438e-5 (2.9606e-6)	(γ) 419.1233 (0.0111)	(λ) -3.5116 (0.0071)		
CFGNIG	1.5931	(α_1) 0.9937 (0.0059)	(β_1) 2.5670e-15 (1.1929e-7)	(β_0) 2.5670e-15 (1.2268e-7)	(γ) -0.2690 (5.5020)	(α) 98.6165 (0.0079)	(β) -67.0771 (0.0146)	(δ) 39.8413 (0.0677)

Table 5.2: *Calibration with Options traded over the period January'2005-June'2005. We consider Options traded on every Wednesday. After all cleaning we have 1179 option contracts with a mean option price of 17.8030 and average implied volatility of 0.0746. Standard errors are obtained by numerically computing the Jacobian of mean squared error(MSE) function. We applied FRFT approach to price options which significantly reduces the calibration time. Discrete time Heston-Nandi (HN) closed form GARCH model requires longer time in calibration than any continuous time model and the requirement is even longer for our CFGNIG model.*

Days to maturity<40	RMSE	AAE	MOE
0.95<=S/K<0.99			
BS	1.0405	0.6750	-0.3051
HS	1.2368	0.9412	-0.0972
HN(R)	1.2721	0.9784	-0.1165
HN(U)	1.3889	1.0589	0.3753
CFG-NIG	1.0334	0.8146	0.3384
0.99<=S/K<1.01			
BS	2.1963	1.7863	-0.5754
HS	2.3926	1.8329	0.6608
HN(R)	2.4326	1.8367	0.8263
HN(U)	2.5346	1.9334	0.9698
CFG-NIG	2.0258	1.3896	0.7578
1.01<=S/K<1.05			
BS	1.8157	1.4391	0.1867
HS	1.7861	1.4695	-0.2292
HN(R)	1.7380	1.3939	-0.0648
HN(U)	2.1064	1.8227	-0.6162
CFG-NIG	1.5677	1.3433	-0.3082
40=< Days to maturity<70			
0.95<=S/K<0.99			
BS	2.0020	1.5002	-0.8276
HS	1.7341	1.3281	-0.1506
HN(R)	1.8120	1.3706	-0.4693
HN(U)	1.7589	1.3837	0.1365
CFG-NIG	1.0261	0.7519	-0.3082
0.99<=S/K<1.01			
BS	2.3768	1.9612	-0.6544
HS	2.3556	1.9295	-0.6386
HN(R)	2.2228	1.8144	-0.2610
HN(U)	2.3292	1.9080	-0.5102
CFG-NIG	2.0453	1.8278	-1.0576
1.01<=S/K<1.05			
BS	2.4496	1.9381	0.6805
HS	2.1738	1.7432	-0.2123
HN(R)	2.2725	1.8186	0.4496
HN(U)	2.7070	2.2402	-1.1189
CFG-NIG	1.8276	1.5552	-0.5203
70=< Days to maturity=<100			
0.95<=S/K<0.99			
BS	2.2828	1.8776	-1.1226
HS	1.9622	1.6261	-0.3238
HN(R)	2.0627	1.6884	-0.8817
HN(U)	1.9968	1.6419	-0.6242
CFG-NIG	2.4701	2.1776	-1.5039
0.99<=S/K<1.01			
BS	2.0780	1.7208	-0.5826
HS	2.3804	2.0773	-1.0115
HN(R)	1.9467	1.5789	-0.4403
HN(U)	2.8036	2.4672	-1.5063
CFG-NIG	3.3416	3.1926	-2.2256
1.01<=S/K<1.05			
BS	2.3978	1.7040	0.7137
HS	2.2601	1.9177	-0.4136
HN(R)	2.3097	1.6734	0.5907
HN(U)	3.8670	3.5808	-2.5468
CFG-NIG	2.3657	2.1428	-0.8189

Figure 5.3: *Out-of-sample valuation errors for Call Options traded in second half of 2005. The models are calibrated on first half of the same year. Total number of contracts available for the second half is 1456. BS stands for Black-Scholes model, HS stands for Hestons'93 stochastic volatility model, HN(R) stands for restricted version of Heston and Nandis 2000 GARCH model, HN(U) stands for unrestricted version of Heston and Nandis 2000 GARCH model, CFG-NIG stands for closed form GARCH model with NIG innovations. RMSE is the root mean square error as defined in*(4.48), *AAE is the average absolute error as defined in*(5.175) *and MOE is the mean outside error as defined in*(5.176).

Model	RMSE	Parameters						
BS	2.9708	(σ)						
		0.0750						
		(0.0050)						
VG	2.9603	(σ)	(θ)	(ν)				
		0.0410	0.9197	0.0046				
		(0.0012)	(0.0084)	(0.0009)				
NIG	2.9622	(α)	(β)	(δ)				
		96.1269	28.6945	0.4713				
		(0.0893)	(0.1018)	(0.0629)				
JD-DE	2.9708	(σ)	(λ)	(p)	(η_1)	(η_2)		
		0.0750	0.0782	0.9998	202.6066	293.0997		
		(0.0050)	(17.2433)	(28.8536)	(18.3369)	(28.4823)		
CGMY	2.9603	(C)	(G)	(M)	(Y)			
		80.6780	1879.1	182.2822	0.1689			
		(0.6236)	(0.3425)	(0.3419)	(0.0272)			
HS	2.9284	(κ)	(θ)	(σ)	(ρ)	(V_0)		
		0.3451	0.0399	0.0255	0.9900	0.0042		
		(0.0808)	(0.0181)	(0.0794)	(0.1040)	(0.0009)		
HN(R)	2.9794	(α_1)	(β_1)	(β_0)	(γ)	(λ)		
		2.2204e-16	2.5670e-9	2.2606e-5	419.1042	-419.6042		
		(1.6569e-5)	(3.0063e-6)	(5.81165e-6)	(1.4440e-4)	(2.0761e-4)		
HN(U)	2.9383	(α_1)	(β_1)	(β_0)	(γ)	(λ)		
		0.6476	2.5670e-9	8.7016e-6	416.6406	-1.3804		
		(0.0082)	(2.4114e-7)	(1.3053e-6)	(0.0073)	(0.0078)		
CFGNIG	1.7388	(α_1)	(β_1)	(β_0)	(γ)	(α)	(β)	(δ)
		0.9999	2.567e-15	2.567e-15	-0.5123	42.6940	2.4246	28.4220
		(0.0039)	(1.0508e-7)	(1.0505e-7)	(1.2015)	(0.5236)	(0.7626)	(0.2833)

Table 5.3: *Calibration with Options traded over the period January'2006-June'2006. We consider Options traded on every Wednesday. After all cleaning we have 1607 option contracts with a mean option price of 30.0293 and average implied volatility of 0.0776. Standard errors are obtained by numerically computing the Jacobian of mean squared error(MSE) function. We applied FRFT approach to price options which significantly reduces the calibration time. Discrete time Heston-Nandi (HN) closed form GARCH model requires longer time in calibration than any continuous time model and the requirement is even longer for our CFGNIG model.*

	RMSE	AAE	MOE
Days to maturity<40			
0.95<=S/K<0.99			
BS	1.3903	0.9944	0.3712
HS	1.7616	1.3949	0.2837
HN(R)	1.9296	1.5183	0.6264
HN(U)	2.1056	1.6556	0.9438
CFG-NIG	1.4193	1.0394	0.2820
0.99<=S/K<1.01			
BS	2.8850	2.0902	0.2575
HS	3.4445	2.5600	1.2225
HN(R)	3.6618	2.8188	1.6466
HN(U)	3.8262	2.9776	1.8701
CFG-NIG	2.9867	1.8272	1.2864
1.01<=S/K<1.05			
BS	1.9321	1.4097	-0.0241
HS	2.2159	1.6252	-0.5075
HN(R)	2.1587	1.5720	-0.3456
HN(U)	2.2117	1.6175	-0.3794
CFG-NIG	1.7697	1.3029	-0.2520
40=< Days to maturity<70			
0.95<=S/K<0.99			
BS	1.8624	1.5761	0.4292
HS	2.0128	1.7381	0.7283
HN(R)	2.0024	1.7141	0.7559
HN(U)	2.1940	1.8767	1.0591
CFG-NIG	0.6285	0.4626	-0.0667
0.99<=S/K<1.01			
BS	2.4526	1.9448	-0.4890
HS	2.4416	1.9765	-0.3513
HN(R)	2.4271	2.0282	-0.1633
HN(U)	2.4347	2.0397	-0.1486
CFG-NIG	1.5991	1.4484	-0.6153
1.01<=S/K<1.05			
BS	2.0036	1.7081	-0.0488
HS	1.9943	1.5672	-0.3800
HN(R)	1.9513	1.6181	-0.2184
HN(U)	2.0288	1.6078	-0.5968
CFG-NIG	1.4420	1.2112	-0.4186
70=< Days to maturity=<100			
0.95<=S/K<0.99			
BS	2.1261	1.8366	0.2257
HS	2.8275	2.5788	1.4105
HN(R)	2.2110	1.9717	0.4555
HN(U)	2.2935	2.0484	0.6292
CFG-NIG	1.1418	0.9538	-0.3438
0.99<=S/K<1.01			
BS	2.0919	1.5134	-0.3394
HS	2.2421	1.8386	0.2431
HN(R)	2.0674	1.5331	-0.2288
HN(U)	2.1045	1.5383	-0.4178
CFG-NIG	1.9404	1.8672	-0.9284
1.01<=S/K<1.05			
BS	2.0482	1.7353	-0.0187
HS	1.9405	1.6451	0.0688
HN(R)	1.9937	1.6607	-0.1313
HN(U)	2.2474	1.6581	-0.8371
CFG-NIG	1.4285	1.2307	-0.2448

Figure 5.4: *Out-of-sample valuation errors for Call Options traded in second half of 2006. The models are calibrated on first half of the same year. Total number of contracts available for the second half is 1606. BS stands for Black-Scholes model, HS stands for Hestons'93 stochastic volatility model, HN(R) stands for restricted version of Heston and Nandis 2000 GARCH model, HN(U) stands for unrestricted version of Heston and Nandis 2000 GARCH model, CFG-NIG stands for closed form GARCH model with NIG innovations. RMSE is the root mean square error as defined in*(4.48), *AAE is the average absolute error as defined in*(5.175) *and MOE is the mean outside error as defined in*(5.176).

Model	RMSE	Parameters						
BS	3.7362	(σ)						
		0.0775						
		(0.0044)						
VG	3.7212	(σ)	(θ)	(ν)				
		0.0185	-1.0296	0.0055				
		(0.0009)	(0.0023)	(0.0007)				
NIG	3.7214	(α)	(β)	(δ)				
		874.7253	-753.9378	0.6988				
		(0.1114)	(0.1092)	(0.0787)				
JD-DE	3.7234	(σ)	(λ)	(p)	(η_1)	(η_2)		
		0.0623	9.7943	0.7093	97.5983	82.5284		
		(0.0035)	(3.0777)	(0.9281)	(17.9164)	(26.8777)		
CGMY	3.7213	(C)	(G)	(M)	(Y)			
		97.3378	153.0980	620.8816	0.0700			
		(0.0216)	(0.0128)	(0.0141)	(0.0106)			
HS	3.6996	(κ)	(θ)	(σ)	(ρ)	(V_0)		
		0.2971	0.0405	0.0343	-0.9900	0.0049		
		(0.0716)	(0.0195)	(0.0698)	(0.0800)	(0.0008)		
HN(R)	3.7021	(α_1)	(β_1)	(β_0)	(γ)	(λ)		
		2.2204e-16	2.5670e-9	2.2606e-5	419.1042	-419.6042		
		(0.0040)	(4.0221e-7)	(1.361e-6)	(0.0020)	(0.0042)		
HN(U)	3.6702	(α_1)	(β_1)	(β_0)	(γ)	(λ)		
		0.4992	2.0502e-6	1.8777e-6	419.2132	-1.4783		
		(0.0055)	(9.4270e-8)	(5.4676e-7)	(0.0054)	(0.0059)		
CFGNIG	2.2825	(α_1)	(β_1)	(β_0)	(γ)	(α)	(β)	(δ)
		0.9999	2.567e-15	7.5968e-8	-0.4985	35.2867	12.1942	26.4248
		(0.0035)	(1.0341e-7)	(1.0371e-7)	(0.2071)	(0.1800)	(0.1616)	(0.1162)

Table 5.4: *Calibration with Options traded over the period January'2007-June'2007. We consider Options traded on every Wednesday. After all cleaning we have 1578 option contracts with a mean option price of 34.7010 and average implied volatility of 0.0793. Standard errors are obtained by numerically computing the Jacobian of mean squared error(MSE) function. We applied FRFT approach to price options which significantly reduces the calibration time. Discrete time Heston-Nandi (HN) closed form GARCH model requires longer time in calibration than any continuous time model and the requirement is even longer for our CFGNIG model.*

Days to maturity<40

	RMSE	AAE	MOE
0.95<=S/K<0.99			
BS	8.7221	6.4323	-5.6875
HS	8.7388	6.7399	-5.9302
HN(R)	8.2805	6.2582	-5.4049
HN(U)	7.7223	5.6501	-4.7166
CFG-NIG	2.5545	1.7976	-0.7966
0.99<=S/K<1.01			
BS	12.8775	11.1315	-9.4685
HS	12.1883	10.3358	-8.2060
HN(R)	11.5708	9.7200	-7.4205
HN(U)	10.7019	8.8163	-6.2401
CFG-NIG	5.3629	4.4576	-1.4436
1.01<=S/K<1.05			
BS	10.0511	8.5113	-6.5124
HS	10.3805	9.0247	-7.1032
HN(R)	9.8502	8.5010	-6.5445
HN(U)	9.0448	7.7015	-5.6951
CFG-NIG	5.0601	4.2035	-1.8791

40=<Days to maturity<70

0.95<=S/K<0.99			
BS	15.4207	14.0656	-12.7740
HS	15.3356	14.0274	-12.7358
HN(R)	14.9604	13.6466	-12.3551
HN(U)	14.3403	13.0507	-11.7591
CFG-NIG	2.8560	2.4703	-1.2939
0.99<=S/K<1.01			
BS	18.6359	17.5317	-16.0970
HS	18.0890	16.9511	-15.5163
HN(R)	17.6968	16.5706	-15.1359
HN(U)	16.6766	15.5474	-14.1127
CFG-NIG	4.7520	4.5271	-3.0955
1.01<=S/K<1.05			
BS	17.8482	17.1204	-15.6441
HS	17.5556	16.8205	-15.3442
HN(R)	17.3400	16.6159	-15.1396
HN(U)	16.6473	15.9286	-14.4522
CFG-NIG	5.4025	5.1531	-3.7234

70=<Days to maturity=<100

0.95<=S/K<0.99			
BS	19.6755	18.4333	-16.9416
HS	18.8199	17.5918	-16.1001
HN(R)	19.2289	18.0164	-16.5248
HN(U)	18.7449	17.5711	-16.0794
CFG-NIG	3.0536	2.5260	-1.2674
0.99<=S/K<1.01			
BS	20.1773	18.9623	-17.4236
HS	18.6527	17.3661	-15.8274
HN(R)	19.3701	18.1232	-16.5845
HN(U)	18.7286	17.4827	-15.9439
CFG-NIG	3.6980	3.5110	-1.9893
1.01<=S/K<1.05			
BS	18.2078	17.1253	-15.3467
HS	16.7919	15.6281	-13.8495
HN(R)	17.4633	16.3725	-14.5939
HN(U)	17.1466	16.1349	-14.3563
CFG-NIG	4.3736	3.8738	-2.3090

Figure 5.5: *Out-of-sample valuation errors for Call Options traded in second half of 2007. The models are calibrated on first half of the same year. Total number of contracts available for the second half is 1505. BS stands for Black-Scholes model, HS stands for Hestons'93 stochastic volatility model, HN(R) stands for restricted version of Heston and Nandis 2000 GARCH model, HN(U) stands for unrestricted version of Heston and Nandis 2000 GARCH model, CFG-NIG stands for closed form GARCH model with NIG innovations. RMSE is the root mean square error as defined in*(4.48), *AAE is the average absolute error as defined in*(5.175) *and MOE is the mean outside error as defined in*(5.176).

Model	RMSE	Parameters						
BS	2.3786	(σ)						
		0.0716						
		(0.0051)						
VG	2.3547	(σ)	(θ)	(ν)				
		0.0191	0.7082	0.0094				
		(0.0020)	(0.0044)	(0.0014)				
NIG	2.3551	(α)	(β)	(δ)				
		928.6679	825.1195	0.4566				
		(0.2445)	(0.2841)	(0.1729)				
JD-DE	2.3786	(σ)	(λ)	(p)	(η_1)	(η_2)		
		0.0716	3.7968	0.4730	1533.5	1533.5		
		(0.0052)	(0.2842)	(0.1913)	(0.2525)	(0.3417)		
CGMY	2.3685	(C)	(G)	(M)	(Y)			
		0.0272	94.9083	53.3706	1.3614			
		(0.0039)	(0.0355)	(0.0219)	(0.0257)			
HS	2.3405	(κ)	(θ)	(σ)	(ρ)	(V_0)		
		0.1870	0.0460	0.0379	0.9900	0.0042		
		(0.0775)	(0.0311)	(0.0744)	(0.1109)	(0.0009)		
HN(R)	2.3647	(α_1)	(β_1)	(β_0)	(γ)	(λ)		
		2.2204e-16	2.5670e-9	2.0504e-5	419.0868	-419.5868		
		(1.9020e-4)	(5.7663e-7)	(2.2987e-6)	(1.0363e-4)	(1.1255e-4)		
HN(U)	2.2776	(α_1)	(β_1)	(β_0)	(γ)	(λ)		
		2.2204e-16	2.5670e-9	2.3835e-5	419.9706	-1.7080		
		(0.0137)	(6.9763e-7)	(3.6808e-6)	(0.0070)	(0.0146)		
CFGNIG	1.7013	(α_1)	(β_1)	(β_0)	(γ)	(α)	(β)	(δ)
		0.9980	2.567e-15	2.5670e-15	-0.5126	296.3627	-0.1393	312.6878
		(0.0045)	(1.0223e-7)	(1.0219e-7)	(0.1251)	(0.0105)	(0.0105)	(0.0049)

Table 5.5: *Calibration with Options traded over the period January'2005-December'2006. We consider Options traded on every Wednesday. After all cleaning we have 5848 option contracts with a mean option price of 20.7565 and average implied volatility of 0.0774. Standard errors are obtained by numerically computing the Jacobian of mean squared error(MSE) function. We applied FRFT approach to price options which significantly reduces the calibration time. Discrete time Heston-Nandi (HN) closed form GARCH model requires longer time in calibration than any continuous time model and the requirement is even longer for our CFGNIG model.*

	RMSE	AAE	MOE
Days to maturity<40			
0.95<=S/K<0.99			
BS	2.2074	1.3461	-0.3481
HS	2.3409	1.7433	-0.2749
HN(R)	2.3316	1.7525	-0.0699
HN(U)	2.3430	1.7792	0.2012
CFG-NIG	1.7614	1.3704	0.5146
0.99<=S/K<1.01			
BS	4.9875	3.6377	-1.1268
HS	5.1212	3.9012	0.0874
HN(R)	5.0971	3.9383	0.3924
HN(U)	5.1114	3.9686	0.4809
CFG-NIG	3.7954	2.4969	1.2636
1.01<=S/K<1.05			
BS	4.1194	2.6545	-0.9448
HS	4.4929	2.9341	-1.5443
HN(R)	4.3820	2.8330	-1.3732
HN(U)	4.5581	3.0042	-1.5867
CFG-NIG	2.7516	1.8567	-0.6468
40=< Days to maturity<70			
0.95<=S/K<0.99			
BS	3.5136	2.5294	-0.5713
HS	3.4392	2.5483	-0.1206
HN(R)	3.4108	2.4899	-0.2603
HN(U)	3.3987	2.5578	0.1331
CFG-NIG	0.8670	0.5876	-0.0750
0.99<=S/K<1.01			
BS	5.1376	3.8793	-2.0494
HS	5.0129	3.7564	-1.8327
HN(R)	4.8939	3.6482	-1.6270
HN(U)	4.9238	3.6647	-1.6885
CFG-NIG	2.5060	1.9955	-1.0942
1.01<=S/K<1.05			
BS	4.5360	3.1752	-0.8124
HS	4.7061	3.2040	-1.3687
HN(R)	4.5830	3.1361	-1.0614
HN(U)	4.9044	3.3146	-1.8031
CFG-NIG	2.8950	1.9977	-1.0025
70=< Days to maturity=<100			
0.95<=S/K<0.99			
BS	6.0532	4.4878	-2.1437
HS	5.6243	4.2388	-1.0760
HN(R)	5.9557	4.4072	-1.8840
HN(U)	5.9631	4.4374	-1.6571
CFG-NIG	1.8078	1.3718	-0.5850
0.99<=S/K<1.01			
BS	7.8154	5.9639	-4.3866
HS	7.4242	5.6176	-3.9375
HN(R)	7.6899	5.8229	-4.2110
HN(U)	8.0528	6.1289	-4.6553
CFG-NIG	3.1106	2.6789	-1.5454
1.01<=S/K<1.05			
BS	2.8609	2.2658	-0.3348
HS	2.9403	2.3457	-0.6483
HN(R)	2.8860	2.2934	-0.4853
HN(U)	3.5061	2.8174	-1.6319
CFG-NIG	1.8051	1.5212	-0.3290

Figure 5.6: *Out-of-sample valuation errors for Call Options traded in first half of 2007. The models are calibrated using 2005 and 2006 contracts. Total number of contracts available for the second half is 1578. BS stands for Black-Scholes model, HS stands for Hestons'93 stochastic volatility model, HN(R) stands for restricted version of Heston and Nandis 2000 GARCH model, HN(U) stands for unrestricted version of Heston and Nandis 2000 GARCH model, CFG-NIG stands for closed form GARCH model with NIG innovations. RMSE is the root mean square error as defined in*(4.48), *AAE is the average absolute error as defined in*(5.175) *and MOE is the mean outside error as defined in*(5.176).

Model	RMSE	Parameters						
BS	5.6804	(σ)						
		0.0895						
		(0.0044)						
VG	5.6785	(σ)	(θ)	(ν)				
		0.0900	-0.0603	0.0283				
		(0.0045)	(0.2097)	(0.1193)				
NIG	5.6786	(α)	(β)	(δ)				
		70.5699	-8.0693	0.5663				
		(0.0852)	(0.2034)	(0.0553)				
JD-DE	5.6804	(σ)	(λ)	(p)	(η_1)	(η_2)		
		0.0892	0.0105	0.4473	13.7655	16.3763		
		(0.0044)	(0.1637)	(0.7376)	(0.8875)	(0.9346)		
CGMY	5.6796	(C)	(G)	(M)	(Y)			
		9.9238e-4	12.9536	90.5655	1.8597			
		(0.0001)	(0.0176)	(0.0662)	(0.0089)			
HS	5.6796	(κ)	(θ)	(σ)	(ρ)	(V_0)		
		8.8081e-4	0.0530	0.0097	-0.9900	0.0080		
		(0.1341)	(0.1491)	(0.0650)	(0.1491)	(0.0009)		
HN(R)	4.8583	(α_1)	(β_1)	(β_0)	(γ)	(λ)		
		0.9861	2.5670e-9	2.5670e-9	418.8949	-419.3949		
		(2.2635e-5)	(1.1691e-7)	(1.1557e-6)	(1.1373e-4)	(2.4801e-5)		
HN(U)	4.8582	(α_1)	(β_1)	(β_0)	(γ)	(λ)		
		0.9863	2.5670e-9	2.5670e-9	417.9965	-0.5434		
		(0.0044)	(1.8702e-8)	(1.3887e-7)	(0.0092)	(0.0039)		
CFGNIG	2.38	(α_1)	(β_1)	(β_0)	(γ)	(α)	(β)	(δ)
		0.99999	8.7545e-8	2.567e-15	-0.5402	36.0597	10.2048	26.8947
		(0.0031)	(1.1055e-7)	(1.1071e-7)	(0.0504)	(0.2314)	(0.0353)	(0.0479)

Table 5.6: *Calibration with Options traded over the period January'2006-December'2007. We consider Options traded on every Wednesday. After all cleaning we have 5848 option contracts with a mean option price of 20.7565 and average implied volatility of 0.0774. Standard errors are obtained by numerically computing the Jacobian of mean squared error(MSE) function. We applied FRFT approach to price options which significantly reduces the calibration time. Discrete time Heston-Nandi (HN) closed form GARCH model requires longer time in calibration than any continuous time model and the requirement is even longer for our CFGNIG model.*

Days to maturity<40

	RMSE	AAE	MOE
0.95<=S/K<0.99			
BS	10.3605	9.3289	-8.5312
HS	10.0263	9.1636	-8.3658
HN(R)	5.8013	4.8923	-2.4570
HN(U)	5.7983	4.8898	-2.4712
CFG-NIG	2.1664	1.9696	-1.2020
0.99<=S/K<1.01			
BS	14.8007	14.3488	-13.1958
HS	13.6761	13.0478	-11.8948
HN(R)	8.0183	7.0852	-5.0756
HN(U)	8.0362	7.1037	-5.1078
CFG-NIG	3.6698	3.5747	-2.4217
1.01<=S/K<1.05			
BS	11.7575	10.7682	-9.4563
HS	11.8149	10.9170	-9.6075
HN(R)	7.4184	6.4184	-3.9779
HN(U)	7.4499	6.4539	-4.0400
CFG-NIG	3.3378	3.0189	-1.7456

40=<Days to maturity<70

0.95<=S/K<0.99			
BS	18.3695	17.9737	-16.7178
HS	18.2109	17.8298	-16.5739
HN(R)	10.4007	9.4663	-6.9594
HN(U)	10.4028	9.4638	-6.9749
CFG-NIG	2.8379	2.6335	-1.5205
0.99<=S/K<1.01			
BS	21.1614	20.8998	-19.4607
HS	20.8318	20.5660	-19.1270
HN(R)	11.6669	10.6320	-7.7611
HN(U)	11.6940	10.6641	-7.8236
CFG-NIG	4.2510	4.1823	-2.7441
1.01<=S/K<1.05			
BS	19.2422	18.8390	-17.0146
HS	19.1059	18.7051	-16.8808
HN(R)	11.2020	10.3059	-7.1731
HN(U)	11.2735	10.3790	-7.2938
CFG-NIG	4.0219	3.8536	-2.0767

70=<Days to maturity=<100

0.95<=S/K<0.99			
BS	22.5842	22.2782	-21.0498
HS	22.5329	22.2344	-21.0060
HN(R)	15.0993	14.1702	-12.5554
HN(U)	15.0953	14.1672	-12.5578
CFG-NIG	1.7891	1.5078	-0.6658
0.99<=S/K<1.01			
BS	23.6267	23.4109	-22.2609
HS	23.3639	23.1433	-21.9933
HN(R)	16.1731	15.4553	-13.6298
HN(U)	16.2079	15.4896	-13.6857
CFG-NIG	3.0466	2.9400	-1.8370
1.01<=S/K<1.05			
BS	20.6092	20.2546	-18.9189
HS	20.3539	19.9962	-18.6605
HN(R)	13.9716	13.1916	-11.3228
HN(U)	14.0613	13.2821	-11.4446
CFG-NIG	2.9846	2.7656	-1.5236

Figure 5.7: *Out-of-sample valuation errors for Call Options traded in first half of 2008. The models are calibrated using 2006 and 2007 contracts. Total number of contracts available for the second half is 943. BS stands for Black-Scholes model, HS stands for Hestons'93 stochastic volatility model, HN(R) stands for restricted version of Heston and Nandis 2000 GARCH model, HN(U) stands for unrestricted version of Heston and Nandis 2000 GARCH model, CFG-NIG stands for closed form GARCH model with NIG innovations. RMSE is the root mean square error as defined in*(4.48), *AAE is the average absolute error as defined in*(5.175) *and MOE is the mean outside error as defined in*(5.176).

Model	RMSE	Parameters						
BS	5.0604	(σ)						
		0.0846						
		(0.0046)						
VG	5.0592	(σ)	(θ)	(ν)				
		0.0801	0.5579	0.0023				
		(0.0028)	(0.1206)	(0.0022)				
NIG	5.0592	(α)	(β)	(δ)				
		252.3847	74.8615	1.5681				
		(0.2445)	(0.2841)	(0.1729)				
JD-DE	5.0604	(σ)	(λ)	(p)	(η_1)	(η_2)		
		0.0845	0.1382	0.4519	155.8497	190.3719		
		(0.0049)	(0.0128)	(0.0094)	(0.0124)	(0.0141)		
CGMY	5.0592	(C)	(G)	(M)	(Y)			
		0.0240	94.9079	53.3726	1.4418			
		(0.0026)	(0.0264)	(0.0201)	(0.0185)			
HS	5.0590	(κ)	(θ)	(σ)	(ρ)	(V_0)		
		0.0015	0.0543	0.0128	0.9900	0.0071		
		(0.1244)	(0.1419)	(0.0683)	(0.1419)	(0.0009)		
HN(R)	4.9648	(α_1)	(β_1)	(β_0)	(γ)	(λ)		
		0.9348	2.5670e-9	1.3022e-6	355.1512	-355.6512		
		(5.9897e-5)	(2.2001e-7)	(1.1262e-6)	(5.9926e-5)	(1.3546e-4)		
HN(U)	4.9543	(α_1)	(β_1)	(β_0)	(γ)	(λ)		
		0.9494	2.5670e-9	1.0208e-6	419.0905	-1.0947		
		(0.0067)	(3.9161e-8)	(2.6777e-7)	(0.0139)	(0.0072)		
CFGNIG	2.2553	(α_1)	(β_1)	(β_0)	(γ)	(α)	(β)	(δ)
		0.99999	2.1836e-8	2.6181e-8	-0.5406	38.5655	7.5215	26.4156
		(0.0034)	(1.0925e-7)	(1.0933e-7)	(0.0518)	(0.0412)	(0.0375)	(0.0380)

Table 5.7: *Calibration with Options traded over the period January'2005-December'2007. We consider Options traded on every Wednesday. After all cleaning we have 8931 option contracts with a mean option price of 22.7247 and average implied volatility of 0.0884. Standard errors are obtained by numerically computing the Jacobian of mean squared error(MSE) function. We applied FRFT approach to price options which significantly reduces the calibration time. Discrete time Heston-Nandi (HN) closed form GARCH model requires longer time in calibration than any continuous time model and the requirement is even longer for our CFGNIG model.*

Days to maturity<40	RMSE	AAE	MOE
0.95<=S/K<0.99			
BS	10.8866	9.7935	-8.9957
HS	10.4743	9.5769	-8.7791
HN(R)	8.9946	8.0134	-7.2156
HN(U)	8.7844	7.8147	-7.0169
CFG-NIG	1.9176	1.6728	-0.9335
0.99<=S/K<1.01			
BS	15.5497	15.0953	-13.9423
HS	14.3446	13.6969	-12.5439
HN(R)	12.4475	11.6801	-10.5272
HN(U)	12.3749	11.6219	-10.4690
CFG-NIG	3.6057	3.5051	-2.3521
1.01<=S/K<1.05			
BS	12.1693	11.1507	-9.8386
HS	12.3755	11.4623	-10.1511
HN(R)	11.0409	10.0724	-8.7675
HN(U)	11.2772	10.3377	-9.0292
CFG-NIG	3.4930	3.1862	-1.9139
40=<Days to maturity<70			
0.95<=S/K<0.99			
BS	19.4016	19.0076	-17.7517
HS	19.0720	18.6809	-17.4250
HN(R)	17.6176	17.2845	-16.0286
HN(U)	17.1889	16.8431	-15.5872
CFG-NIG	2.7134	2.4753	-1.3712
0.99<=S/K<1.01			
BS	22.3122	22.0660	-20.6269
HS	22.0745	21.8246	-20.3856
HN(R)	20.6600	20.4276	-18.9885
HN(U)	20.5368	20.2986	-18.8595
CFG-NIG	4.3538	4.2914	-2.8523
1.01<=S/K<1.05			
BS	20.0306	19.6221	-17.7978
HS	20.2490	19.8550	-18.0307
HN(R)	19.3674	18.9695	-17.1451
HN(U)	19.8140	19.4397	-17.6153
CFG-NIG	4.3167	4.1733	-2.3606
70=<Days to maturity=<100			
0.95<=S/K<0.99			
BS	24.0277	23.7262	-22.4978
HS	23.8118	23.4998	-22.2714
HN(R)	23.9746	23.6832	-22.4548
HN(U)	23.7480	23.4377	-22.2093
CFG-NIG	1.8936	1.5875	-0.7416
0.99<=S/K<1.01			
BS	25.0446	24.8383	-23.6883
HS	25.0835	24.8767	-23.7267
HN(R)	25.0379	24.8427	-23.6927
HN(U)	25.2600	25.0650	-23.9150
CFG-NIG	3.3625	3.2757	-2.1603
1.01<=S/K<1.05			
BS	21.6040	21.2467	-19.9110
HS	21.9060	21.5565	-20.2208
HN(R)	21.7753	21.4382	-20.1025
HN(U)	22.6029	22.2926	-20.9569
CFG-NIG	3.4187	3.2422	-1.9301

Figure 5.8: *Out-of-sample valuation errors for Call Options traded in first half of 2008. The models are calibrated using options traded on 2005-2007. Total number of contracts available for the second half is 943. BS stands for Black-Scholes model, HS stands for Hestons'93 stochastic volatility model, HN(R) stands for restricted version of Heston and Nandis 2000 GARCH model, HN(U) stands for unrestricted version of Heston and Nandis 2000 GARCH model, CFG-NIG stands for closed form GARCH model with NIG innovations. RMSE is the root mean square error as defined in*(4.48), *AAE is the average absolute error as defined in*(5.175) *and MOE is the mean outside error as defined in*(5.176).

Part II

Risk Management

Chapter 6

Risk Measures: Extreme Value Versus Lévy

This chapter revisits the basics of risk management in financial context. It then investigates Lévy spectral risk measure as coherent alternative to Generalized Pareto spectral risk measure. In particular we consider implementation of expected shortfall(ES) and spectral risk measure(SRM) for conditional distributions belonging to Generalized Hyperbolic family of Lévy processes and compare their risk-management features with traditional unconditional extreme value(EV) approach. This reveals the relative performance of fitting the entire distributions and fitting only the tail; with their associated impact in risk-management. For frequently used risk measure VaR, backtesting performance of Lévy and EV approaches is investigated.

6.1 Introduction

Risk is a factor which plays an important role in our everyday dealings. Risk in economic and financial dealings needs to be modeled by financial institutions.By now risk modeling has become an integral part of most, if not all, financial institutions. The purpose of such modeling is not always to eliminate the risk but to have a good perception of and control on it. The quantitative idea about risk, used in such modeling, is about the probability that an investments actual return will be different from its expected return. Parallel to its definition as "subjective phenomena" involving exposure and uncertainty, the working idea about risk is captured in terms of changes in values, of some underlying, between two dates. Uncertainty and risk are very close in intuition , however, "uncertainty" can't be measured, in any form what soever, where as "risk" is a quantity which could possibly be measured. Since such a measure is closely related with the variability of the future value of position (or portfolio) due to the market changes or more generally due to uncertain events, the quantification of such a measure should naturally involve the future values only. Thus quantitatively the study of risk involves the study of random variables(underlying general stochastic process) on the set of the nature at a future date interpreted as possible future values of positions or portfolios currently held.

If we count the parameters of a particular model under consideration there could possibly be a good number of parameters involved in risk management. However there are always some generic parameters independent of the choice of the model. A fixed duration over which a model's riskiness is assessed is always a parameter. This parameter is usually referred as time horizon or holding period. Another parameter is the level of acceptance of risk. Roughly this acceptance level indicates beyond what level of downward returns risk-management might be a real concern. However the central tool to take care of in risk management is the financial random variable. This is the random variable used in describing the return process of the underlings. As mentioned earlier it is the random variable on the set of the nature at a future date interpreted as possible future values of positions or portfolios currently held.

6.1.1 Various Risk Measures

In this section we discuss the traditional risk measure VaR and its coherent versions ES and SRM. A comprehensive practical survey of these risk measures can be found in Dowd(2005)[42], Christoffersen(2003)[34]. Cotter and Dowd(2006)[39] studied these risk measures with an application to fixing clearing house margin requirement. Cotter and Dowd(2006)[39] consider extreme value(EV) model particularly suitable in this context. More recently Sorwar and Kevin(2010)[110] further studied these risk measures in option model framework under CEV dynamics. In this chapter we document that tail based risk measures perform well under tail based model and SRM performs well with models which are calibrated on the entire data. Let us first explain the intuitions behind these risk measures.

Value-at-Risk(VaR)

Under the consideration of static version of risk measure, VaR–for a given fixed time period and coverage–provides us an estimate of the magnitude of the expected potential loss. In plain words it provides us, over an specific time interval and for a given confidence level, the worst expected loss under normal market condition. So clearly VaR has three parameters, as mentioned earlier; relatively high level of confidence $(1-\alpha)$(typically 95% or 99%), the time period of projection T,(day, month, year) and the estimate of investment loss L. If X_0 and X_T denote the values of the investment at time 0 and T, respectively, the loss function is defined as follows.

$$L = X_0 e^{rT} - X_T \tag{6.1}$$

where "r" is the constant rate of interest. Then formally VaR can be defined as:

Definition 6.1 *VaR of any risky investment at the confidence level $1-\alpha$, for $\alpha \in (0,1)$, is given by the smallest number "D" such that the probability that the loss "L" exceeds "D" is not greater than α:*

$$\begin{aligned} VaR_T^{\alpha} &= \inf\{D \mid P(L > D) < \alpha\} \\ &= \inf\left\{D \mid P(X_0 e^{rT} - X_T > D) < \alpha\right\} \end{aligned} \tag{6.2}$$

So according to Christoffersen(2003)[34], VaR answers the question "what dollar loss is such that it will only be exceeded $\alpha \times 100\%$ of the time in next "T" trading days?"

Similarly, in case of portfolio, if we denote the portfolio return as R_{PF} then we can write: \$L= $-X_0 * R_{PF}$. Then equation (6.2) implies:

$$P(-X_0 * R_{PF} > VaR_T^{\alpha}) = \alpha \tag{6.3}$$

That is:

$$P\left(R_{PF} < \frac{VaR_T^{\alpha}}{-X_0}\right) = \alpha$$

Thus defining VaR with respect to the current value of the portfolio, X_0, we write: $\widehat{VaR}_T^{\alpha} = \frac{VaR_T^{\alpha}}{X_0}$. Then from equation (6.3) we obtain:

$$P\left(R_{PF} < -\widehat{VaR}_T^{\alpha}\right) = \alpha. \tag{6.4}$$

As in [34], writing VaR relative to the current value of the portfolio makes it much easier to think about. e.g. knowing that the VaR_T^{α} is equal to \$50000 doesn't mean much unless we know the current value of the portfolio. However knowing that $\widehat{VaR}_T^{\alpha}$ is 50% of the value of the portfolio conveys much more information.

Let $\widehat{VaR}_T^{\alpha}$ denote the 1%($\alpha = 0.01$) VaR for the one day ahead return ($T = 1$). If returns are normally distributed with mean zero and standard deviation $\sigma_{PF,T}$ then it can be shown, see e.g. Dowd(2005)[42], Christofferse(2003)[34], that:

$$\begin{aligned} \widehat{VaR}_T^{\alpha} &= -\sigma_{PF,T} \times \Phi_{\alpha}^{-1} \\ &= -\sigma_{PF,T} \times (-2.33) \end{aligned} \tag{6.5}$$

So the only information we need to obtain the VaR is tomorrow's($T = 1$) variance forecast. As Φ_{α}^{-1} is always negative for $\alpha < 0.5$, the minus sign in front of the VaR formula ensures that the VaR itself is a positive number. Thus in precise terms considering one day ahead uncertainty($T = 1$), the one day VaR with 99% coverage gives us a number $\widehat{VaR}_T^{\alpha}$ such that there is 1% chance of losing more than $[\sigma_{PF,T} \times (-2.33)] \times 100\%$ of the today's portfolio value. However this simplicity is because of the assumption of normality in return and for non-normal models things are not so straightforward.In this chapter we will deal with the difficulties in implementation of VaR and other risk measures for Lévy models. In next chapter we will show how an engineering tool FRFT can help us overcome those difficulties. This will pave the path to a practical implementation of the risk measures expected shortfall(ES) and spectral risk measures(SRM) for the Lévy models; which we believe is a contribution to the literature of risk management.

However, VaR, has serious drawbacks. One of the remarkable drawbacks is that it ignores the extreme losses. According to equation (6.4), it only tells us that 1% of the time we will get a return below the reported VaR number $\widehat{VaR}_T^{\alpha}$, but it says nothing about what will happen in those 1% worst cases. Thus in theoretical terms, among others two key desirable characteristics of a risk measure, namely coherence and subadditivity, are not satisfied by VaR. VaR fixes tail events corresponding to a given confidence level but leaves the tail completely unattended. Knowing an amount of possible loss to the occurrence of extreme event is important but what is more important is to have the idea of how catastrophic the loss could be once such an event occurs. Moreover VaR assumes that the portfolio is constant in next "T" trading days which is unrealistic in many cases when "T" is larger than a day or a week. Finally it may not be clear how one should choose "T" and "α".

Despite all its limitations the tail based risk measure VaR has seen a great amount of applications in many stochastic environments where risk management makes a difference and undoubtedly becomes the industry benchmark for risk calculation. This is because it captures the important aspect of risk namely how bad things can get with a certain probability,α. Also it can be easily communicated and understood.

We close the discussion on VaR with an obvious result which further reinforces the underlying intuition.

Proposition 6.1 *VaR of a risk free asset is zero.*

Proof. We know in case of risk free asset, say bond, we have:

$$B_t = B_0 e^{rt} \qquad t \in [0, T].$$

Since source of randomness "z" has no role to play in case of risk free asset, the loss function "L", as defined in (6.1), turns out to be:

$$L(z) = B_0 e^{rT} - B_T(z) = 0.$$

Therefore without loss of generality we can assume that the constant random variable "L(z)" can be described by a distribution of the form:

$$P(L(z) \leq l) = \begin{cases} 1 & \text{if } \ l > 0 \\ 0 & \text{otherwise} \end{cases} \tag{6.6}$$

Now applying to the definition of VaR, we obtain:

$$
\begin{aligned}
VaR_T^{\alpha} &= \inf\{D \mid P(L > D) < \alpha\} \\
&= \inf\{D \mid P(L \leq D) > 1-\alpha\} \\
&= \inf\left\{\{D \leq 0 \mid P(L \leq D) > 1-\alpha\} \bigcup \{D > 0 \mid P(L \leq D) > 1-\alpha\}\right\} \\
&= \inf\left\{\{D \leq 0 \mid 0 > 1-\alpha\} \bigcup \{D > 0 \mid 1 > 1-\alpha\}\right\} \qquad [using (6.6)] \\
&= \inf\left\{(0,\infty) \bigcup \emptyset\right\} \\
&= \inf(0,\infty) \\
&= 0.
\end{aligned}
$$

The proof is complete. □

Expected Shortfall(ES)

Mathematically it can be argued that for quantile based estimate it is possible to have similar 1% VaR for two portfolios having completely different 0.1% or 0.01% VaR. See Christoffersen(2003)[34]. That is VaR estimate with 1% coverage rate completely fails to reveal the fact that the tail shapes of the distribution may be completely different corresponding to portfolios with different risk exposures. This translates into the great limitation of VaR namely it concerns only with number of losses exceeding the VaR but not the magnitude of those losses. However the magnitude should be of serious concern to risk-manager as large VaR exceedence are much more likely to cause financial distress, such as bankruptcy , than those of small exceedence. Thus a risk measure accounting for both frequency and magnitude of large losses is very much expected. This is exactly what Expected shortfall does. Once modelled correctly the tail of the portfolio return distribution bears significant information to risk managers about the future losses. But with VaR, getting the idea of the shape of the entire tail of the return distribution is equivalent to computing it for various coverage levels, which is certainly less effective as a reporting tool. Expected shortfall bears this significance as a convenient reporting tool. It has the formal mathematical expression as:

$$ES_{t+1}^{p} = -\mathbb{E}_t[R_{t+1} \mid R_{t+1} < -VaR_{t+1}^{p}] \tag{6.7}$$

The negative sign in front of the VaR and expectation signifies that both VaR and expected shortfall are defined as positive numbers. Hence the underlying intuition is that expected shortfall represents the expected value of those future return's which are worse than VaR. The tail losses and its distribution can be thought of as a two dimensional object which gives us the information about the range of possible losses along x-axis and probability associated with each outcome along the y-axis. The measure expected shortfall aggregates these two dimensions into a single number by computing the average of the tail outcomes weighted by their probabilities.Thus when VaR gives us the loss such that only 1% of the extreme losses will be worse than it, ES gives us the expected value of those extreme losses exceeding the VaR. Thus the up-shoot is that though ES is not providing complete information about the shape of the tail, the shape beyond the VaR measure, however, is now being accounted in quantifying the risk. Expected shortfall, coherent version of VaR, is developed to provide the investors the idea of how severe the loss could be, on an average, once extreme event occurs.

Corollary 6.1 *The ES of a risk less investment is zero.*

The proof is obvious from the definition of ES and the Proposition6.1.

Spectral Risk Measure(SRM)

Expected Shortfall assigns equal weight to the losses in excess of VaR, which doesn't reflect investors relative risk appetite. Instead one can define more general risk measures M_ϕ that are weighted averages of quantiles of the loss distribution:

$$M_\phi = \int_0^1 \phi(p) VaR(p) dp \tag{6.8}$$

Here ϕ is a general weighting function which, in its general forms, assigns different weights to different quantiles reflecting investors appetite for underlying risk. The VaR and ES are the special cases of SRM. For the SRM to correspond to ES the weighting function ϕ is required to have the following form:

$$\phi(p) = \begin{cases} 0 & p < \alpha \\ \frac{1}{1-\alpha} & p \geq \alpha \end{cases}$$

VaR being a single quantile corresponds to the SRM when ϕ is Dirac delta function putting all the mass to the particular event $\{p = \alpha\}$ and zero mass to all other events $\{p \neq \alpha\}$. So by it's very definition VaR ignores the quantiles in the tail as it assigns zero mass to all the quantiles other than VaR itself where as ES assigns equal mass of $\frac{1}{1-\alpha}$ to all quantiles in the tail specified by the VaR.

Roughly speaking a coherent risk measure basically ensures that higher losses are assigned the weights which are at least not less than any weight assigned to lower losses. To ensure this the weighting function ϕ is required to satisfy the following conditions:

- Non-negativity: $\phi(p) \geq 0, \quad \forall p \in [0, 1]$
- Normalization: $\int_0^1 \phi(p)dp = 1$
- Monotonicity: for any two $p_1, p_2 \in [0, 1]$, with $p_1 \geq p_2$, $\phi(p_1) \geq \phi(p_2)$

See e.g. Acerbi(2004)[1]. Clearly the monotonicity reflects the investors risk averse attitude. However if there is no risk, we have nothing to worry about:

Corollary 6.2 *For any risk less investment the SRM is identically zero for any choice of weighting function.*

The proof follows from the definition of SRM and the Proposition6.1.

Kevin, Cotter and Sorwar(2008)[44] investigates spectral risk measure with respect to exponential risk aversion function and suggest using exponential risk aversion function in modelling financial risk. It provides investors the flexibility to choose their individual degree of aversions to risk, in contrast to obtaining the estimates of other risk measures corresponding to a given coverage level. This generalization comes with a parameter and a function of it, known as "risk aversion function". Off course the computational hassle also increases and complexity discourages to seek a closed form formula. Even the numerical schemes to evaluate the associated integral exhibits different degrees of perfection. Cotter and Dowd(2006)[39] investigate this risk measure for extreme Value(EV) model and compare its estimation performance with those of other risk measures such as VaR and ES. This chapter focuses on estimating SRM for Lévy models and addresses the related subjectivity's of implementation.

Purely tail based EV model is frequently used to model extreme events in many applications such as weather extremes, reserve extreme, financial extreme etc. However recently conditional models of Lévy type are also in extensive use. While EV is purely tail based, Lévy models utilize the entire data to estimate the parameters of the models. These Lévy models are also developed with a view towards improved tail modelling. We investigate the relative performance of tail modelling by these two categories of models and contrast one another on the basis of their tail based risk measures. Our approach is to fix the tail as it is used in EV calibration and then consider the similar tails obtained from Lévy models with calibration based on entire data. This reveals the fact whether considering the observations discarded by EV does make any difference in performance, in other words we are interested in investigating whether extreme observations alone suffice modelling the extreme behavior or information does carry from the discarded observations as well.

Lévy approach mathematically appeals more than EV approach and provides plethora of alternatives to try for. However, Lévy models have the obvious limitation of non-availability of closed form formulas for risk measures. As such even the relatively straightforward VaR looks cumbersome to implement. Naturally as other risk measures such as ES and SRM are some sort of compounded versions of VaR, their implementations become even more cumbersome. We enrich the literature by considering the implementations of ES and SRM for Lévy models. Though requirement of huge computational times renders the use of these risk measures less appealing there is, however, positive side as well. Furthermore computational hassles so far discourages researchers to backtest Lévy based VaR models. This paper, probably for the first time in literature, reveals the backtesting performance of Lévy based VaR models. We contrast the backtesting performance of Lévy and EV models. In doing so we consider a rolling window of long four years to calibrate the models and obtain VaR's in dynamic fashion. The long length of window is considered basically to obtain enough observations on tails corresponding to EV model. Even so we need to consider 30% of the observations to fix tail, in calibrating EV model. This is to ensure that in long eight year period of backtesting we are not having convergence problem in estimation even for a day.

6.2 Characterization and estimation in Lévy Framework

The general expression of characteristic function of a stochastically continuous process starting at zero and having stationary independent increments is known in the literature as Lévy-Kintchine formula.

Equation (2.23) is the celebrated Lévy–Khinchine representation of a Lévy process. The theory of Lévy process is well established (see Bertoin(1996)[15], Sato(1999)[100], or Kyprianou(2006)[76]) and recently has seen great applications both in finance and insurance.

Given the transition density of a process on $[t_1, t_2]$,say, the characteristic function (2.23) of the conditional distribution of the process at t_2, given the information available up to t_1, can be obtained by so called Fourier transform. However the transition density itself suffices the estimation of different risk measures. Availability of closed form transition density assumes the underlying process is closed under convolution. When processes are not closed under convolution and only density at $t = 1$ is available then infinite divisibility of Lévy processes can be used to obtain the conditional characteristics of the process on an interval of length t:

$$\Phi_{X_t}(s) = [\Phi_{X_1}(s)]^t. \tag{6.9}$$

Finally when even density at $t = 1$ is not available, inverse Fourier transform can be used to numerically obtain the transition density from the characteristic function (2.23) with a given Lévy measure of a process, which is always available. Then these numerical transition densities can be utilized to estimate the risk measures under different model assumptions corresponding to different Lévy measures.

In this chapter our interest is limited to those members of Generalized Hyperbolic(GH) family of Lévy processes which are extensively studied in recent time for financial modeling. See e.g. Fusai and Meucci(2008)[60], Schoutens and Cariboni(2009)[101], Fusai and Roncoroni(2008)[59] and the numerous references therein. The original introduction of GH family of Lévy processes took place in modeling grain-size distribution of wind-blown sands, see Barndorff(1977)[12]. Later Eberlein and Prause(1998)[50] and Prause(1999)[91] studied the whole family of GH distributions as a tool to model log-returns of financial assets. Some of its subclasses were separately studied in financial context. Eberlien and

Keller(1995)[52], Bingham and Kiesel(2001)[18] studied the Hyperbolic distribution and Barndorff(1995)[12] studied NIG for the first time in literature to model financial data. Seneta(2004) [103] separately studied VG subclass to model financial return. Eberlien and Hammerstein(2002)[51] provide a complete and useful overview of limiting cases for this rich family of processes. Our focus is concentrated on exploring this family of Lévy processes for recently introduced spectral risk measures(SRM) in financial context, see Cotter and Dowd(2006)[39], Kevin, Cotter and Sorwar(2008)[44]. Recently SRM is further explored, see Sorwar and Dowd(2010)[110]. Confinement to the subclasses Variance Gamma(VG), Normal Inverse Gaussian(NIG), Hyperbolic(HYP) and Generalized Hyperbolic(GH), itself, allows us to obtain either the transition densities across time for processes closed under convolution or at least the densities at time $t = 1$ for those which are not closed under convolution. Furthermore in our empirical study we will be using daily return data for the indices under consideration and will be keeping the time scale in *days* so that $t = 1$ in equation(6.9) ensures that we are not required to use any inversion to obtain the transition densities numerically even when the process is not closed under convolution.

For us $X_1 = log\left(\frac{S_{t+1}}{S_t}\right)$, for any non-negative integer t and is characterized by Lévy Kintchine Formula (2.23). For the models we consider, the equivalent processes are given, more effectively, by their densities.

For VG:

$$f^{vg}_{X_1}(x) = \frac{2e^{\frac{\theta x}{\sigma^2}}}{\sigma\sqrt{2\pi}\nu^{\frac{1}{\nu}}}\left(\frac{x}{\sqrt{\frac{2\sigma^2}{\nu}+\theta^2}}\right)^{\frac{1}{\nu}-\frac{1}{2}} K_{\frac{1}{\nu}-\frac{1}{2}}\left(\frac{x\sqrt{\frac{2\sigma^2}{\nu}+\theta^2}}{\sigma^2}\right) \tag{6.10}$$

with $\sigma > 0$, $\theta \in \mathbb{R}$ and $\nu > 0$. Here $K_I()$ is the modified Bessel function of third kind with index I. From the expressions of higher moments it become apparent that θ controls the skewness in this model. In symmetric case σ solely characterizes the volatility where as in case of asymmetry the volatility characterization involves other parameters as well. See Schoutens(2003)[102].

For NIG:

$$f^{nig}_{X_1}(x) = \frac{\alpha\delta}{\pi}e^{\left(\delta\sqrt{\alpha^2-\beta^2}+\beta x\right)}\frac{K_1\left(\alpha\sqrt{\delta^2+x^2}\right)}{\sqrt{\delta^2+x^2}} \tag{6.11}$$

with $\alpha > 0$, $\mid \beta \mid < \alpha$ and $\delta > 0$. Expressions of higher moments show that β characterizes the skewness in this model and $\beta = 0$ corresponds to the symmetric case. Restrictions on other parameters are obvious from the expression of density function of the model and property of Bessel function.

For HYP:

$$f_{X_1}^{hyp}(x) = \frac{\sqrt{\alpha^2 - \beta^2}}{2\delta\alpha K_1\left(\delta\sqrt{\alpha^2 - \beta^2}\right)} e^{\left(-\alpha\sqrt{\delta^2+x^2}+\beta x\right)} \tag{6.12}$$

with $\alpha > 0$, $\mid \beta \mid < \alpha$ and $\delta > 0$. Here the parameter restrictions are obvious. However in this case it is not obvious how and which parameter('s) characterize the higher moments.

Finally for GH:

$$f_{X_1}^{gh}(x) = \frac{(\alpha^2 - \beta^2)^{\frac{\nu}{2}}}{\sqrt{2\pi}\alpha^{(\nu-\frac{1}{2})}\delta^\nu K_\nu\left(\delta\sqrt{\alpha^2 - \beta^2}\right)} \left(\delta^2 + x^2\right)^{\frac{\nu-\frac{1}{2}}{2}} K_{\nu-\frac{1}{2}}\left(\alpha\sqrt{\beta^2 + x^2}\right) e^{\beta x} \tag{6.13}$$

where

$$\alpha \geq 0, \quad \mid \beta \mid < \alpha \text{ if } \nu > 0$$

$$\alpha > 0, \quad \mid \beta \mid < \alpha \text{ if } \nu = 0$$

$$\alpha > 0, \quad \mid \beta \mid \leq \alpha \text{ if } \nu < 0.$$

Here the parameter restrictions are a bit complicated and inequalities are basically required for the modified Bessel function to be well defined. However like Hyperbolic model it is not obvious how and which parameters characterize the higher moments.

The availability of closed form densities make it relatively easier to obtain the standard errors of each parameter through Fishers information matix.

The benchmark for this paper is the extreme value model which only considers extreme returns in calibration. Consequently only the extreme returns characterize the performance of risk measures in this model. As explained in Dowd(2005)[42] ,Embrechts(1997)[53], and subsequently used in Cotter and Dowd(2006)[39], perhaps the most elegant approach to such purposes is to utilize the peaks-over-threshold(POT). The essence of POT approach lies in the fact that as the threshold u gets larger the distribution of exceedances converge to a two parameter Generalized Patero(GP) distribution:

Index	Position	u	Prob	N_u	β	ξ
S&P500	long	2	0.040	130	0.604(0.079)	0.182(0.099)
	short	2	0.035	118	0.759(0.130)	0.127(0.146)
FTSE100	long	1.5	0.077	250	0.707(0.074)	0.097(0.084)
	short	1.5	0.085	276	0.727(0.065)	0.022(0.067)
DAX	long	2	0.072	235	1.190(0.099)	0.012(0.052)
	short	2	0.072	237	1.000(0.097)	0.048(0.072)
Hang Seng	long	2	0.111	353	1.184(0.096)	0.127(0.062)
	short	2	0.116	367	1.148(0.086)	0.143(0.055)
Nikkei225	long	2	0.088	277	0.891(0.074)	-0.012(0.058)
	short	2	0.081	255	1.045(0.085)	-0.068(0.052)

Table 6.1: *Maximum likelihood estimation for EV model using futures indexes. Estimated standard error of each parameter appears in bracket.*

$$GP_{\xi,\beta}(x) = \begin{cases} 1-(1+\frac{\xi x}{\beta})^{-\frac{1}{\xi}} & \text{if } \xi \geq 0 \\ 1-\exp\left(\frac{-x}{\beta}\right) & \text{if } \xi < 0 \end{cases} \tag{6.14}$$

Here:

$$x \in \begin{cases} [0,\infty) & \text{if } \xi \geq 0 \\ \left[0, \frac{-\beta}{\xi}\right] & \text{if } \xi < 0 \end{cases} \tag{6.15}$$

The parameters ξ and $\beta > 0$ are respectively shape and scale parameters, contingent upon the threshold u.

Table6.1 shows the calibration results for EV model.Tables6.2-6.5 show the calibration results for various Lévy models of our consideration.

The very first observation is that when for EV, tail based calibration provides significantly different estimates for long and short positions(obviously for having different number of tail observations corresponding to long and short positions), that is not the case for Lévy based calibration on entire data sets. In this case long and short positions just alter the sign of the parameter characterizing the skewness of the model. This point could be explained further. For Lévy models with calibration on complete data 'short' and 'long' positions just causes the densities to be reflected along y-axis. So for a particular model, e.g.'VG'

Index	Position	u	Prob	σ	θ	ν
S&P500	long	2	0.034	1.094(0.020)	-0.037(0.019)	0.834(0.055)
	short	2	0.040	1.094(0.020)	0.037(0.019)	0.834(0.055)
FTSE100	long	1.5	0.084	1.164(0.019)	-0.022(0.020)	0.474(0.048)
	short	1.5	0.090	1.164(0.019)	0.022(0.020)	0.474(0.048)
DAX	long	2	0.076	1.505(0.026)	-0.032(0.026)	0.736(0.053)
	short	2	0.082	1.505(0.026)	0.032(0.026)	0.736(0.053)
Hang Seng	long	2	0.113	1.905(0.034)	-0.046(0.034)	0.808(0.053)
	short	2	0.122	1.905(0.034)	0.046(0.034)	0.808(0.053)
Nikkei225	long	2	0.088	1.529(0.024)	0.026(0.027)	0.398(0.048)
	short	2	0.083	1.529(0.024)	-0.026(0.027)	0.398(0.048)

Table 6.2: *Maximum likelihood estimation for VG model using futures indexes. Estimated standard error of each parameter appears in bracket.*

Index	Position	u	Prob	α	β	δ
S&P500	long	2	0.033	0.744(0.054)	-0.030(0.016)	0.926(0.044)
	short	2	0.039	0.744(0.054)	0.030(0.016)	0.926(0.044)
FTSE100	long	1.5	0.081	1.039(0.079)	-0.016(0.015)	1.427(0.087)
	short	1.5	0.088	1.039(0.079)	0.016(0.015)	1.427(0.087)
DAX	long	2	0.072	0.587(0.042)	-0.014(0.011)	1.374(0.068)
	short	2	0.077	0.587(0.042)	0.014(0.011)	1.374(0.068)
Hang Seng	long	2	0.106	0.428(0.031)	-0.012(0.009)	1.622(0.077)
	short	2	0.116	0.428(0.031)	0.012(0.009)	1.622(0.077)
Nikkei225	long	2	0.086	0.907(0.079)	0.011(0.012)	2.136(0.157)
	short	2	0.081	0.907(0.079)	-0.011(0.012)	2.136(0.157)

Table 6.3: *Maximum likelihood estimation for NIG model using futures indexes. Estimated standard error of each parameter appears in bracket.*

Index	Position	u	Prob	α	β	δ
S&P500	long	2	0.034	1.338(0.035)	-0.031(0.016)	0.238(0.063)
	short	2	0.040	1.338(0.035)	0.031(0.016)	0.238(0.063)
FTSE100	long	1.5	0.083	1.475(0.063)	-0.016(0.015)	0.874(0.116)
	short	1.5	0.089	1.475(0.063)	0.016(0.015)	0.874(0.116)
DAX	long	2	0.075	0.999(0.029)	-0.014(0.011)	0.488(0.095)
	short	2	0.081	0.999(0.029)	0.014(0.011)	0.488(0.095)
Hang Seng	long	2	0.111	0.773(0.021)	-0.013(0.009)	0.465(0.108)
	short	2	0.121	0.773(0.021)	0.013(0.009)	0.465(0.108)
Nikkei225	long	2	0.087	1.189(0.063)	0.011(0.012)	1.406(0.197)
	short	2	0.082	1.189(0.063)	-0.011(0.012)	1.406(0.197)

Table 6.4: *Maximum likelihood estimation for Hyperbolic model using futures indexes. Estimated standard error of each parameter appears in bracket.*

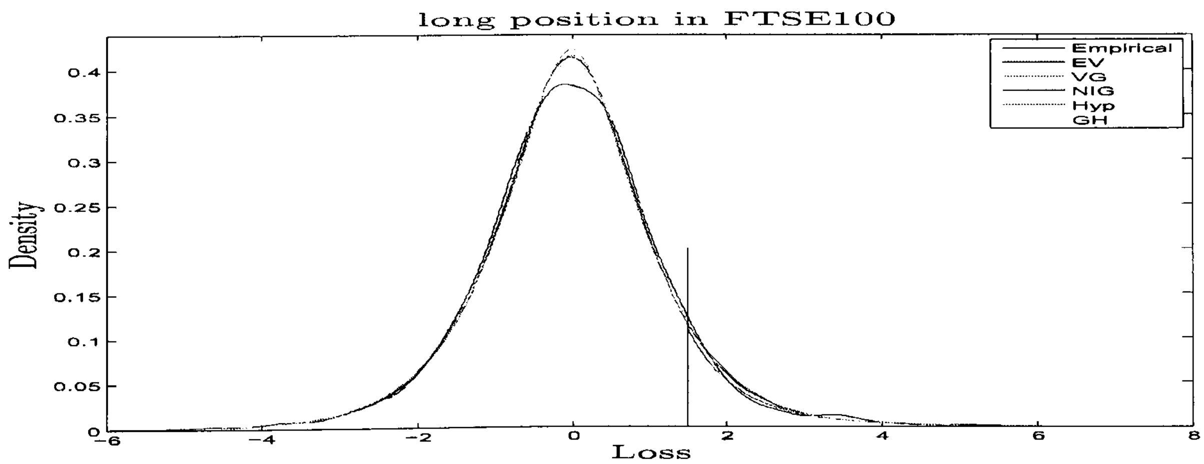


Figure 6.1: *Tail fit(EV) and total fit(Lévy) for long position on FTSE100. The threshold for EV model(red) is 1.5.*

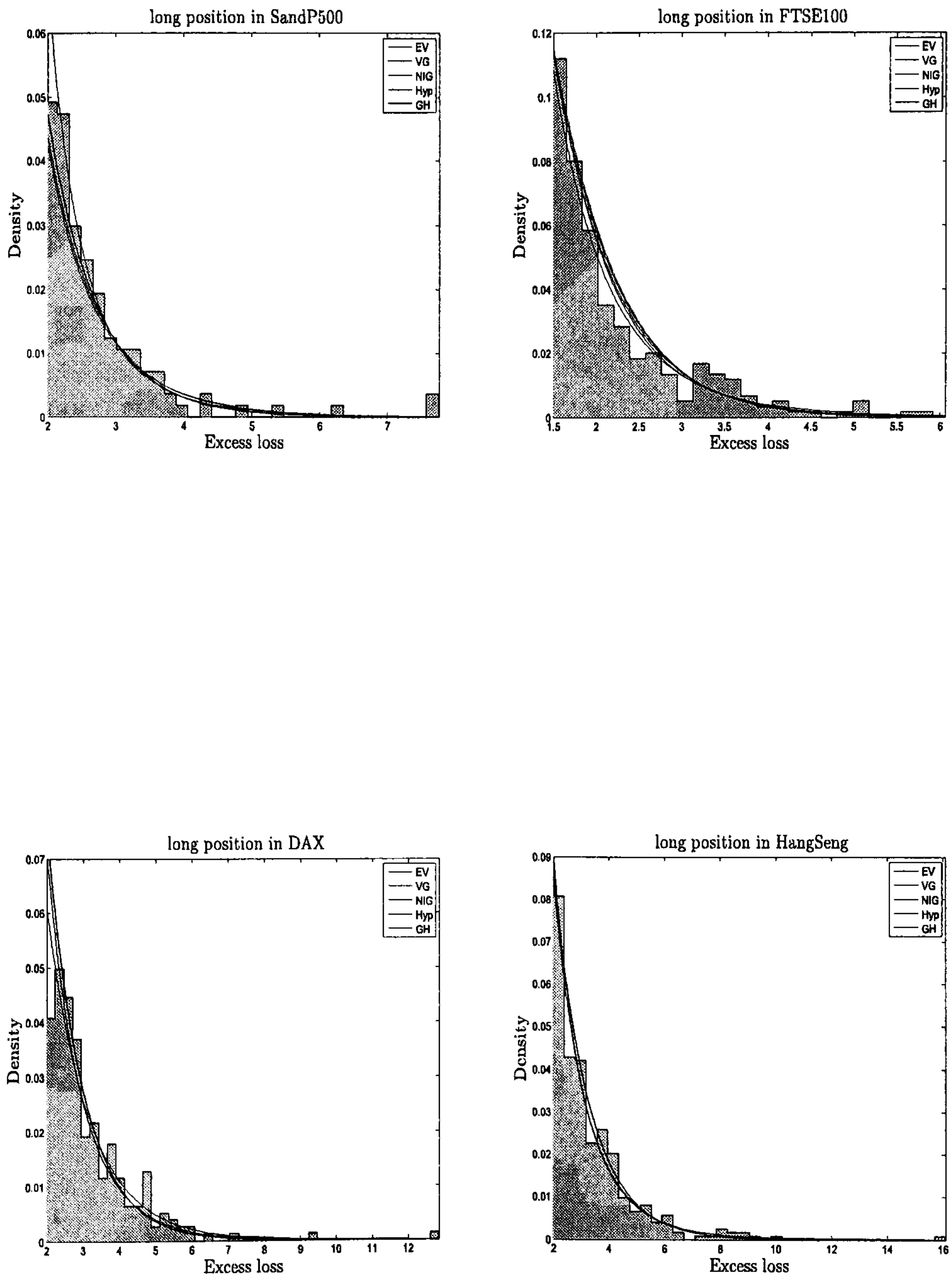


Figure 6.2: *Long positions in S&P500, FTSE100, DAX, and HangSeng. The thresholds are 2(S&P500),1.5(FTSE100),2(DAX) and 2(HangSeng) .*

Index	Position	u	Prob	α	β	δ	μ
S&P500	long	2	0.034	0.937(0.182)	-0.030(0.016)	0.712(0.204)	0.005(0.460)
	short	2	0.040	0.937(0.181)	0.030(0.016)	0.712(0.204)	0.005(0.460)
FTSE100	long	1.5	0.080	0.539(0.362)	-0.016(0.015)	1.920(0.300)	-1.956(0.848)
	short	1.5	0.086	0.539(0.362)	0.016(0.015)	1.920(0.300)	-1.956(0.848)
DAX	long	2	0.072	0.629(0.155)	-0.014(0.011)	1.289(0.313)	-0.345(0.553)
	short	2	0.078	0.629(0.155)	0.014(0.011)	1.289(0.313)	-0.345(0.553)
Hang Seng	long	2	0.106	0.347(0.132)	-0.012(0.009)	1.872(0.387)	-0.854(0.537)
	short	2	0.115	0.347(0.132)	0.012(0.009)	1.872(0.387)	-0.854(0.537)
Nikkei225	long	2	0.087	1.031(0.317)	0.011(0.012)	1.834(0.806)	0.152(1.648)
	short	2	0.081	1.031(0.349)	-0.011(0.012)	1.834(0.888)	0.152(1.824)

Table 6.5: *Maximum likelihood estimation for GH model using futures indexes. Estimated standard error of each parameter appears in bracket.*

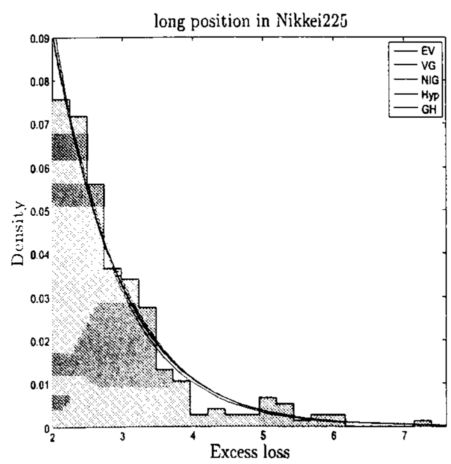


Figure 6.3: *Long position in Nikkei225. The threshold is 2(Nikkei225).*

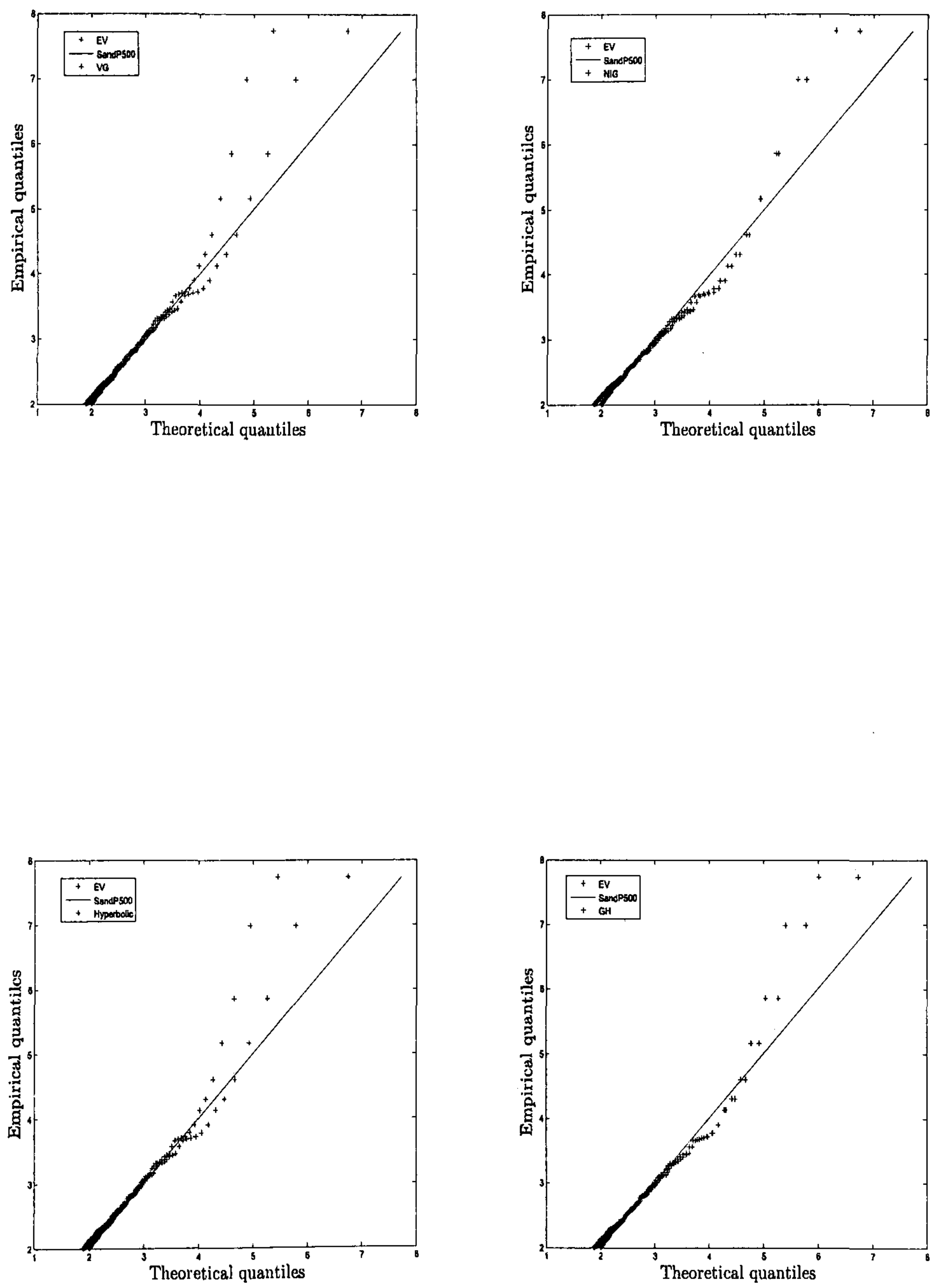


Figure 6.4: *EV and Lévy quantiles in excess of threshold(2): long position in S&P500.*

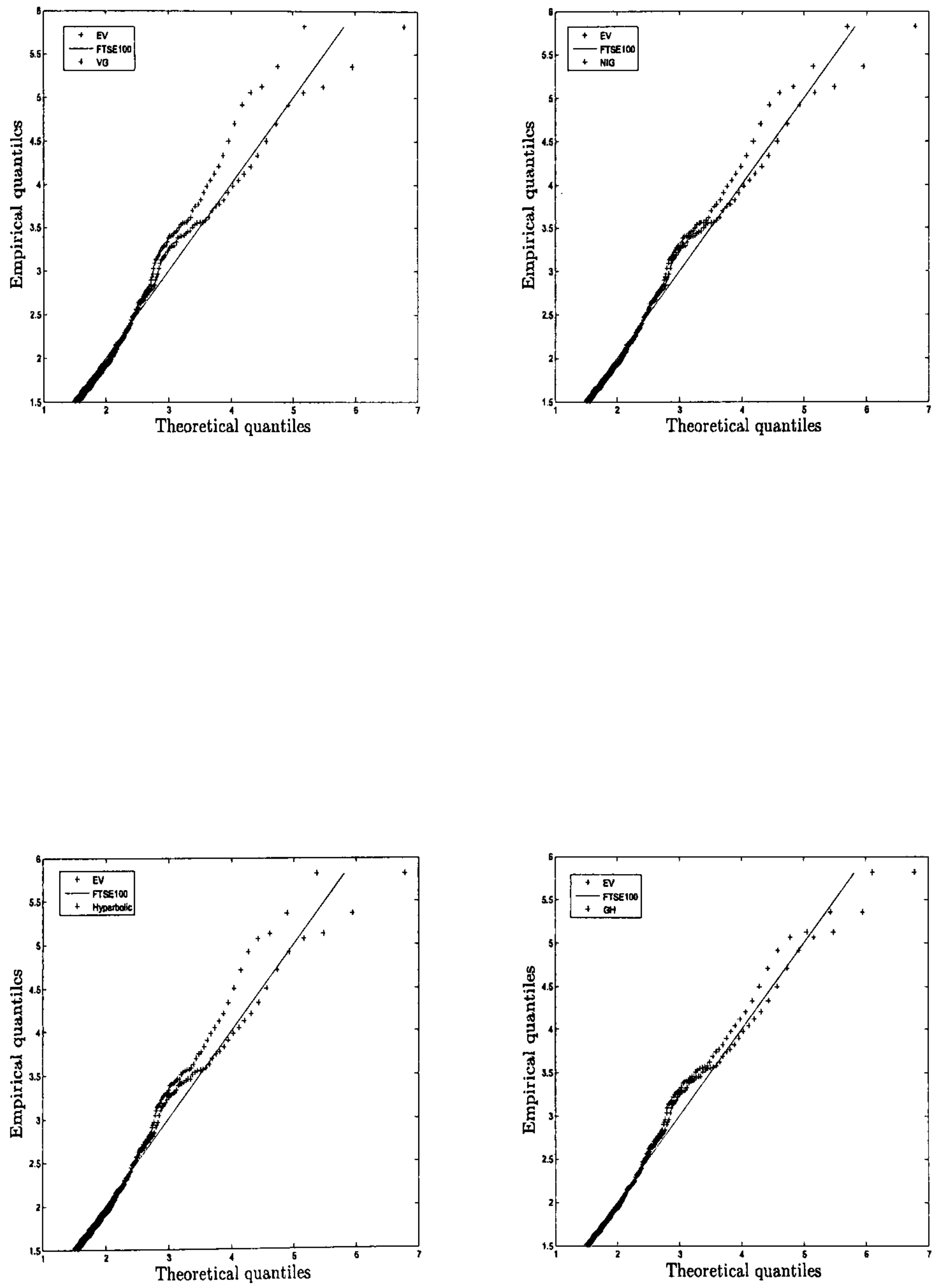


Figure 6.5: *EV and Lévy quantiles in excess of threshold(1.5): long position in FTSE100.*

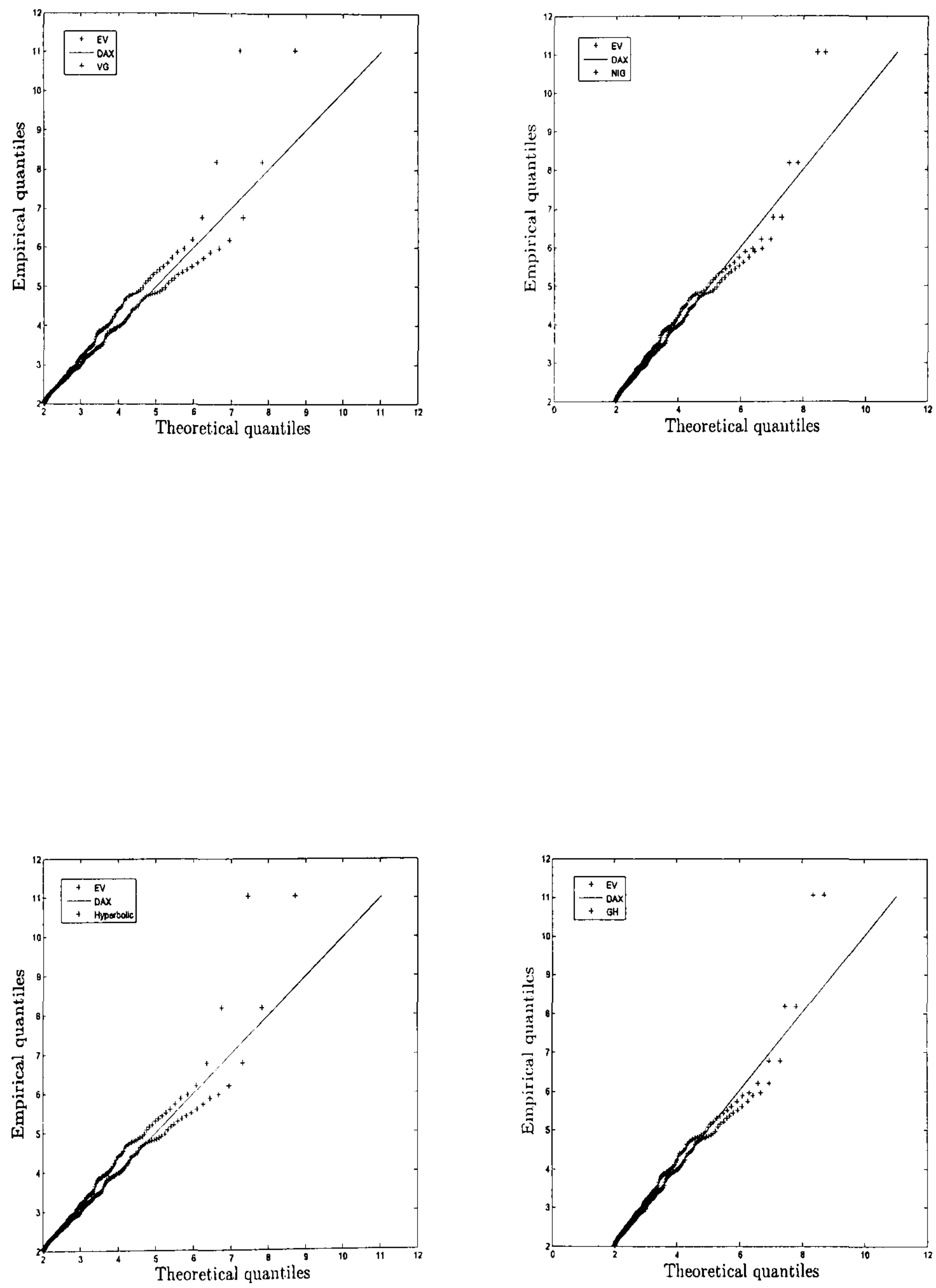


Figure 6.6: *EV and Lévy quantiles in excess of threshold(2): long position in DAX.*

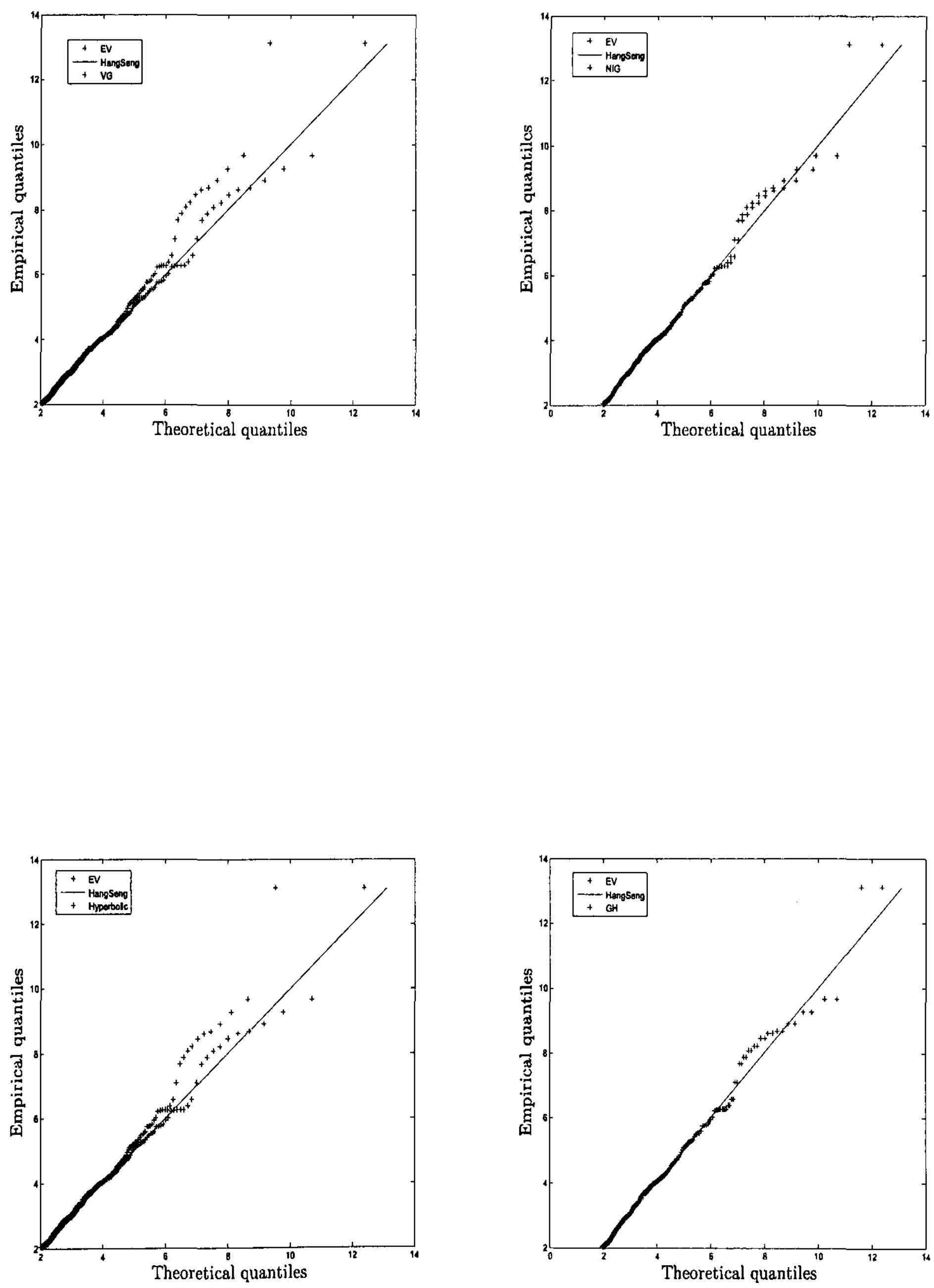


Figure 6.7: *EV and Lévy quantiles in excess of threshold(2): long position in Hang Seng.*

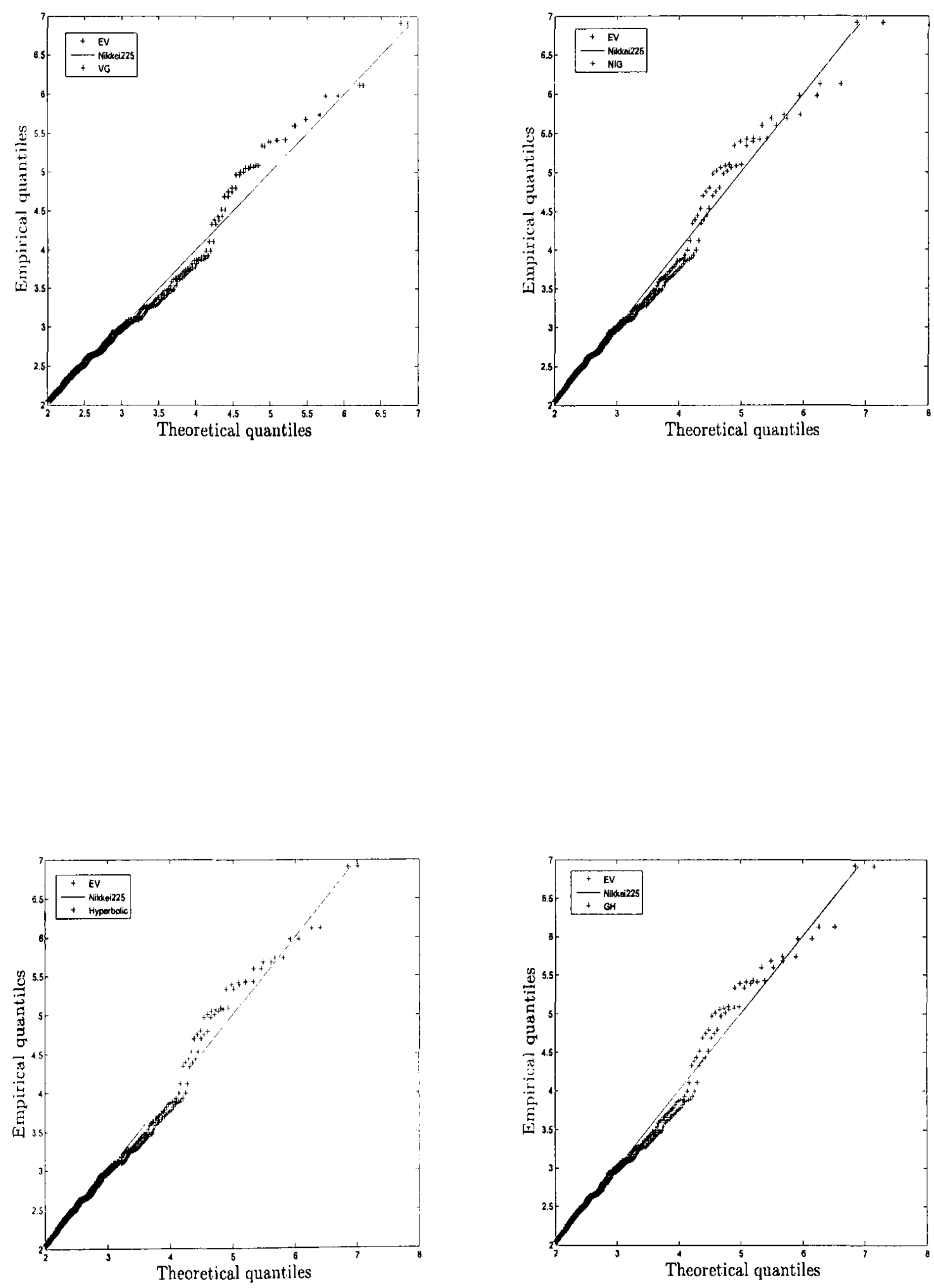


Figure 6.8: *EV and Lévy quantiles in excess of threshold(2): long position in Nikkei225.*

when 'long' position gives a left skewed density, for 'short' position the shape of the density remains same but becomes right skewed. Thus 'long' and 'short' positions just correspond to a sign change of the skewness characterizing parameter. On the other hand for tail based EV when the left skewed density become right skewed the tail observations for 'long' and 'short' positions could be significantly different in numbers and hence the estimates. The asymmetry of the distributions, both for EV and all Lévy, is further confirmed through difference in tail masses corresponding to long and short positions, see tables6.2-6.5.

The tail masses, for observations in excess of thresholds, are observed to be different for extreme-value and Lévy models. Also under different Lévy models, corresponding to same threshold, the tail masses exhibit further difference. As a consequence the corresponding quantiles of extreme-value and Lévy models, as well as those among different Lévy models, do not lie along a vertical line. See figures6.4-6.8. For the sake of illustration with same number of tail observations we use the tail mass of EV model and obtain the QQ-plot of EV model with each of the Lévy models separately.

To visualize tail fits of the models we separately present the EV, that is Generalized Pareto, tail with the tail of each of our considered Lévy model. We do this for all the indices under investigation. Figures6.4-6.8 show the tails for S&P500, FTSE100, DAX, HangSeng and Nikkei225 respectively. The VaR engines are used to obtain the quantiles. We obtain the EV quantiles in excess of thresholds and then obtain the corresponding quantiles from Lévy models. In other words we do not fix the tail mass but fix the thresholds[1]. The consequence is that some of the Lévy quantiles closed to EV thresholds are in fact slightly smaller than the threshold. This is due to the difference in tail masses covered by EV and Lévy models as reported in tables6.1-6.5.

In excess of extreme value thresholds, theoretical quantiles of EV appear to fare well with their empirical as well as Lévy counterparts. In estimation of Lévy models we do have some effects of observations which do not exceed the threshold and are discarded by EV model. This is what individually depicted in figures6.4-6.8 and is reflected in table6.6

[1]Cotter and Dowd(2006)[39] studied the rationale behind these fixations. The idea is to get enough observations on the tails so that the estimation deem credible. Since for this chapter we are using the same data as used by Cotter and Dowd(2006)[39] we simply use the same thresholds.

Index	Model	AD-stat	1%CV	5%CV	10%CV	p-value
S&P500	EV	0.253	2.953*	1.868*	1.490*	0.984
	VG	1.645	2.889*	1.727*	1.220	0.056
	NIG	1.499	3.001*	1.689*	1.179	0.064
	HYP	1.527	2.902*	1.722*	1.212	0.065
	GH	1.461	2.989*	1.706*	1.188	0.069
FTSE100	EV	0.738	3.095*	1.811*	1.317*	0.782
	VG	0.649	2.970*	1.875*	1.374*	0.251
	NIG	0.350	3.026*	1.855*	1.353*	0.353
	HYP	0.469	2.993*	1.867*	1.365*	0.309
	GH	0.239	3.027*	1.838*	1.346*	0.396
DAX	EV	1.014	4.383*	2.440*	1.615*	0.646
	VG	1.410	3.971*	2.410*	1.719*	0.136
	NIG	1.093	4.119*	2.364*	1.668*	0.180
	HYP	1.244	4.003*	2.402*	1.707*	0.158
	GH	1.091	4.115*	2.370*	1.671*	0.180
HangSeng	EV	1.324	5.299*	2.966*	2.099*	0.682
	VG	0.713	5.057*	3.038*	2.153*	0.298
	NIG	0.496	5.289*	2.979*	2.081*	0.353
	HYP	0.565	5.084*	3.028*	2.138*	0.333
	GH	0.537	5.287*	2.954*	2.070*	0.344
Nikkei225	EV	0.798	3.885*	2.474*	1.858*	0.731
	VG	0.715	3.946*	2.537*	1.884*	0.300
	NIG	0.633	3.996*	2.516*	1.864*	0.323
	HYP	0.650	3.980*	2.527*	1.871*	0.317
	GH	0.636	3.993*	2.522*	1.867*	0.321

Table 6.6: *Anderson Darling and left truncated Anderson Darling tests for Lévy and EV models respectively. The p-value for left truncated Anderson Darling test is obtained by Bootstrapping with 1000 resampling.(*) implies that the model survives the test to the corresponding significance level.*

where we produced the results for tail-emphasized Anderson-Darling goodness of fit test. This indicates that tail observations alone suffice tail modeling and observations outside tail do not necessarily carry any positive information for models tail behavior. Thus to grasp extreme nature of random outcome extreme observations alone contain sufficient information. However any statistic which takes other quantiles, in addition to those in the extreme tail, into consideration might loose the reliability in case of EV model.

Among the Lévy models, however, across the indices NIG and GH often show better fits than VG and HYP models which themselves are hardly distinguishable.

6.3 Estimation of risk measures : Methodology and Performance

Apart form few standard cases VaR, is obtained in general as the solution of quantile-integral equation:

$$\int_{x_{min}}^{VaR} f(u)du - \alpha = 0 \tag{6.16}$$

where α is the coverage level.

The perennial problem with VaR is that when it gives the magnitude of loss to a certain level it remains totally non-informative about the extent of losses which could possibly exceed the level.In other words VaR specifies the tail to a given level but leaves the tail completely unattended. On the other hand, in addition to specifying the tail to a given level, ES provides the average of those losses belonging to the specified tail. Thus ES does not consider only the exact quatile to the level but also those quantiles which exceed the exact one.

As in Cotter and Dowd(2006)[39] significantly high αth quantile in EV model, which is also VaR at high confidence level α, is given by:

$$VaR(\alpha) = u + \frac{\beta}{\xi}\left\{\left(\frac{n}{N_u}\alpha\right)^{-\xi} - 1\right\} \tag{6.17}$$

and the expected shortfall(ES), in the same model, with a coverage to the level of α is:

$$ES^{gp}(\alpha) = \frac{VaR(\alpha)}{1-\xi} + \frac{\beta - \xi u}{1-\xi} \tag{6.18}$$

In equation (6.17), n is the total number of observations and N_u is the number of observations which exceeds the threshold u.

In general when VaR does not have any analytic expression estimating expected shortfall could be rather time consuming and is obtained as:

$$ES(\alpha) = \frac{1}{1-\alpha}\int_{\alpha}^{1} VaR(u)du \tag{6.19}$$

$$= \frac{1}{1-\alpha}\int_{\alpha}^{1}\left\{S \Big| \int_{x_{min}}^{S} f(x)dx = u\right\} du \tag{6.20}$$

For Variance Gamma model ES can then be obtained from the equation:

$$ES^{vg}(\alpha) = \frac{1}{1-\alpha}\int_{\alpha}^{1}\left\{S \Big| \int_{x_{min}}^{S}\left[\frac{2e^{\frac{\theta x}{\sigma^2}}\left(\frac{x}{\sqrt{\frac{2\sigma^2}{\nu}+\theta^2}}\right)^{\frac{1}{\nu}-\frac{1}{2}}}{\sigma\sqrt{2\pi}\nu^{\frac{1}{\nu}}} K_{\frac{1}{\nu}-\frac{1}{2}}\left(\frac{x\sqrt{\frac{2\sigma^2}{\nu}+\theta^2}}{\sigma^2}\right)\right] dx = u\right\} du \tag{6.21}$$

Similarly for NIG model the ES is obtained as:

$$ES^{nig}(\alpha) = \frac{1}{1-\alpha}\int_{\alpha}^{1}\left\{S \Big| \int_{x_{min}}^{S}\left[\frac{\alpha\delta e^{\left(\delta\sqrt{\alpha^2-\beta^2}+\beta x\right)}}{\pi}\frac{K_1\left(\alpha\sqrt{\delta^2+x^2}\right)}{\sqrt{\delta^2+x^2}}\right] dx = u\right\} du \tag{6.22}$$

The density of hyperbolic(HYP) model gives the ES for HYP model:

$$ES^{hyp}(\alpha) = \frac{1}{1-\alpha}\int_{\alpha}^{1}\left\{S \Big| \int_{x_{min}}^{S}\left[\frac{\sqrt{\alpha^2-\beta^2}e^{\left(-\alpha\sqrt{\delta^2+x^2}+\beta x\right)}}{2\delta\alpha K_1\left(\delta\sqrt{\alpha^2-\beta^2}\right)}\right] dx = u\right\} du \tag{6.23}$$

Finally the ES of GH model is given as:

$$ES^{gh}(\alpha)$$

$$= \frac{1}{1-\alpha}\int_{\alpha}^{1}\left\{S\Big|\int_{x_{min}}^{S}\left[\frac{(\alpha^2-\beta^2)^{\frac{\nu}{2}}\left(\delta^2+x^2\right)^{\frac{\nu-\frac{1}{2}}{2}}e^{\beta x}}{\sqrt{2\pi}\alpha^{(\nu-\frac{1}{2})}\delta^{\nu}K_{\nu}\left(\delta\sqrt{\alpha^2-\beta^2}\right)}K_{\nu-\frac{1}{2}}\left(\alpha\sqrt{\beta^2+x^2}\right)\right]dx = u\right\}du \tag{6.24}$$

Spectral risk measure, however, does not consider any particular coverage level. Instead given a parameter characterizing the degree of investors risk aversion, SRM considers the whole spectrum of losses with weights obtained as a function of the investors risk aversion parameters. For our benchmark EV model, the closed form VaR formula provides relatively simple expression for SRM as well:

$$M_{\phi}^{gp}(R) = \int_0^1 \frac{Re^{-R(1-\alpha)}}{1-e^{-R}}\left[u+\frac{\beta}{\xi}\left\{\left(\frac{n}{N_u}\alpha\right)^{-\xi}-1\right\}\right]d\alpha \tag{6.25}$$

In case of Lévy models, however, computation of SRM is very time consuming given we are not equipped with any closed form VaR measure:

$$M_{\phi}(R) = \int_0^1 \frac{Re^{-R(1-u)}}{1-e^{-R}}VaR(u)du \tag{6.26}$$

For Variance Gamma model SRM can then be obtained from the equation:

$$M_{\phi}^{vg}(R)$$

$$= \int_0^1 \frac{Re^{-R(1-u)}}{1-e^{-R}}\left\{S\Big|\int_{x_{min}}^{S}\left[\frac{2e^{\frac{\theta x}{\sigma^2}}\left(\frac{x}{\sqrt{\frac{2\sigma^2}{\nu}+\theta^2}}\right)^{\frac{1}{\nu}-\frac{1}{2}}}{\sigma\sqrt{2\pi}\nu^{\frac{1}{\nu}}}K_{\frac{1}{\nu}-\frac{1}{2}}\left(\frac{x\sqrt{\frac{2\sigma^2}{\nu}+\theta^2}}{\sigma^2}\right)\right]dx = u\right\}du \tag{6.27}$$

Similarly for NIG model the SRM is obtained as:

$$M_{\phi}^{nig}(R)$$

$$= \int_0^1 \frac{Re^{-R(1-u)}}{1-e^{-R}}\left\{S\Big|\int_{x_{min}}^{S}\left[\frac{\alpha\delta e^{\left(\delta\sqrt{\alpha^2-\beta^2}+\beta x\right)}}{\pi}\frac{K_1\left(\alpha\sqrt{\delta^2+x^2}\right)}{\sqrt{\delta^2+x^2}}\right]dx = u\right\}du \tag{6.28}$$

The density of HYP model gives the expression of SRM for HYP model:

$$M_{\phi}^{hyp}(R)$$
$$= \int_0^1 \frac{Re^{-R(1-u)}}{1-e^{-R}} \left\{ S \Big| \int_{x_{min}}^{S} \left[\frac{\sqrt{\alpha^2-\beta^2} e^{\left(-\alpha\sqrt{\delta^2+x^2}+\beta x\right)}}{2\delta\alpha K_1\left(\delta\sqrt{\alpha^2-\beta^2}\right)} \right] dx = u \right\} du \tag{6.29}$$

Finally the SRM of GH model is given by:

$$M_{\phi}^{gh}(R)$$
$$= \int_0^1 \frac{Re^{-R(1-u)}}{1-e^{-R}} \left\{ S \Big| \int_{x_{min}}^{S} \left[\frac{(\alpha^2-\beta^2)^{\frac{\nu}{2}} (\delta^2+x^2)^{\frac{\nu-\frac{1}{2}}{2}} e^{\beta x}}{\sqrt{2\pi}\alpha^{(\nu-\frac{1}{2})}\delta^{\nu} K_{\nu}\left(\delta\sqrt{\alpha^2-\beta^2}\right)} K_{\nu-\frac{1}{2}}\left(\alpha\sqrt{\beta^2+x^2}\right) \right] dx = u \right\} du \tag{6.30}$$

The ϕ symbolizes that this version of SRM is contingent upon a particular choice of exponential risk aversion function $\phi(R) = \frac{Re^{-R(1-p)}}{1-e^{-R}}$. Other choices of risk aversion functions are available in the literature, see for example Kevin, Cotter and Sorwar(2006)[44].

The parametric bootstrap is applied to obtain the standard error(SE) and confidence interval(CI) of each risk measure. However given the fact that we are dealing with Lévy models which have no closed form expressions for risk measures, it is practically infeasible to implement bootstrap with large number of resampling.For each resample we draw the same number of uniform(0,1) random numbers as the size of the sample in hand. After sorting these uniform(0,1) numbers in ascending order we find the relevant quantile corresponding to the given coverage level of VaR. This quantile is then used as bootstrap coverage level corresponding to which we obtain the bootstrap VaR and bootstrap ES, for a particular resample, using equations(6.16) and (6.20). Since for any coverage level the VaR equation needs to be solved numerically, the corresponding ES computations takes huge time to find a converging value. This is because any numerical scheme applied to obtain the ES, search the converging limit by evaluating the integrand "vector by vector" and for each element of a vector the VaR needs to be obtained as a solution of the quantile integral equation (6.16).The same process is repeated for each resample to obtain a bootstrap VaR and a bootstrap ES.

The variation in bootstrapped VaR and bootstrapped ES is caused by the variation in bootstrapped confidence level. Since SRM doesn't depend on any particular confidence level, to obtain bootstrapped estimates of SRM we need to randomize the whole spectrum. This forces us to approximate the integral in (6.26) by slicing the spectrum.The reason goes with the fact that because of randomization the integral often fails to converge under numerical routines available in standard softwares such as "quadl" in matlab or those available in "CompEcon" toolbox of Miranda and Fackler(2002)[84]. Thus to have things manageable we consider bootstrapping with 100 resampling.[2]

6.4 VaR based goodness of fit tests

Among different goodness of fit(GOF) tests Anderson Darling(AD) test is particularly suitable to assess the tail based performance of risk management models.[3] Anderson and Darling(1952)[4], Anderson and Darling(1954)[3] proposed a weighing rule in distance based Kolmogorov-Smirnov test which puts more emphasis on the tail observations. For a complete data this test is well established in the literature. Recently Anna,Rachev and Fabozzi(2005)[36] provide a formula for AD-test statistic when observations only on the extreme tail are available and distribution of the complete sample is unknown. It is referred as Anderson Darling test for left-truncated data. This is precisely the case with EV model. For the AD-test, corresponding to complete distributions with closed form densities (as is the case with our Lévy models), the p-values can be obtained analytically. However p-values for the AD-test with left truncated data needs to be obtained either by Monte carlo

[2]Even with 100 resampling we found that a machine with sophisticated configuration takes considerable time to provide SE and CI of ES for any Lévy model, corresponding to a given coverage level. The same is true for SRM with each particular choice of risk aversion parameter. However as Cotter and Dowd(2006) reported SE and CI with 5000 resampling for EV model–which has closed form expressions both for VaR and ES and closed form VaR helps us calculate SRM as well in seconds–we can see that difference with 100 and 5000 resampling is not necessarily significant for VaR and ES. However in case of SRM the difference is enormous. This is because in addition to considering small number of resampling, we evaluate the integral in SRM by considering only 100 slices.This makes the estimation performance of SRM comparable only among Lévy models.

[3]Other GOF tests such as Chi-square test is not comparable between Lévy models which are calibrated on complete data and EV model which is calibrated on left truncated incomplete data.

simulation or Bootstrapping. We carry out Bootstrapping with 1000 resampling to obtain the p-values for the EV model. The critical values are obtained using the VaR engines of the respective models. In case of left truncated AD test, VaR still works as critical value because VaR is computed from left truncated density.

$$AD^2 = -N - \sum_{i=1}^{N} \frac{(2i-1)}{N} \Big[log(F(x_i)) + log(1 - F(x_{N+1-i})) \Big] \tag{6.31}$$

$$\begin{aligned} AD_{ev}^2 &= -n + 2nlog(1 - F(u)) - \frac{1}{n} \sum_{i=1}^{n} (1 + 2(n-i)) \, log(1 - F(x_j)) \cdots \\ &\quad + \frac{1}{n} \sum_{i=1}^{n} (1 - 2i) \, log(F(x_j) - F(u)) \end{aligned} \tag{6.32}$$

Here u is the level of truncation and x_j is the jth observed value of the order statistic $X_1 \leq X_2 \leq \cdots \leq X_n$ and n is the total number of observations available on the tail.

Table6.6 provides evidence to the fact that EV model with calibration based on few extreme observations is well comparable to Lévy models with calibration based on full density. Otherwise said for tail modeling or estimation of solely tail based statistic, full density based models may not deem essential. Thus observations outside the tail do not essentially bear any information which could necessarily improve models likelihood to describe the extreme events better. This point will become clear when we will investigate the tail based risk measures VaR and ES in next section. We will see that for estimating VaR and ES, EV model with calibration on few extreme observations compares pretty well with Lévy models which are calibrated on entire data set.

6.5 Backtesting Risk Models Under Dynamic Calibration

The simple promise of a daily VaR(α) measure is that under all possible extremity the loss from holding an asset for one day could possibly exceed VaR(α)-at most $\alpha \times 100\%$ of the times provided the VaR(α) is estimated on daily returns of the same asset. Given we have an indicator variable ,describing the so called hit sequence which identifies the days of VaR violation in next T trading days, we should not be able to predict on which particular day or days the VaR violation will occur (i.e. the indicator variable will assume '1'); but should

Model	Risk Measures	0.99(or R=20)	0.995(or R=100)	0.999(or R=200)
empirical	VaR	2.952	3.434	5.647
	ES	3.902	4.613	7.246
	SRM	2.265	3.685	4.476
GP	VaR	2.919(0.138) [0.935 1.075]	3.489(0.204) [0.908 1.089]	5.125(0.659) [0.797 1.189]
	ES	3.862(0.162) [0.935 1.070]	4.559(0.237) [0.905 1.093]	6.558(0.802) [0.804 1.175]
	SRM	2.251(0.153) [0.898 1.097]	3.288(0.508) [0.781 1.259]	3.703(0.773) [0.690 1.292]
VG	VaR	2.889(0.121) [0.935 1.072]	3.384(0.165) [0.926 1.094]	4.528(0.374) [0.870 1.145]
	ES	3.604(0.121) [0.948 1.052]	4.099(0.160) [0.943 1.073]	5.243(0.313) [0.904 1.111]
	SRM	2.274(0.952) [0.365 2.007]	3.484(3.256) [0.063 2.819]	4.001(5.080) [0.005 3.333]
NIG	VaR	3.001(0.146) [0.925 1.085]	3.617(0.203) [0.883 1.086]	5.139(0.547) [0.840 1.180]
	ES	3.920(0.155) [0.939 1.061]	4.569(0.214) [0.932 1.087]	6.154(0.445) [0.885 1.136]
	SRM	2.353(1.056) [0.409 1.802]	3.818(3.699) [0.057 2.832]	4.500(5.899) [0.004 3.590]
HYP	VaR	2.902(0.123) [0.934 1.073]	3.409(0.164) [0.898 1.073]	4.587(0.371) [0.889 1.172]
	ES	3.634(0.122) [0.954 1.064]	4.141(0.164) [0.942 1.073]	5.318(0.321) [0.903 1.113]
	SRM	2.282(0.973) [0.431 1.731]	3.518(3.297) [0.063 2.830]	4.049(5.155) [0.005 3.358]
GH	VaR	2.989(0.139) [0.928 1.081]	3.569(0.189) [0.888 1.081]	4.967(0.454) [0.869 1.179]
	ES	3.843(0.142) [0.949 1.071]	4.440(0.195) [0.936 1.082]	5.868(0.395) [0.892 1.126]
	SRM	2.338(1.031) [0.416 1.764]	3.729(3.567) [0.059 2.831]	4.358(5.644) [0.004 3.525]

Table 6.7: *Performance of risk measures on extreme tail:long position in S&P500. SE's are reported besides each estimate and 90% normalized (by means of bootstrapped estimates)CI's are reported right below. The parameters used are those obtained through calibrations.*

Model	Risk Measures	0.99(or R=20)	0.995(or R=100)	0.999(or R=200)
empirical	VaR	3.307	3.738	5.118
	ES	4.018	4.532	5.590
	SRM	2.338	3.739	4.377
GP	VaR	3.058(0.147) [0.924 1.086]	3.674(0.243) [0.908 1.119]	5.273(0.556) [0.837 1.190]
	ES	4.009(0.163) [0.940 1.071]	4.689(0.259) [0.929 1.089]	6.459(0.666) [0.849 1.176]
	SRM	2.234(0.147) [0.889 1.113]	3.415(0.551) [0.766 1.232]	3.848(0.808) [0.699 1.323]
VG	VaR	2.969(0.111) [0.942 1.064]	3.424(0.145) [0.938 1.078]	4.453(0.324) [0.882 1.129]
	ES	3.617(0.109) [0.953 1.052]	4.062(0.143) [0.925 1.053]	5.078(0.326) [0.904 1.102]
	SRM	2.319(0.803) [0.529 1.685]	3.423(3.147) [0.067 2.738]	3.881(4.864) [0.005 3.217]
NIG	VaR	3.026(0.124) [0.937 1.071]	3.542(0.168) [0.931 1.088]	4.772(0.452) [0.866 1.191]
	ES	3.782(0.127) [0.948 1.058]	4.309(0.172) [0.916 1.060]	5.565(0.414) [0.889 1.118]
	SRM	2.352(0.832) [0.521 1.701]	3.574(3.349) [0.063 2.824]	4.112(5.245) [0.005 3.403]
HYP	VaR	2.993(0.116) [0.939 1.067]	3.470(0.153) [0.935 1.082]	4.569(0.300) [0.887 1.123]
	ES	3.679(0.115) [0.951 1.054]	4.153(0.153) [0.922 1.056]	5.249(0.359) [0.893 1.119]
	SRM	2.331(0.814) [0.525 1.694]	3.478(2.924) [0.052 2.567]	3.964(4.998) [0.005 3.279]
GH	VaR	3.027(0.131) [0.933 1.075]	3.579(0.184) [0.926 1.096]	4.980(0.412) [0.863 1.156]
	ES	3.865(0.140) [0.944 1.062]	4.462(0.197) [0.908 1.067]	5.966(0.523) [0.866 1.156]
	SRM	2.363(0.846) [0.518 1.708]	3.658(3.217) [0.049 2.616]	4.262(5.539) [0.004 3.597]

Table 6.8: *Performance of risk measures on extreme tail:long position in FTSE100. SE's are reported besides each estimate and 90% normalized (by means of bootstrapped estimates)CI's are reported right below. The parameters used are those obtained through calibrations.*

Model	Risk Measures	0.99(or R=20)	0.995(or R=100)	0.999(or R=200)
empirical	VaR	4.397	4.940	6.594
	ES	5.511	6.330	9.738
	SRM	3.185	5.219	6.276
GP	VaR	4.330(0.195) [0.926 1.074]	5.178(0.326) [0.914 1.115]	7.174(0.659) [0.860 1.151]
	ES	5.563(0.212) [0.939 1.063]	6.421(0.276) [0.919 1.076]	8.441(0.691) [0.858 1.116]
	SRM	3.035(0.218) [0.889 1.103]	4.761(0.684) [0.787 1.249]	5.336(1.036) [0.692 1.255]
VG	VaR	3.971(0.165) [0.919 1.063]	4.634(0.257) [0.925 1.102]	6.157(0.556) [0.872 1.165]
	ES	4.922(0.160) [0.949 1.056]	5.579(0.212) [0.919 1.057]	7.096(0.411) [0.907 1.108]
	SRM	3.072(1.292) [0.438 1.723]	4.668(3.866) [0.104 2.434]	5.345(5.831) [0.012 3.374]
NIG	VaR	4.119(0.206) [0.931 1.079]	4.932(0.282) [0.909 1.085]	6.928(0.709) [0.828 1.149]
	ES	5.328(0.204) [0.941 1.066]	6.179(0.279) [0.905 1.069]	8.248(0.581) [0.888 1.132]
	SRM	3.162(1.398) [0.417 1.776]	5.055(4.388) [0.096 2.481]	5.925(6.665) [0.011 3.447]
HYP	VaR	4.003(0.171) [0.917 1.065]	4.690(0.240) [0.915 1.085]	6.284(0.548) [0.851 1.126]
	ES	4.994(0.167) [0.948 1.058]	5.680(0.222) [0.917 1.059]	7.273(0.433) [0.904 1.111]
	SRM	3.088(1.310) [0.434 1.727]	4.733(3.947) [0.103 2.441]	5.440(5.962) [0.012 3.388]
GH	VaR	4.115(0.203) [0.932 1.078]	4.916(0.277) [0.910 1.084]	6.865(0.689) [0.831 1.146]
	ES	5.300(0.199) [0.942 1.065]	6.133(0.272) [0.907 1.067]	8.141(0.560) [0.890 1.129]
	SRM	3.158(1.390) [0.418 1.766]	5.027(4.341) [0.097 2.476]	5.879(6.591) [0.011 3.441]

Table 6.9: *Performance of risk measures on extreme tail:long position in DAX. SE's are reported besides each estimate and 90% normalized (by means of bootstrapped estimates)CI's are reported right below. The parameters used are those obtained through calibrations.*

Model	Risk Measures	0.99(or R=20)	0.995(or R=100)	0.999(or R=200)
empirical	VaR	5.269	6.273	9.169
	ES	7.099	8.454	11.791
	SRM	3.974	6.682	8.169
GP	VaR	5.230(0.279) [0.912 1.099]	6.385(0.367) [0.907 1.092]	9.494(1.083) [0.817 1.198]
	ES	7.056(0.296) [0.932 1.079]	8.379(0.515) [0.916 1.114]	11.939(1.156) [0.818 1.154]
	SRM	3.757(0.253) [0.899 1.110]	5.943(1.026) [0.749 1.256]	6.781(1.779) [0.685 1.346]
VG	VaR	5.057(0.209) [0.936 1.071]	5.917(0.278) [0.901 1.071]	7.901(0.642) [0.878 1.129]
	ES	6.302(0.209) [0.948 1.057]	7.169(0.280) [0.918 1.059]	9.221(0.697) [0.889 1.120]
	SRM	3.919(1.659) [0.435 1.727]	5.995(4.984) [0.103 2.438]	6.879(7.524) [0.012 3.382]
NIG	VaR	5.288(0.258) [0.925 1.084]	6.372(0.356) [0.883 1.085]	9.052(0.901) [0.853 1.159]
	ES	6.906(0.289) [0.942 1.066]	8.048(0.395) [0.922 1.073]	10.835(0.989) [0.846 1.133]
	SRM	4.059(1.820) [0.409 1.800]	6.581(5.769) [0.094 2.491]	7.754(8.777) [0.010 3.459]
HYP	VaR	5.084(0.215) [0.935 1.073]	5.968(0.286) [0.899 1.073]	8.019(0.666) [0.876 1.132]
	ES	6.359(0.214) [0.948 1.058]	7.243(0.286) [0.917 1.059]	9.294(0.666) [0.893 1.113]
	SRM	3.937(1.678) [0.431 1.730]	6.064(5.067) [0.102 2.444]	6.978(7.658) [0.011 3.393]
GH	VaR	5.285(0.266) [0.923 1.087]	6.413(0.374) [0.878 1.089]	9.281(0.980) [0.845 1.169]
	ES	6.997(0.306) [0.939 1.069]	8.215(0.424) [0.918 1.077]	11.266(1.102) [0.837 1.143]
	SRM	4.068(1.848) [0.405 1.832]	6.678(5.957) [0.092 2.543]	7.927(9.069) [0.010 3.480]

Table 6.10: *Performance of risk measures on extreme tail:long position in HangSeng. SE's are reported besides each estimate and 90% normalized (by means of bootstrapped estimates)CI's are reported right below. The parameters used are those obtained through calibrations.*

Model	Risk Measures	0.99(or R=20)	0.995(or R=100)	0.999(or R=200)
empirical	VaR	3.748	4.780	5.891
	ES	4.856	5.536	6.614
	SRM	3.009	4.535	5.289
GP	VaR	3.848(0.144) [0.946 1.062]	4.447(0.223) [0.927 1.090]	5.821(0.472) [0.882 1.134]
	ES	4.706(0.153) [0.951 1.060]	5.299(0.200) [0.938 1.055]	6.657(0.466) [0.892 1.101]
	SRM	2.892(0.206) [0.876 1.095]	4.118(0.578) [0.772 1.243]	4.500(0.946) [0.693 1.339]
VG	VaR	3.946(0.158) [0.942 1.063]	4.526(0.189) [0.937 1.069]	5.833(0.382) [0.887 1.108]
	ES	4.769(0.139) [0.955 1.050]	5.336(0.182) [0.928 1.051]	6.622(0.345) [0.916 1.097]
	SRM	2.944(1.015) [0.532 1.679]	4.317(2.889) [0.049 2.576]	4.882(5.787) [0.003 3.370]
NIG	VaR	3.996(0.163) [0.928 1.064]	4.636(0.206) [0.936 1.083]	6.142(0.541) [0.879 1.159]
	ES	4.926(0.157) [0.951 1.055]	5.574(0.210) [0.921 1.057]	7.098(0.500) [0.895 1.112]
	SRM	2.963(1.246) [0.396 1.692]	4.427(4.112) [0.065 2.797]	5.058(5.983) [0.010 3.058]
HYP	VaR	3.980(0.148) [0.931 1.057]	4.592(0.180) [0.934 1.059]	5.995(0.434) [0.877 1.135]
	ES	4.857(0.148) [0.953 1.052]	5.463(0.195) [0.948 1.066]	6.861(0.457) [0.896 1.117]
	SRM	2.956(1.134) [0.486 1.665]	4.380(3.435) [0.045 2.382]	4.979(6.260) [0.005 3.255]
GH	VaR	3.993(0.160) [0.928 1.063]	4.622(0.201) [0.937 1.081]	6.083(0.520) [0.882 1.155]
	ES	4.901(0.153) [0.952 1.054]	5.531(0.204) [0.922 1.056]	6.997(0.478) [0.898 1.108]
	SRM	2.962(1.241) [0.397 1.689]	4.411(4.084) [0.066 2.779]	5.029(5.917) [0.010 3.032]

Table 6.11: *Performance of risk measures on extreme tail:long position in Nikkei225. SE's are reported besides each estimate and 90% normalized (by means of bootstrapped estimates)CI's are reported right below. The parameters used are those obtained through calibrations.*

just be able to say that only $\alpha \times 100\%$ of those T days VaR could possibly be violated. In other words the hit sequence is Bernoulli distributed with a probability α of assuming '1', see Christoffersen(2003)[34], McNeil et al(2005)[83].This lays the foundation to verify the performance of any VaR model and is known in the literature as backtesting. In most applications α is very low usually 1% or 5% (or very high usually 99% or 95% if VaR is estimated on returns multiplied by -1). See Christoffersen(2003)[34].

There are three useful hypotheses used in backtesting risk models. The unconditional hypothesis does not have any assumption on today's violation status when it provides statistical evidence whether the likelihood that tomorrow will be a violation is significantly different from VaR models promised fraction α. This evidence is provided through the test statistic asymptotically following χ^2 with one degree of freedom:

$$LR_{uc} = -2log\left[\frac{(1-\alpha)^{T_0}\alpha^{T_1}}{\left(1-\frac{T_1}{T}\right)^{T_0}\left(\frac{T_1}{T}\right)^{T_1}}\right] \sim \chi^2(1) \tag{6.33}$$

Here $T = T_1 + T_0$ is assumed to be significantly large and T_0 and T_1 are the number of days with no violation(i.e. number of zeros in the hit sequence) and number of days with violation(i.e. number of ones in the hit sequence).

The requirement that T needs to be large enough is hardly met in application which forces us to compute the Monte-Carlo(MC) p-value and rely on it more than χ^2 p-value. One frequently used technique to compute MC p-value is to simulate 999 test values $\{LR(i)\}_{i=1}^{999}$, each of which is based on a Bernoulli(α) sample of hit sequence having the same size as the original sample in hand:

$$p^{mc} = \frac{1}{1000}\left\{1 + \sum_{i=1}^{999}\mathbb{I}_{\left[\hat{LR}(i) > LR_{uc}\right]}\right\} \tag{6.34}$$

That is the simulated p-value is the ratio of simulated test value, to the number of simulation, given that simulated test value is more significant than the test value associated with original sample.

Independence test helps us verify whether VaR violations are truly occurring randomly and are not clustered over time. Assets exhibiting clustering of variance are highly likely to exhibit the clustering of VaR violations which is an indication of misspecification of risk model. This is because under such circumstances it is possible to predict with some certainty that if today is a violation tomorrow is highly likely to be a violation too i.e. more than $\alpha \times 100\%$ likely. If any forecast is available regarding high volatility then the risk model should use that information and adjust the VaR accordingly. This will ensure that if the risk model is correctly specified than the violation of VaR should remain an unpredictable event. As in Chriftoffersen(2003)[34] the test statistic associated with Independence test is given by :

$$LR_{ind} = -2log\left[\frac{L(p)}{L(\hat{\prod}_1)}\right] \sim \chi^2 \tag{6.35}$$

where matrix of transitional probabilities of conditional violations is given by:

$$\hat{\prod}_1 = \begin{bmatrix} \frac{T_{00}}{T_{00}+T_{01}} & \frac{T_{01}}{T_{00}+T_{01}} \\ \frac{T_{10}}{T_{10}+T_{11}} & \frac{T_{11}}{T_{10}+T_{11}} \end{bmatrix} \tag{6.36}$$

Thus:

$$L(\hat{\prod}_1) = \left(\frac{T_{00}}{T_{00}+T_{01}}\right)^{T_{00}} \left(\frac{T_{01}}{T_{00}+T_{01}}\right)^{T_{01}} \left(\frac{T_{10}}{T_{10}+T_{11}}\right)^{T_{10}} \left(\frac{T_{11}}{T_{10}+T_{11}}\right)^{T_{11}} \tag{6.37}$$

and p is such that it characterizes the matrix of transitional probabilities of violation assuming there is no dependence between 0 and 1 in the hit sequence:

$$\hat{\prod} = \begin{bmatrix} 1-p & p \\ 1-p & p \end{bmatrix} = \begin{bmatrix} \frac{T_0}{T_0+T_1} & \frac{T_1}{T_0+T_1} \\ \frac{T_0}{T_0+T_1} & \frac{T_1}{T_0+T_1} \end{bmatrix} \tag{6.38}$$

Consequently $L(p)$ is similar to the one given for unconditional hypothesis in (6.33). Thus LR_{ind} provides statistical significance of likelihood of independence in hit sequence over the likelihood of dependence. As usual MC p-values are usually much more reliable than $\chi^2(1)$ p-values.

Finally conditional coverage test helps us verify whether the average number of violations is consistent with the coverage level of risk model. However this test comes jointly with independence test. Conditional coverage test statistic has the similar expression as the independence test statistic with $p = \frac{T_1}{T}$ of independence statistic replaced by the coverage level α of risk model:

$$\begin{aligned} LR_{cc} &= -2log\left[\frac{LR(\alpha)}{LR(\hat{\prod}_1)}\right] & (6.39)\\ &= LR_{uc} + LR(ind) \sim \chi^2(2) & (6.40)\end{aligned}$$

See Chiristoffersen(2003)[34], Dowd(2005)[42] e.g. Again MC p-values are often found reliable in application.

Risk measures sensitivity to new observation can make significant difference in models backtesting performance. This section focuses on the risk measure VaR and its sensitivity to new observations for dynamic calibration with a rolling window of four business years, 1008 trading days.The choice of length of a rolling look-back window is subjective.[4] The significance of fixed-length look-back window goes with the fact that backtesting results remain equally decisive to the "news" in any given years observation. An window expanding from an initial fixed length one may seem easier and more comfortable to implement but the downside of doing so is that as time goes by the "news" contained in the most recent observations receives less weight relative to the increasing number of older observations in the look-back sample. The rolling approach is recently used in actuarial literature and is found providing respectable results, see e.g. Dowd et al(2010)[43]. For EV, extreme 30% observations are considered for calibration in a dynamic fashion and for Lévy models all the observations available in a rolling window of four business years are considered. Given the fact that we would like to assess the Lévy models on extreme tail with the EV model, we consider the usual coverage of 95% and 99%. The dynamic calibration starts at 1st January 1995 and ends at 31st December, 2003. On 31st December 1994 we calibrate all the models on the time series of returns for the period 1991-1994 and we use the calibrated parameters to predict the VaR for 1st January 1995. However on 1st Jan 1995 we have one new observation of return, so we remove the oldest observation to accommodate this new observation in our fixed length look-back window. We then calibrate the models with observations in new window and predict the VaR for 2nd Jan 1995. The same mechanism of calibrating the models and predicting the VaR's continues in dynamic fashion until the end of 2003. Figures6.9-6.13 show the VaR measures for long positions in S&P500, FTSE100, DAX, HangSeng, Nikkei225 respectively. In each figure we present the 95% coverage VaR on

[4]We start with a two year look-back window and keep increasing the window length as long as we become satisfied with the performance of dynamic calibration. The problem arises basically with the convergence in calibration corresponding to EV model, as it considers only extreme observations. To overcome this problem, in addition to expanding the length of look-back window we increase the proportion of extreme observations (used in calibration) by adjusting the threshold.

top and the 99% coverage VaR at the bottom. Subsequently Unconditional, Independence and Conditional coverage hypotheses are implemented for 95% and 99% VaR's. Backtesting results appearing in tables6.12-6.16 are respectively for long positions in S&P500, FTSE100, DAX, HangSeng and Nikkei225.

6.6 Estimation of Coherent Risk Measures Using FRFT

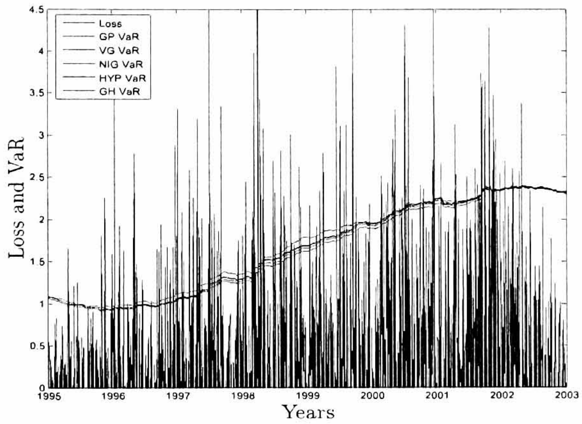
This section is work in progress. It introduces an alternative technique to estimate risk measure VaR and coherent risk measures ES and SRM using models characteristic functions and FRFT engine.This approach is found much faster than the approach we used in previous sections where we used numerical root search technique to find VaR using complicated density expressions. The FRFT based estimation of risk measures becomes feasible because of the following relation involving the characteristic function and the CDF of an underlying processes. Some simple transformations applied to Fourier inversion formula yield the classic relation , see e.g.Roussas(2005)[99]:

$$F(x) = \frac{e^{\alpha x}}{2\pi} \int_{-\infty}^{\infty} e^{-iux} \frac{\Phi(u + i\alpha)}{\alpha - iu} du \tag{6.41}$$

Here α is a constant and is known as dampening factor. The role of α is to ensure that the integral converges indeed. Often this α takes a value in $(0, 5]$ depending on the data in hand. For a reasonably fine grid of Φ, FRFT can be utilized to obtain a vector of values of F in seconds[5]. Then for any arbitrary coverage of VaR it just remains to figure out between which two F nodes the coverage level falls, and to collect the corresponding x values. An interpolation will then yield the particular x, corresponding to the given coverage $p(say)$, such that $F(x) = p$. This x is the value of our risk measure VaR. This way we are completely avoiding the complicated density expressions and are just using the discretized values of characteristic function.[6] For a vector of coverage probabilities Matlab can perform such interpolation in fraction of a second. Performance of FRFT VaR compare to traditional root search VaR is presented in figure6.14. This figure also shows the grid size VaR as well as grid size computational time trade off.

[5]Thc details on how to apply FRFT are discussed in chapter3.

[6]For risk management we do not use risk neutralized dynamics, that is the characteristic function must not be risk-neutral one.



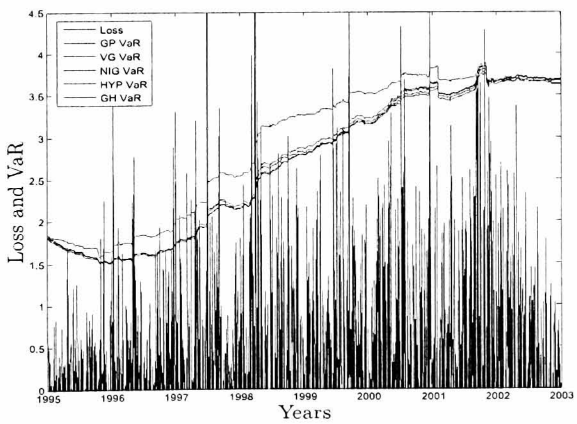


Figure 6.9: *Dynamic VaR for Extreme Value and Lévy Models: S&P500 case. A rolling window of four years is considered. The top panel is for* 95% *coverage and the bottom panel is for 99% coverage.*

		Stat.	1%CV	5%CV	10%CV	p-χ^2	p-MC	PV
GP	UC(95%)	21.47	6.63	3.84	2.70	3.6e-6	1e-3	7.27
	Ind(95%)	0.12	6.63	3.84	2.70	0.723	0.738	7.27
	CC(95%)	21.59	9.21	5.99	4.60	2e-5	1e-3	7.27
	UC(99%)	0.89	6.63	3.84	2.70	0.346	0.358	1.20
	Ind(99%)	0.93	6.63	3.84	2.70	0.334	0.163	1.20
	CC(99%)	1.82	9.21	5.99	4.60	0.402	0.302	1.20
VG	UC(95%)	30.08	6.63	3.84	2.70	4e-8	1e-3	7.71
	Ind(95%)	1.08	6.63	3.84	2.70	0.296	0.318	7.71
	CC(95%)	31.17	9.21	5.99	4.60	2e-7	1e-3	7.71
	UC(99%)	23.78	6.63	3.84	2.70	1e-6	1e-3	2.18
	Ind(99%)	0.67	6.63	3.84	2.70	0.411	0.256	2.18
	CC(99%)	24.45	9.21	5.99	4.60	5e-6	1e-3	2.18
NIG	UC(95%)	33.89	6.63	3.84	2.70	6e-9	1e-3	7.89
	Ind(95%)	1.26	6.63	3.84	2.70	0.260	0.267	7.89
	CC(95%)	35.15	9.21	5.99	4.60	2e-8	1e-3	7.89
	UC(99%)	19.21	6.63	3.84	2.70	1e-5	1e-3	2.05
	Ind(99%)	0.94	6.63	3.84	2.70	0.332	0.192	2.05
	CC(99%)	20.15	9.21	5.99	4.60	4e-5	1e-3	2.05
HYP	UC(95%)	30.08	6.63	3.84	2.70	4e-8	1e-3	7.71
	Ind(95%)	1.08	6.63	3.84	2.70	0.296	0.297	7.71
	CC(95%)	31.17	9.21	5.99	4.60	2e-7	1e-3	7.71
	UC(99%)	20.69	6.63	3.84	2.70	5e-6	1e-3	2.09
	Ind(99%)	0.85	6.63	3.84	2.70	0.357	0.179	2.09
	CC(99%)	21.53	9.21	5.99	4.60	2e-5	1e-3	2.09
GH	UC(95%)	36.87	6.63	3.84	2.70	1e-9	1e-3	8.02
	Ind(95%)	1.56	6.63	3.84	2.70	0.211	0.223	8.02
	CC(95%)	38.43	9.21	5.99	4.60	4e-9	1e-3	8.02
	UC(99%)	20.69	6.63	3.84	2.70	5e-6	1e-3	2.10
	Ind(99%)	0.84	6.63	3.84	2.70	0.357	0.187	2.10
	CC(99%)	21.53	9.21	5.99	4.60	2e-5	1e-3	2.10

Table 6.12: *Backtesting results for conditional and unconditional models: S&P500. UC stands for Unconditional Coverage,Ind stands for Independence Test, CC stands for Conditional Coverage. PV stands for proportion of VaR violation.p-values from both Chisquare and Monte-Carlo simulations are reported.*

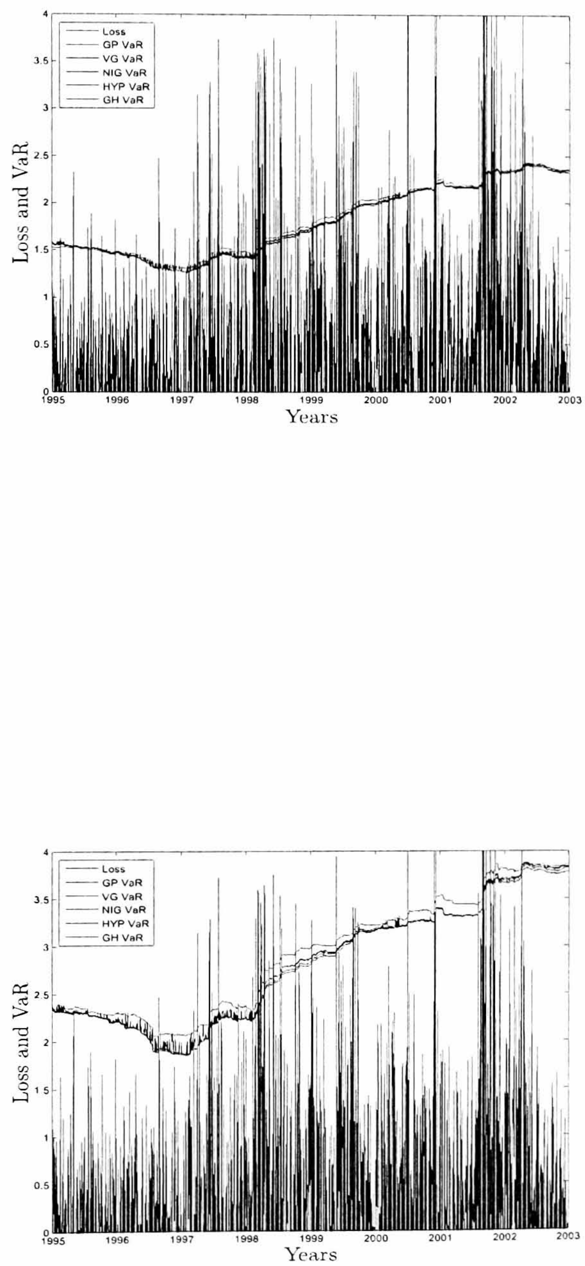


Figure 6.10: *Dynamic VaR for Extreme Value and Lévy Models: FTSE100 case. A rolling window of four years is considered. The top panel is for* 95% *coverage and the bottom panel is for 99% coverage.*

		Stat.	1%CV	5%CV	10%CV	p-χ^2	p-MC	PV
GP	UC(95%)	4.96	6.63	3.84	2.70	0.025	0.027	6.06
	Ind(95%)	8.51	6.63	3.84	2.70	0.003	0.009	6.06
	CC(95%)	13.47	9.21	5.99	4.60	0.001	0.002	6.06
	UC(99%)	13.53	6.63	3.84	2.70	6e-4	1e-3	1.79
	Ind(99%)	1.61	6.63	3.84	2.70	0.203	0.082	1.79
	CC(99%)	13.15	9.21	5.99	4.60	0.001	1e-3	1.79
VG	UC(95%)	6.25	6.63	3.84	2.70	0.012	0.013	6.19
	Ind(95%)	9.34	6.63	3.84	2.70	0.002	0.004	6.19
	CC(95%)	15.60	9.21	5.99	4.60	4e-4	0.002	6.19
	UC(99%)	19.53	6.63	3.84	2.70	1e-6	1e-3	2.06
	Ind(99%)	2.98	6.63	3.84	2.70	0.083	0.026	2.06
	CC(99%)	12.52	9.21	5.99	4.60	1e-5	1e-3	2.06
NIG	UC(95%)	6.72	6.63	3.84	2.70	0.009	0.010	6.24
	Ind(95%)	9.02	6.63	3.84	2.70	0.002	0.004	6.24
	CC(95%)	15.74	9.21	5.99	4.60	4e-4	1e-3	6.24
	UC(99%)	18.08	6.63	3.84	2.70	2e-5	1e-3	2.02
	Ind(99%)	3.18	6.63	3.84	2.70	0.074	0.023	2.02
	CC(99%)	21.26	9.21	5.99	4.60	2e-5	1e-3	2.02
HYP	UC(95%)	6.25	6.63	3.84	2.70	0.012	0.015	6.19
	Ind(95%)	9.34	6.63	3.84	2.70	0.002	0.007	6.19
	CC(95%)	15.60	9.21	5.99	4.60	4e-4	1e-3	6.19
	UC(99%)	18.08	6.63	3.84	2.70	2e-5	1e-3	2.02
	Ind(99%)	3.18	6.63	3.84	2.70	0.074	0.023	2.02
	CC(99%)	21.26	9.21	5.99	4.60	2e-5	1e-3	2.02
GH	UC(95%)	5.81	6.63	3.84	2.70	0.015	0.014	6.15
	Ind(95%)	9.67	6.63	3.84	2.70	0.001	0.005	6.15
	CC(95%)	15.48	9.21	5.99	4.60	4e-4	1e-3	6.15
	UC(99%)	16.67	6.63	3.84	2.70	4e-5	1e-3	1.97
	Ind(99%)	3.38	6.63	3.84	2.70	0.066	0.017	1.97
	CC(99%)	20.06	9.21	5.99	4.60	4e-5	1e-3	1.97

Table 6.13: *Backtesting results for conditional and unconditional models: FTSE100. UC stands for Unconditional Coverage,Ind stands for Independence Test, CC stands for Conditional Coverage. PV stands for proportion of VaR violation.p-values from both Chisquare and Monte-Carlo simulations are reported.*

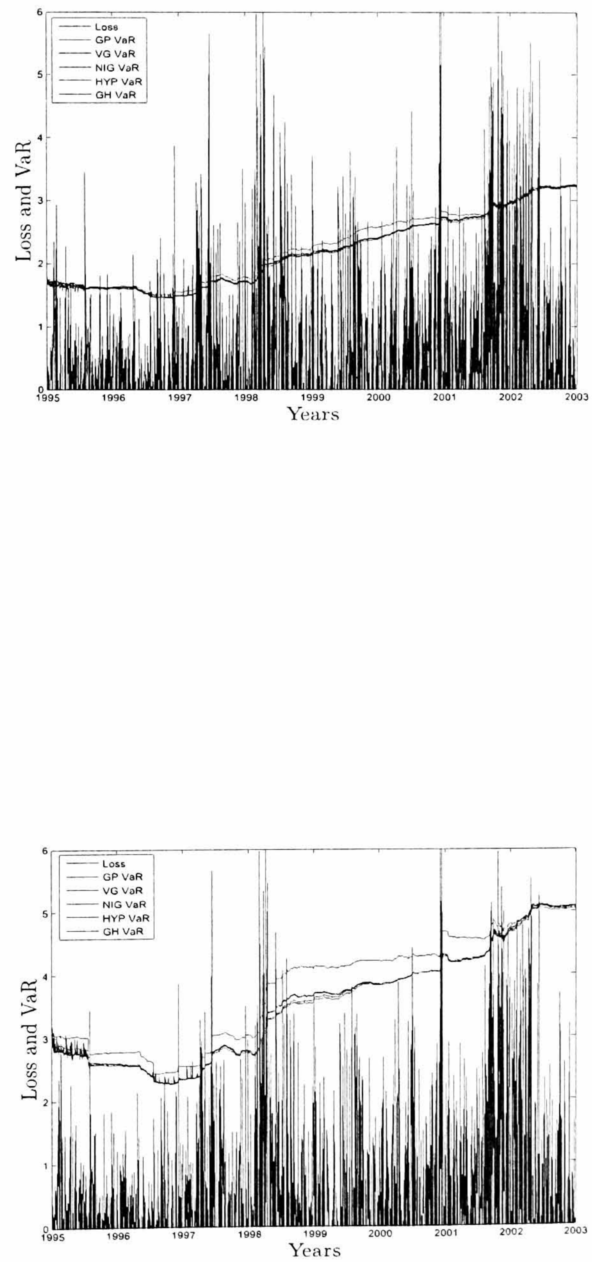


Figure 6.11: *Dynamic VaR for Extreme Value and Lévy Models: DAX case. A rolling window of four years is considered. The top panel is for* 95% *coverage and the bottom panel is for 99% coverage.*

		Stat.	1%CV	5%CV	10%CV	p-χ^2	p-MC	PV
GP	UC(95%)	21.56	6.63	3.84	2.70	3e-6	1e-3	7.28
	Ind(95%)	17.38	6.63	3.84	2.70	3e-5	1e-3	7.28
	CC(95%)	38.94	9.21	5.99	4.60	3e-9	1e-3	7.28
	UC(99%)	9.07	6.63	3.84	2.70	0.003	0.002	1.69
	Ind(99%)	4.81	6.63	3.84	2.70	0.028	0.011	1.69
	CC(99%)	13.89	9.21	5.99	4.60	9e-4	1e-3	1.69
VG	UC(95%)	25.72	6.63	3.84	2.70	4e-7	1e-3	7.50
	Ind(95%)	17.18	6.63	3.84	2.70	3e-5	1e-3	7.50
	CC(95%)	42.90	9.21	5.99	4.60	5e-10	1e-3	7.50
	UC(99%)	25.44	6.63	3.84	2.70	4e-7	1e-3	2.23
	Ind(99%)	15.42	6.63	3.84	2.70	8e-5	1e-3	2.23
	CC(99%)	40.86	9.21	5.99	4.60	1e-9	1e-3	2.23
NIG	UC(95%)	27.47	6.63	3.84	2.70	1e-7	1e-3	7.58
	Ind(95%)	16.29	6.63	3.84	2.70	5e-5	1e-3	7.58
	CC(95%)	43.77	9.21	5.99	4.60	3e-10	1e-3	7.58
	UC(99%)	20.73	6.63	3.84	2.70	5e-6	1e-3	2.09
	Ind(99%)	12.78	6.63	3.84	2.70	3e-4	0.002	2.09
	CC(99%)	33.51	9.21	5.99	4.60	5e-8	1e-3	2.09
HYP	UC(95%)	27.47	6.63	3.84	2.70	1e-7	1e-3	7.58
	Ind(95%)	16.29	6.63	3.84	2.70	5e-5	1e-3	7.58
	CC(95%)	43.77	9.21	5.99	4.60	3e-10	1e-3	7.58
	UC(99%)	25.44	6.63	3.84	2.70	4e-7	1e-3	2.23
	Ind(99%)	15.45	6.63	3.84	2.70	8e-5	1e-3	2.23
	CC(99%)	40.86	9.21	5.99	4.60	1e-9	1e-3	2.23
GH	UC(95%)	29.28	6.63	3.84	2.70	6e-8	1e-3	7.67
	Ind(95%)	15.44	6.63	3.84	2.70	8e-5	0.002	7.67
	CC(95%)	44.72	9.21	5.99	4.60	2e-10	1e-3	7.67
	UC(99%)	22.26	6.63	3.84	2.70	2e-6	1e-3	2.14
	Ind(99%)	12.31	6.63	3.84	2.70	4e-4	1e-3	2.14
	CC(99%)	34.58	9.21	5.99	4.60	3e-8	1e-3	2.14

Table 6.14: *Backtesting results for conditional and unconditional models: DAX. UC stands for Unconditional Coverage,Ind stands for Independence Test, CC stands for Conditional Coverage. PV stands for proportion of VaR violation.p-values from both Chisquare and Monte-Carlo simulations are reported.*

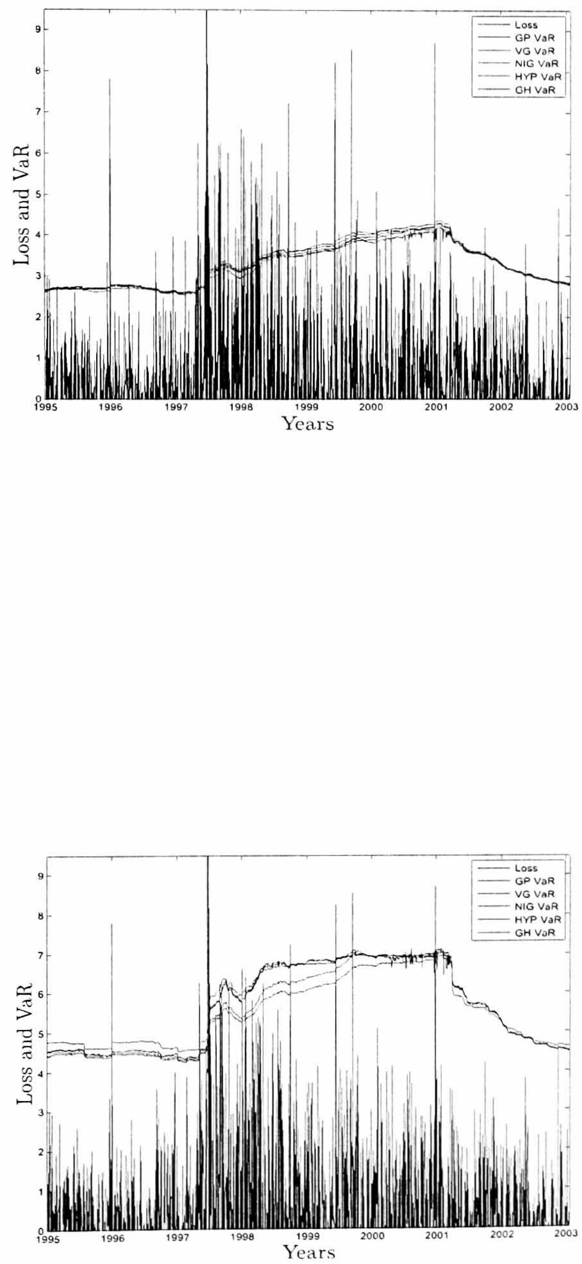
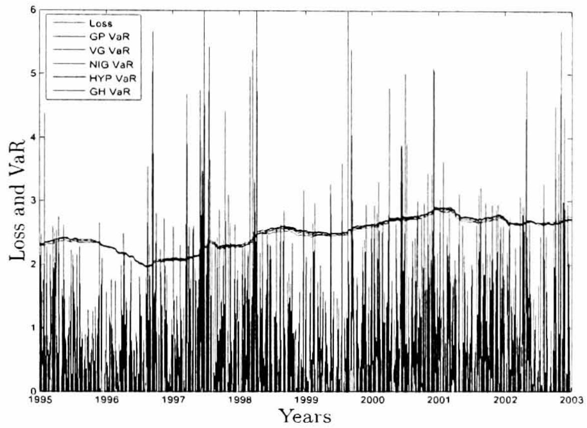


Figure 6.12: *Dynamic VaR for Extreme Value and Lévy Models: HangSeng case. A rolling window of four years is considered. The top panel is for* 95% *coverage and the bottom panel is for 99% coverage.*

		Stat.	1%CV	5%CV	10%CV	p-χ^2	p-MC	PV
GP	UC(95%)	1.79	6.63	3.84	2.70	0.180	0.154	4.39
	Ind(95%)	22.67	6.63	3.84	2.70	2e-6	1e-3	4.39
	CC(95%)	24.46	9.21	5.99	4.60	5e-6	1e-3	4.39
	UC(99%)	1.09	6.63	3.84	2.70	0.295	0.250	0.78
	Ind(99%)	7.52	6.63	3.84	2.70	0.006	0.003	0.78
	CC(99%)	8.61	9.21	5.99	4.60	0.013	0.006	0.78
VG	UC(95%)	2.07	6.63	3.84	2.70	0.149	0.140	4.34
	Ind(95%)	23.26	6.63	3.84	2.70	1e-6	1e-3	4.34
	CC(95%)	25.33	9.21	5.99	4.60	3e-6	1e-3	4.34
	UC(99%)	0.49	6.63	3.84	2.70	0.481	0.461	1.15
	Ind(99%)	9.24	6.63	3.84	2.70	0.002	0.002	1.15
	CC(99%)	9.74	9.21	5.99	4.60	0.007	0.006	1.15
NIG	UC(95%)	1.52	6.63	3.84	2.70	0.216	0.202	4.43
	Ind(95%)	22.09	6.63	3.84	2.70	2e-6	1e-3	4.43
	CC(95%)	23.62	9.21	5.99	4.60	7e-6	1e-3	4.43
	UC(99%)	0.13	6.63	3.84	2.70	0.716	0.710	0.92
	Ind(99%)	11.93	6.63	3.84	2.70	5e-4	1e-3	0.92
	CC(99%)	12.07	9.21	5.99	4.60	0.002	1e-3	0.92
HYP	UC(95%)	1.79	6.63	3.84	2.70	0.180	0.175	4.39
	Ind(95%)	22.67	6.63	3.84	2.70	2e-6	1e-3	4.39
	CC(95%)	24.46	9.21	5.99	4.60	5e-6	1e-3	4.39
	UC(99%)	0.49	6.63	3.84	2.70	0.481	0.442	1.15
	Ind(99%)	9.24	6.63	3.84	2.70	0.002	1e-3	1.15
	CC(99%)	9.74	9.21	5.99	4.60	0.008	0.006	1.15
GH	UC(95%)	0.86	6.63	3.84	2.70	0.352	0.330	4.57
	Ind(95%)	23.56	6.63	3.84	2.70	1e-6	1e-3	4.57
	CC(95%)	24.43	9.21	5.99	4.60	5e-6	1e-3	4.57
	UC(99%)	0.13	6.63	3.84	2.70	0.716	0.670	0.92
	Ind(99%)	11.93	6.63	3.84	2.70	5e-4	1e-3	0.92
	CC(99%)	12.07	9.21	5.99	4.60	0.002	0.003	0.92

Table 6.15: *Backtesting results for conditional and unconditional models: HangSeng. UC stands for Unconditional Coverage,Ind stands for Independence Test, CC stands for Conditional Coverage. PV stands for proportion of VaR violation.p-values from both Chisquare and Monte-Carlo simulations are reported.*



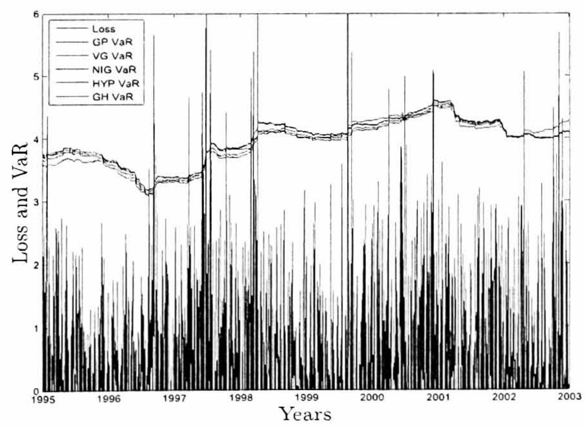


Figure 6.13: *Dynamic VaR for Extreme Value and Lévy Models: Nikkei225 case. A rolling window of four years is considered. The top panel is for* 95% *coverage and the bottom panel is for 99% coverage.*

		Stat.	1%CV	5%CV	10%CV	p-χ^2	p-MC	PV
GP	UC(95%)	1.23	6.63	3.84	2.70	0.267	0.227	5.53
	Ind(95%)	0.94	6.63	3.84	2.70	0.330	0.349	5.53
	CC(95%)	2.17	9.21	5.99	4.60	0.336	0.333	5.53
	UC(99%)	0.60	6.63	3.84	2.70	0.437	0.375	1.17
	Ind(99%)	0.59	6.63	3.84	2.70	0.441	0.319	1.17
	CC(99%)	1.19	9.21	5.99	4.60	0.549	0.499	1.17
VG	UC(95%)	0.67	6.63	3.84	2.70	0.412	0.424	5.39
	Ind(95%)	1.25	6.63	3.84	2.70	0.262	0.265	5.39
	CC(95%)	1.93	9.21	5.99	4.60	0.381	0.408	5.39
	UC(99%)	0.60	6.63	3.84	2.70	0.436	0.389	1.17
	Ind(99%)	0.59	6.63	3.84	2.70	0.441	0.287	1.17
	CC(99%)	1.19	9.21	5.99	4.60	0.549	0.492	1.17
NIG	UC(95%)	1.03	6.63	3.84	2.70	0.311	0.296	5.48
	Ind(95%)	1.04	6.63	3.84	2.70	0.306	0.296	5.48
	CC(95%)	2.07	9.21	5.99	4.60	0.355	0.361	5.48
	UC(99%)	0.32	6.63	3.84	2.70	0.569	0.503	1.12
	Ind(99%)	0.546	6.63	3.84	2.70	0.459	0.360	1.12
	CC(99%)	0.87	9.21	5.99	4.60	0.647	0.540	1.12
HYP	UC(95%)	1.02	6.63	3.84	2.70	0.311	0.310	5.48
	Ind(95%)	0.31	6.63	3.84	2.70	0.306	0.307	5.48
	CC(95%)	2.07	9.21	5.99	4.60	0.355	0.375	5.48
	UC(99%)	0.32	6.63	3.84	2.70	0.568	0.513	1.12
	Ind(99%)	0.546	6.63	3.84	2.70	0.459	0.379	1.12
	CC(99%)	0.87	9.21	5.99	4.60	0.647	0.560	1.12
GH	UC(95%)	1.23	6.63	3.84	2.70	0.367	0.287	5.53
	Ind(95%)	0.95	6.63	3.84	2.70	0.330	0.373	5.53
	CC(95%)	2.18	9.21	5.99	4.60	0.336	0.381	5.53
	UC(99%)	0.32	6.63	3.84	2.70	0.569	0.530	1.12
	Ind(99%)	0.546	6.63	3.84	2.70	0.459	0.356	1.12
	CC(99%)	0.87	9.21	5.99	4.60	0.647	0.573	1.12

Table 6.16: *Backtesting results for conditional and unconditional models: Nikkei225. UC stands for Unconditional Coverage,Ind stands for Independence Test, CC stands for Conditional Coverage. PV stands for proportion of VaR violation.p-values from both Chisquare and Monte-Carlo simulations are reported.*

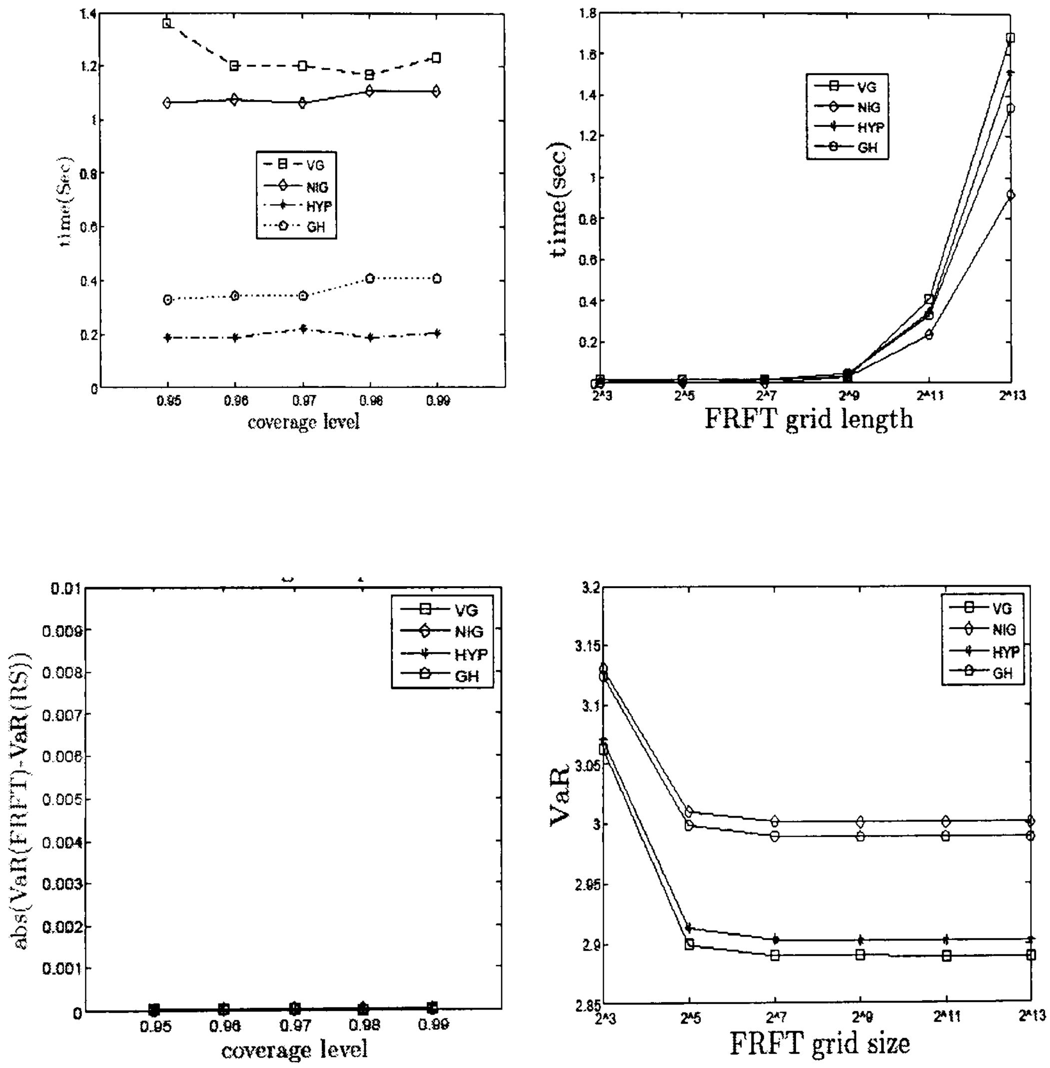


Figure 6.14: *The top panel shows the time-grid size trade-off in FRFT VaR computation. The figure on the right hand side shows that the time required to compute the FRFT VaR with different coverage levels remains similar. For this figure we consider FRFT grid size of* 2^{12}. *The bottom panel shows the performance-grid size trade-off. We consider grid size of* 2^9 *to obtain the figure on the right hand side which shows that root search(RS) and FRFT approach provide very similar estimates, though one has excellent computational superiority over other. For this illustration we used the parameters reported in Schoutens (2003)[102].*

Once the FRFT engine is ready it gives us a vector estimates of risk measure VaR– corresponding to a vector of coverage probabilities– in fraction of a second. Thus estimation of coherent risk measures become computationally feasible. This feasibility arises from the fact that we can now avoid estimation of ES based on equations of the form(6.20) and can directly use equation of the form(6.19). This way quantile by quantile root search involving complicated density expressions is totally ignored. Similarly estimation of SRM can now be efficiently accomplished using equation (6.26), avoiding equations of the form (6.27). We leave the estimation of all three risk measures using the FRFT approach for near future.

6.6.1 Risk Measures for GARCH-Lévy Dynamics Using FRFT

In discrete time GARCH area explicit modeling with non-normal innovations is still under-developed. Academics are still confined to models with student-t, GED type innovations–at the best–under various GARCH specifications. Consequently GARCH dynamics with explicit Lévy type innovations(VG, NIG,CGMY) are yet to be investigated for risk management. Further attractions of such dynamics in risk management are apparent from chapter5. First of all for each of our GARCH-Lévy dynamics we have analytic characteristic function. In first part of this chapter we implemented FRFT based quick estimation of risk measures which consider only characteristic function values as input. So with an explicit dynamic for GARCH-Lévy volatility characterization–which is available for each of our proposed dynamics in chapter5–we can use characteristic function of the form(5.151), for GARCH-NIG model e.g., to estimate risk measures in semi analytic fashion. An apparent limitation is, however, fathomed from the recursive relations embedded in characteristic functions, e.g. from relations in(5.152) which are embedded in (5.151) in case of GARCH-NIG model. Namely the terminal conditions on 'A' and 'B'[7] implies that structural parameters α_1 and β_0 will be ignored in predicting the risk measures one period ahead. This will be misleading. However for multi-period ahead forecast we can see the feasibility of estimating semi-analytic risk measures in GARCH-Lévy settings; which we expect will be a breakthrough.

In standard GARCH dynamics though innovations are often normal, the non-normality is not completely ignored. This is because structural GARCH parameters introduce skew-

[7] $A(T,T,u) = 0, \;\; B(T,T,u) = 0$ as required for the development of the relations in (5.152).

ness and kurtosis even when innovations are normal. Perignon et al(2008)[92] applied GARCH(1,1) dynamics to see the efficiency of VaR forecast in commercial banks. They observed that GARCH(1,1) VaR, with normal innovations, performed better than those based on Black-Scholes model. However they found it necessary to explore more rich characterization of return dynamics. The main difference our GARCH-Lévy dynamics exhibit with respect to usual GARCH(1,1) dynamics is due to the fact that we have conditional characterization of skewness and kurtosis which is not the case with GARCH(1,1) dynamics. This allows us to expect that theoretical values of risk measures from our model will be more appreciable predictions of reality even when there are few outlier like observations. In our future works we will consider the estimated volatility from GARCH-Lévy models and will use them to compute the characteristic function values to estimate the risk measures in semi-analytic fashion.

One can in principle considers the risk neutral parameters calibrated from option prices and use the market price of risk characterization and switch back to statistical parameters to investigate the risk measures. However in practice it is important that parameters are calibrated directly from historical returns instead of risk-neutralizing those calibrated from forward looking option prices. The difference between "principle" and "practice" arises from the fact that the information contents of stocks and options could be significantly different.

6.7 Discussion

Some important observations stand out when we compare the risk measures with their empirical counterparts. Though VaR and ES depend only on extreme quantiles, EV and Lévy models often fare similar to each other. Though these two risk measures consider only extreme quantiles of Lévy densities, nonetheless, the Lévy models are calibrated on the entire data. That is where EV differs from Lévy: in addition to considering only extreme quantiles for estimating risk measures it's calibration as well considers only extreme observations. Thus tail based risk measures often fare similar no matter whether models are calibrated using few tail observations alone or using all observations i.e. discarded quantiles do not carry any information in modeling extreme quantiles. See tables6.7-6.11. However for these tail based risk measures VaR and ES precision estimates bootstrapped SE's and bootstrapped CI's are overall more satisfactory for Lévy models than EV model. This implies that full density based Lévy estimates of VaR and ES are slightly more stable than their corresponding EV estimates. This is roughly true for all indices. See tables6.7-6.11.

In case of SRM, however, estimation considers all the quantiles with corresponding probabilities covering the entire spectrum. Again comparing with corresponding empirical values it stands out that estimates of SRM for Lévy models often outperform similar estimates for EV model. See tables6.7-6.11 To look deep into this fact we recall that SRM is not a tail based risk measure but EV parameters are calibrated using only extreme quantiles. Such a calibration could yield highly misleading quantiles, especially those falling far outside extreme tail. On the other hand Lévy models, considering entire data set in calibration, are expected to consistently generate the quantiles even when the quantiles fall far outside the extreme tail. These quantiles are in turn used in estimation of SRM.[8] In terms of bootstrapped precision estimates SE and CI for SRM; Lévy models are much worse compare to EV model. This is more so with the increase of risk aversion parameter "R". This feature is observed across all indices. The reason goes with the fact that to have the computation manageable we approximate the SRM integral with 100 slices only. Furthermore VaR for EV model is given in closed form and VaR of Lévy models are obtained through numerical search. These VaR numbers are directly feeded into SRM calculation. This adds to the loss of precision.

On the basis of VaR and its backtesting information model's static calibrations(on entire data) and dynamic calibrations(on a four year rolling window) often indicate similar preference. Proportion of violation's(PV) indicate that often the frequency of EV VaR violations are closest to the promised fraction α across indexes, except Nikkei225. For Nikkei225, on our observation period Japanese economy was booming thus the return density is skewed

[8]However we recall that we considerd only 100 slices in evaluating the integrals of SRM for Lévy models. But then this is uniformly considered for all the models. Furthermore Cotter and Dowd(2006)[39] reports the SRM for EV model considering one million slices(which is almost impossible to apply for Lévy models), see their table5 at page 3481. We can now compare SRM estimate of Lévy models with 100 slicing for numerical integration and see that they are often fairly comparable with their EV counterparts. Moreover we clearly notice that even with 100 slicing NIG and GH estimates systematically outperform EV estimates.

towards profit and not towards loss. Hence extreme loss events being non-substantial , EV and Lévy give similar backtesting evidence with comparable proportion of violation. Here we note that in static calibration for EV we used index-wise fixed threshold as studied and fixed by Cotter and Dowd(2006)[39] for the same data set. In case of dynamic calibration, however, we considerd fixed 30% of four years observations and didn't fix tails.

Backtesting results are often found unrelated across indexes. When for DAX all the test of hypothesis of backtesting often fail, for Nikkei225 all the test of hypotheses often pass; again stronger Japanese economy in our sample period implies no unusual fall in their asset prices, so standard hypotheses of backtesting are supposed to exhibit expected evidence. Other indices, however, provide mixed evidence with respect to different hypotheses of backtesting. Another evidence that stands out is that when to the 95% coverage Independence test fails(i.e. VaR violations are clustered and to some extent becomes a predictable event) , to the 99% coverage the test often pass; which is pretty understandable. It is rare to see that the test of conditional coverage hypothesis passed; it requires both the test of Unconditional and the test of Independence hypotheses to pass which is very rare across indices. The failure of Unconditional coverage hypothesis test is often justifiably implied by significantly deviated observed fraction of violation(PV) from promised fraction of violation(complement of coverage level).

The chi-square and simulated p-values are often very similar which further reinforce the effectiveness of all the test statistics under consideration.

6.8 Conclusion

We provide methodologies to estimate coherent risk measures ES and SRM for Lévy models. Considering the empirical evidence from major indices our study suggests that extreme spectral risk measure has some inconsistencies. EV model's calibration on restricted tail alone, provides poor estimate of quantiles outside the fixed tail which in turn yields the poor estimates of spectral risk measure itself. Lévy spectral risk measures , in contrast, are often found performing better than extreme spectral risk measure. This feature becomes increasingly apparent as the investors become more and more risk averse. However tail-based risk measures VaR and ES are often found similar for both tail based EV and Lévy models; observations discarded by EV but incorporated by Lévy models are not essentially making any improvement in the performance of tail-based risk measures. As a consequence we put forward a strong recommendation for the risk departments in financial establishments: if it is the strategy of the establishment to use EV model to measure the propensity of catastrophe, do not use SRM. Confine to the tail based coherent risk measure ES, if VaR deems non-informative; as it is often the case. However if attracted by the use of more realistic risk aversion functions the risk departments prefer using SRM, then do not use EV model to quantify the underlying risk and go for Lévy models.

For frequently used risk measure VaR (which is ,however, strongly criticized as not being "respectable") we present the backtesting results for Lévy and EV risk models.

One of the drawback of this estimation methodology is quite halting. Traditional root search method renders implementation of Lévy-coherent risk measures, ES and SRM, virtually infeasible. There arise the necessity of quick VaR estimation method especially for Lévy models. In this regard work in progress considers an innovative implementation using FRFT. This will facilitate practical implementation of ES and SRM for Lévy models. The estimates of ES and SRM for Lévy models reported in this chapter took huge computational times, so huge that it discourages us to report the computational times. Mathematically speaking FRFT approach skips computations delineated in equations(6.21)-(6.24) and (6.27)-(6.30). It just considers the characteristic function of the process and generates sufficiently fine cumulative distribution function(CDF) grids, from which a simple interpolation can give vector of VaR's corresponding to a vector of coverage probabilities. In a machine with sophisticated configuration the whole process can be accomplished in fraction of a second. The ultimate goal of this chapter is to capitalize on this characteristic function based FRFT VaR engine and use the conditional moment generating functions of GARCH-Lévy models in place of characteristic function to obtain the VaR in semi-analytic fashion. This is an ongoing work.

Conclusion

We conclude this thesis focusing on three important components, namely main contributions to the literature, limitations of the research and future works.

Main Contributions

Our first contribution is to have demonstrated how the standard Lévy-Kintchine formula may be interpreted as a series of shocks superimposed on a normal distribution, and how it can be used to value options using an illustrative example of a Variance-Gamma process. Using this derivation we have also been able to offer a correct solution to the misspecification in the Lévy measure for the VG model derived by Geman(2002)[62]. We also calibrated the VG and BS models considering weekly options data using both FFT and FRFT methods. We found that the FRFT is much faster than the FFT approach, saving 97-98% of calibration time. These findings have important implications for the calibration of options models and for options risk management in general. We observed that fractional parameter of FRFT generates some extra noise along the strike dimension and some systematic deviation along the maturity dimension which are, however, related with other parameters in not obvious ways.

Our second contribution is have investigated a number of available approaches to non-normality in option pricing as well as in option portfolio approximations. We compared the overall performance of a wide range of models – including Gram-Charlier, stochastic volatility, GARCH and Lévy models, as well as BS – using a common data set. We found a number of notable differences between them and in particular, that the BS and Gram-Charlier models often perform less well than the GARCH, stochastic volatility and Lévy models. Of these the stochastic volatility model performs robustly well and the Lévy models often perform well too. We have also shown that satisfactory estimates of these model's hedge ratio deltas and gammas can be obtained using traditional finite difference methods, and that these can be used to value portfolios of options. In this respect, our study extends the earlier work Britten and Schaefer(1999)[22], who only considered such problems in the context of Black-Scholes. We found that regardless of the model used, delta and delta-gamma approaches can yield inaccurate approximations of option portfolio values, especially in the face of large swings in the price of the underlying. These findings suggest that delta and delta-gamma approximations can be very misleading and reinforce the need for full-valuation methods instead. They also remind us - as if we didn't need yet another reminder! - that even the most (otherwise) sophisticated models can be very inaccurate during times of financial market turbulence. Furthermore we found that GARCH models perform comparably well– both in pricing and approximations–with other approaches to non-normality. This is the case even when the innovations of GARCH model are normal. Thus GARCH with non-normal innovations will presumably take the lead among different approaches.

Chapter5 includes our main contribution. In this, Lévy innovations are incorporated into the GARCH noise structure– replacing the normally distributed one– with a view towards analytic pricing of derivatives. Detailed mathematical developments provide complete GARCH characterizations with innovations from four frequently referred Lévy processes. Three of these innovations are from Lévy processes which are not subordinator thus having the potential to exhibit both positive and negative jumps. These are found to be mathematically cumbersome to deal with; as these require an approximation of volatility dynamics to uphold the analytic valuation methodology in GARCH-Lévy framework. Such innovations are from Lévy processes which are Brownian motion stochastically time changed by subordinators: VG–Brownian motion time changed by Gamma subordinator, NIG–Brownian motion time changed by inverse Gaussian subordinators, CGMY–Brownian motions time changed by tempered stable subordinators. The empirical part implements these three frequently refereed Lévy processes in highly volatile market; together with the implementation of only GARCH-NIG analytic option pricing model (CFG-NIG). Future empirical works will consider analytic GARCH option pricing with other innovations.

In addition to detailed theoretical developments, we also obtained some interesting empirical results. First of all our pilot survey shows in turbulent market most of the otherwise sophisticated models may perform poorly and models with time varying volatility can provide a better response to market turbulence. Moreover a strong indication is there to suggest that non-normal innovations can make a huge difference. Some studies observed this little earlier i.e. after Heston and Nandi's(2000)[70] model was proposed. But those studies were mostly based on Monte Carlo simulation to price options in GARCH model with non-normal innovations. GARCH-IG and GARCH-Gamma are two such models. Their appealing performance creates enough motivation to deal with "not-so-comfortable" mathematics in order to be free of reliance on simulations; this results in quick pricing and so makes it possible to consider options traded on long time windows.

Our empirical study considers options traded on a three-year window of turbulent market performance: January 2005 to December 2007. Considering options on different cross-sections we found that our semi-analytic GARCH-NIG model performs significantly well compare to Heston and Nandi's semi analytic GARCH-Normal model. Capitalizing on analytic valuation we are able to verify and compare GARCH-NIG model with GARCH-Normal and Lévy models on different information aggregation schemes. Overall we observed that the GARCH-NIG model performs better with an increasing amount of information aggregation.

The reason for CFG-NIG model's excellent performance basically comes from its volatility characterization and time-varying higher moment features. The separate characterizations to describe conditional evolution of skewness (equation (5.57)) and conditional evolution of kurtosis (equation (5.58)), help the model accommodate the cross-strike and cross-maturity features more consistently. That's where CFG-NIG basically differs from Heston and Nandi's(2000)[70] model. As reported by Bates(2003)[11] equation regarding continuous time SVJ (or SVJJ) model: "having jump components addresses moneyness biases, while having stochastic variables allow distributions to evolve stochastically over time". This is precisely the same for our CFG-NIG model as well.

The last chapter comprises the risk-management part of this research. In this chapter we provide methodologies to estimate coherent risk measures ES and SRM for Lévy models. For the first time in literature–to the best of our knowledge–we provide estimates of coherent risk measures ES and SRM for Lévy models, using most of the leading indices over the world, and compare them with corresponding estimates from EV model. Based on such estimation, empirical evidence from major indices suggests that extreme spectral risk measure has some inconsistencies. Namely, extreme value model's calibration based on a few tail observations provides a poor estimate of quantiles beyond the tail which in turn yields poor estimates of spectral risk measure itself. In contrast we noticed that Lévy spectral risk measures often perform better than the extreme spectral risk measure and this becomes increasingly apparent as the investors become more risk averse. We also present some backtesting results for Lévy and EV risk models.

The computational efficiency of the estimation routines in chapter6 could be significantly improved. The traditional way of estimating VaR makes estimation of ES and SRM practically infeasible with Levy processes. However, as we discussed, we could tackle this problem through an application of FRFT which helps us obtain a vector of VaR's corresponding to a vector of coverage levels in a fraction of a second. We, moreover, observed that it is possible to obtain similar estimates of VaR under both traditional root search and FRFT approaches by appropriately choosing FRFT parameters. But no matter how little the difference between two VaR estimates is, the errors will be aggregated when VaR's will be integrated on the tail for ES and along the entire spectrum for SRM. These will give us much faster FRFT estimates of ES and SRM with little price. That is after all there will be a trade-off between better estimation performance of the root search approach and much faster estimation obtained under the FRFT approach.

Limitations of the Research

Our study in chapter4 raises a pertinent question: *how robust is our preferential ordering of alternative models?*Though we did report the results using options traded on a particular day, we calibrate all eight models on various other days as well. The evidence which stands out is the ordering of approaches we report: stochastic volatility model(Heston(1993)[69] in this case) and GARCH volatility model(Heston and Nandi(2000)[70] in this case) always standing out as two best competing models; pure jump Lévy as second preferred approach(VG, NIG, CGMY in this case, however we did observe within this approach individual model's ordering is not robust); Gram-Charlier and Jump-diffusion models are found to be variable–on some days they are well comparable with Lévy models but on other days they are performing much worse than Lévy models. However we did not find any day when either of these two models outperformed Heston's (1993)[69] stochastic volatility model and/or the Heston and Nandi(2000)[70] GARCH volatility model; but on every single day both of the former outperform at least the benchmark Black-Scholes model. Of course this observation is contingent upon our primary focus: investors wish to choose between models using only the most minimal recent point-in-time data set.

The limitations of our main contribution in chapter5 are well noted. First of all, from a computational point of view, or more precisely from programming point of view, it is a real challenge to implement the model. The results we reported require huge programming concentrations; they also require sophisticated knowledge and patience to keep trying and checking for bugs until it was clear that all the programs are doing exactly what the derived mathematics expects them to do. Saying this it means that computational hassles might make the model less appealing at first instance but once functions are written (in any programming language) calculations become manageable in real time[9]. Regarding other limitations we must say that the explanation of model's true potentials could have been more precise and this would require more econometric oriented interpretations of our results. Moreover we must agree that other related facets of the empirical investigations, e.g. empirical volatility related features, need to be carried out in due course.

Possible Future Work

Our study in chapter4 can be extended in a number of ways. To date there are no systematic comparisons of option risk measures (such as VaR or Expected Shortfall) based on all eight models: it would be useful to compare these on common data sets encompassing both stable

[9]We used a machine with"Intel(R) Core(TM)2Duo CPU T5800" processor, 320 ROM and 3GB RAM.

and turbulent market conditions. Second, it would be useful to examine the performance of different numerical schemes to calculate the Greeks. Quick and accurate calculation of these would help in hedging and risk-managing the options involved.

Following the evidence from chapter5, we are looking forward to conduct empirical studies with other GARCH-Lévy dynamics, namely: GARCH-VG and GARCH-CGMY models. Moreover an innovation from a positive Lévy process (subordinator) is also mathematically detailed, GARCH-TS model, and we will consider empirical investigation with this GARCH-TS dynamics too.

Though SVJ (or SVJJ) models of continuous time have some analogy to these discrete time GARCH-Lévy models with jumps components, nonetheless they are premised on a rigid Markovian structure. Since CFG-NIG model incorporates non-Markovian time series properties through heteroskedasticity, presumably this will give CFG-NIG type models a preferential edge over SVJ (or SVJJ) model. We look forward to clarify this presumption in a future empirical work.

To draw meaningful conclusion regarding GARCH-Lévy models relative performance–w.r.t. different Lévy innovations–we need to consider other GARCH-Lévy models which are by now available in the literature, e.g. GARCH-IG, GARCH-Gamma etc. From our experience we can see that it will be a computationally challenging work to consider all such dynamics on common data sets covering smooth and turbulent market conditions but it will be worth considering. Of course this future work will be a follow up work only after other dynamics (CFG-VG, CFG-CGMY, CFG-TS) have been be implemented.

Regarding hoped for future engagement it is also our intention to implement the dynamics in the credit derivatives area. Initial feasibility arises from the fact that credit derivatives are like option but the details need to be figured out. We do not rule out the application of the dynamics in other derivative pricing as well, e.g. weather derivatives, commodity derivatives and agricultural derivatives.

From chapter6 we should see an article submitted soon. However another follow-up work from this chapter will include the immediate future work where we will consider FRFT estimation of ES and SRM for the Lévy models which we believe will be practically implementable when estimated with appropriate choice of FRFT parameters. We will do all the estimations using FRFT, parallel to what we have in chapter six. Then we will reveal the performance-speed trade-off for some illustrative cases.

Another ambitious direction is to use the numerical CDF of Lévy models in the peak-over-threshold (POT) theory, introducing Lévy-extreme value models in tail-based risk management. However it will be computationally burdensome and it is not clear whether that will compensate empirically.

Finally our GARCH-Lévy dynamics are ready to apply in risk management if we consider the usual Monte-Carlo approach to obtain VaR in a GARCH set up. Naturally, since conditional skewness and conditional kurtosis have time varying characterizations –with variation resulting from time varying volatility–the dynamics are expected to outperform those GARCH characterizations which do not have these features. However the possible breakthrough we are expecting to make is not through Monte Carlo approach; we are looking forward to capitalizing on the characteristic function based FRFT approach to analytically estimate the risk measures under GARCH-Lévy set up.

Bibliography

[1] Acerbi, C.(2004) *Coherent representation of subjective risk aversion.* pp 147-207 in G.Szegö (ed.) *Risk Measures for the 21st Century.* Chichester: John Wiley and Sons, Ltd.

[2] Applebeum, David (2004). *Lévy Processes and Stochastic Calculus.* Cambridge University Press: Cambridge, UK.

[3] Anderson, T.W.,Darling,D.A.(1952).Asymptotic Theory of Certain 'Goodness of Fit' Criteria Based on Stochastic Processes.*The Annals of Mathematical Statistics.* 23(2),193-212.

[4] Anderson, T.W.,Darling,D.A.(1954).A test of Goodness of Fit.*Journal of American Statistical Society.* 49(268),765-769.

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

[6] Bailey, D., Swartztrauber, P.(1991). The fractional fourier transform and applications. *SIAM Review 33.* 3, 389-404.

[7] Bakshi, G.,C. Cao, Z., Chen.(1997). Empirical Performance of Alternative Option Pricing Models.*Journal of Finance.*52, 2003-2049.

[8] Bakshi, G., Cao,C., Chen, Z.(2000). Do call prices and the underlying stock always move in the same direction?*Review of Financial Studies* 13, 549-584.

[9] Backus, D., Foresi, S.,Li,K., Wu, L.(1997). Accounting for Biases in Black-Scholes. *Manuscript, The Stern School of Business ,New York University.*

[10] Bates, David, S.(2000). Post-'87 Crash fears in the S&P500 futures option market. *Journal of Econometrics* 94, 181-238.

[11] Bates, David S.(2003). Empirical option pricing: a retrospection. *Journal of Econometrics* 116, 387-404.

[12] Barndorff-Nielsen, O.E.(1995). Normal Inverse Gaussina distributions and the modeling of stock returns. *Research report no 300,Department of Theoretical Statistics, Aarhus University.* A 353, 401-419.

[13] Bera, A. K., Higgins, M. L.(1993). ARCH Models: Properties, Estimation and Testing. *Journal of Economic Surveys.* Vol. 7,No. 4, 307-366.

[14] Berkes, I., Horvath, L., Kokoszka, P.(2003). GARCH processes: structure and estimation. *Bernoulli.* 9 201-227.

[15] Bertoin, Jean (1996). *Lévy Processes.* Cambridge University Press: Cambridge, UK.

[16] Bellini, F.,Mercuri, L.(2007).Option pricing in GARCH models. *Working Paper.*

[17] Billingsley, P.(1999). Convergence of Probability Measures. 2nd edition. *Wiley, New York.*

[18] Bingham, N.H., Kiesel, R.(2001). Modeling asset returns with hyperbolic distributions. In *Return Distributions on finance* 1-20. Butterworth-Heinemann.

[19] Black, F., Scholes, M.(1973).The pricing of options and corporate liabilities. *Journal of Plitical Economics.* 81, 637-659.

[20] Bollerslev, T. (1986). Generalized autoregressive conditional heteroeskedasticity. *Journal of Econometrics* 31 307-327.

[21] Bollerslev, T.(1987). A conditional heteroskedastic time series model for speculative prices and rates of return. *Review of Economic and Statistics.* 69, 542-547.

[22] Britten-Jones, M., Schaefer, S.(1999). Non-Linear Value at Risk. *European Finance Review.* 2, 161-187.

[23] Buhlmann, H.Delbaen, F., Embrechts, P.,Shiryaev, A.N. (1996). No arbitrage, change of measure and conditional Esscher transform. *CWI Quarterly.* 9 291-317.

[24] Bulm, J.R. and Rosenblat, M. (1959). *On the structure of infinitely divisible distributions.* Pacific J. Math. Vol. 9, 1 , 1-7.

[25] Byan S.J., Min B.(2010). Conditional Volatility and the GARCH Option Pricing Model with Non-normal Innovations. *Working Paper* KAIST Business School, Seoul, Korea.

[26] Carr, P., Geman, H.,Madan, D.B.,Yor, M.(2002) The Fine Structure of Asset Returns: An Empirical Investigation. *Journal of Business.*75(2), 305-332.

[27] Carr,P., Madan, D.(1999).Option valuation using the fast fourier transform. *Journal of Computational Finance.* 2, 61-73.

[28] Campbell, J., Hentschel, L.(1992). No news is good news : An asymetric model of changing volatility in stock returns. . *Journal of Financial Economics* 31 281-318.

[29] Chourdakis, K.M. (2005). Option Pricing using the fractional FFT. *Journal of Computational Finance.* 8, 1-18.

[30] Chourdakis, K.M. (2008). Financial Engineering : A Brief Introduction Using the Matlab System. *Working Book.* 8, 1-18.

[31] Christoffersen, P., K. Jacobs.(2004).Which GARCH model for option valuation? *Management Sci.* 44 1218-1233.

[32] Christoffersen, P., Elkamhi,R., Feunou, B. , Jacobs, K. (2010).Option Valuation with Conditional Heteroskedasticity and Nonnormality. *Review of Financial Studies.* 23(5) 2139-2183.

[33] Christoffersen, P.,Heston, S.L., Jacobs, C.(2006).Option valuation with conditional skewness. *Journal of Econometrics.* 131 253-284.

[34] Christoffersen, P.(2003). *Element of Financial Risk Management.* Academic Press.

[35] Chernov, M., Gallant, A.R., Ghysels, E., Tauchen, G.(2003). Alternative models for stock price dynamics. *Journal of Econometric.* 116.

[36] Chernobai, A., Rachev, S.,Fabozzi, F.(2005).Composite Goodensss-of-Fit Tests for Left-Truncated Loss Samples. *Working Paper, Department of Statistics and Applied Probability. University of California, Santa Barbara.*

[37] Clark, P. (1973). A Subordinated Stochastic Process with Finite Variance for Speculative Prices. *Econometrica.* 41, 135-155.

[38] Cont, Rama and Tankov, Peter (2004). *Financial Modelling With Jump Processes.* Chapman & Hall/CRC Financial Mathematics Series: London.

[39] Cotter, J., Dowd, K.(2006). Extreme spectral risk measures : An application to futures clearinghouse margin requirements. *Journal of Banking& Finance.*30, 3469-3485.

[40] Cox, J.,C. Rubinstein, M.(1985). *Options Markets.* Prentice Hall Englewood Cliffs, NJ.

[41] Derman, E., Kani, I.(1994).Riding on a Smile. *Risk Magazine.* Vol.7,No.2,February,32-40.

[42] Dowd, K.(2005). *Measuring Market Risk.* John Wiley & sons. Ltd.

[43] Dowd, K., Cairns, A.J.G., Blake, D., Coughlan, G.D., Epstein, D., Khalaf-Allah, M.(2010).Backtesting Stochastic Mortality Models: An Ex Post Evaluation of Multiperiod-Ahead Density Forecasts. *North American Actuarial Journal.* 14(3), 281-298.

[44] Dowd, K., Cotter, J., Sorwar, G.(2008). Spectral Risk Measures: Proeprties and Limitations. *Jouranl of Financial Services and Research.* 34, 61-75.

[45] Duan, J.-C. (1995). The GARCH option pricing model. *Mathematical Finance* 5 13-32.

[46] Duan, J.-C., P. Ritchken, Z. Sun.(2004). Jump starting GARCH: Pricing and hedging options with jumps in the returns and volatilities. *Manuscript, University of Toronto, Toronto, Ontario, Canada.*

[47] Duan, J., C.(1996).A unified theory of option pricing under stochastic volatility- from GARCH to diffusion. *Working paper.* Hong-Kong university of science and technology.

[48] Duffy, D., J.(2006) *Finite Difference Methods in Financial Engineering: A Partial Differential Equation Approach* John Wiley & sons. Ltd.

[49] Dupire, B.(1994).Pricing with a smile. *RISK.* 7, 18-20.

[50] Eberlein, E., Prause, K.(1998). The Generalized Hyperbolic model: financial derivatives and risk measures. *FDM preprint 56, University of Freiburg.*

[51] Eberlein, E., Hammerstein, E. A.(2002). The Generalized Hyperbolic and Inverse Gaussian Distributions: limiting cases and approximation of processes. *FDM preprint 80, University of Freiburg.*

[52] Eberlein, E., Keller, U.(1995). Hyperbolic distributions in finance. *Bernoulli* 1, 281-299.

[53] Embrechts,P.,Kluppelberg, C.,Mikosch, T.(1997).Modelling Extremal Events for Insurance and Finance.*Sringer-Verlag,Berlin.*

[54] Engle, R.(1982). Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflations.. *Econometrica* 50, 987-1008.

[55] Fama, E. F.(1965). The Behavior of Stock Market Returns. *Journal of Business.* 38, 34-105.

[56] Fama, E. F.(1976). Foundations of Finance. Basic Books.

[57] Feller, W. (1971). *An introduction to probability theory and its application.* Vol. 2, 2nd ed., Wiley, New york.

[58] French. K., G. W. Schwert, R. Stambaugh. (1987). Expected stock returns and volatility. *journal of Financial Economics* 19 3-30.

[59] Fusai. G, Roncoroni. A.(2008) *Implementing Models in Quantitative Finance : Methods and Cases.* Springer Finance.

[60] Fusai, G., Meucci, A.(2008) Pricing Discretely Monitored Asian Options Under Lévy Processes. *Journal of Banking & Finance* 32, 2076-2088.

[61] Garcia, R., E. Renault. (1998). A note on hedging in ARCH and stochastic volatility option pricing model. *Mathematical Finance* 8 153-161.

[62] Geman, H.(2002). Pure Jump Lévy processes for asset price modelling. *Journal of Banking & Finance.* 26, 1297-1316.

[63] Geman, H., Madan, D., Yor, M.(2001).Time changes for Lévy processes. *Mathematical Finance.* 11, 79-96.

[64] George, L.(2001). An introduction to GARCH models in finance. *Financial Engineering News.* 22.

[65] Gerber, H.U., Shiu, E.S.W.(1994). Option Pricing by Esscher Transforms. *Transactions of The Society of Actuaries.* 46, 99-191.

[66] Gil-Pelaez, J.(1951).Note on the Inversion Theorem.*Biometrika.* Vol.38, No 3/4,481-482.

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

[68] Gut, Allun (2005). *Probability : A Graduate Course.* Springer Texts in Statistics.

[69] Heston, S.(1993).A closed form solutions for options with stochastic volatility with applications to bond and currency options. *Review Financial Studies.* 6, 327-343.

[70] Heston, S.L., Nandi, S.(2000).A closed form GARCH option valuation model. *Review of Financial Studies.* 13, 585-625.

[71] Hull, J.,White, A.(1987). The pricing of options on assests with stochastic volatilities. *Journal of Finance.* XLII, 281-300.

[72] Huynh, H.,T., Lai, V.,S., Soumare, I.(2008).*Stochastic simulation and applications in finance with MATAB programs.* Wiley Finance.

[73] Jondeau, E., Poon, S.H., Rockinger, M.(2007). *Financial Modeling Under Non-Gaussian Distributions.* Springer Finance.

[74] Kim, Y.S., Rachev, S.T., Chung, D.M.(2006).The modified tempered stable distribution, GARCH-models and option pricing. *Technical report, School of economics and Business Engineering. University of Karlsruhe.*

[75] Kou, S.(2002).A jump diffusion model for option pricing.*Management Science.* 48, 1086-1101.

[76] Kyprianou, A.(2006).*Introductory Lectures on Fluctuations of Lévy Processes with Applications.*Springer

[77] Liuren, Wu.(2006). Modelling Financial Security Returns Using Lévy Processes. *Working paper, Zicklin School of Business, City University of New York.*

[78] Lee, R.,W.(2004). Option Pricing by Transform Methods: Extensions, Unification, and Error Control.*Journal of Computational Finance.* 7(3), 51-86.

[79] Madan, D.B., Senata, E.(1990)The veriance gamma model for share market returns. *Journal of Business* 63(4), 511-524.

[80] Madan,D., Carra, P., Chang, E.(1998). The variance Gamma Process and Option Pricing Model. *European Finance Review.* 2, 79-105.

[81] Major, Péter. Limit theorems and infinitely divisible distributions. *The mathematical Institute of Hungarian Academy of Sciences.* (Available on his website, http://www.renyi.hu/ major/probability/divisible.html)

[82] Mandelbrot, B.(1963). The variation of certain speculative prices. *Journal of Business.* 36, 394-419.

[83] McNeil, A., Frey, R., Embrechts, P.(2005). *Quantitative Risk Management: Concepts, Techniques and Tools.* Princton Series in Finance.

[84] Miranda, M.,J., Fackler, P.,L. (2002) *Applied Computational Economics and Finance.* MIT Press, Cambridge MA and London.

[85] Merton,R.(1976).Option pricing when underlying stock retunrs are discontinuous. *Journal of financial economics.* 3, 125-144.

[86] Mozumder, Md. S.U.(2007). Estimation of the Lévy Measure for the Aggregate Claims Process in Risk Theory. *MS Thesis, Concordia University* Montreal, Canada.

[87] Nelson, Daniel, B.(1991).Conditional Heteroeskedasticity in Asset Returns:A new Approach.*Econometrica.* Vol.59, 347-370.

[88] Nelson, D. B., Cao, C. Q.(1992).Inequality Constraints in the Univariate GARCH Model. *Journal of Business and Economic Statistics.* 10(2), 229-235.

[89] Ornthanalai,C.(2010). A New Class of Asset Pricing Models with Lévy Processes: Theory and Applications. *Working Paper, Georgia Institute of Technology.*

[90] Pagan, A., G. W. Schwert. (1990). Alternative models for conditional stock volatility. Journal Econometrics 45, 267-290.

[91] Prause, K.(1999). The Generalized Hyperbolic Model: Estimation, financial derivatives and risk measures. *PhD Thesis, University of Freiburg.*

[92] Perignon, C., Deng, Z.,Y., Wang, Z.,J. (2008). Do banks overstate their Value-at-risk? Journal of Banking & Finance. 32 784-793.

[93] Poon, S., H.(2005) *A Practical Guide for Forecasting Financial Market Volatility.* John Wiley & Sons Ltd.

[94] Ramezani, C.A., Zeng, Y.(1999). Maximum likelihood estimation of asymmetric jump-diffusion processes: Application to security prices. *Working paper, Department of Statistics, University of Winsconsin, Madison, WI.*

[95] Rubinstein, M.(1976).Nonparametric Tests of Alternative Option Pricing Models Using All Reported Trades and Quotes on the 30 Most Active Option Classes from August 23,1976 through August 31,1978.*Journal of Finance.* 40, 455-480.

[96] Rubinstein, M.(1994).Implied Binomial Trees.*Journal of Finance.* 69, 771-818.

[97] Rouah, F., D., Vainberg, G.(2007). *Option Pricing Models & Volatility.* Wiley Finance

[98] Scott, L.(1987). Option Pricing When the Variance Changes Randomly: Theory, Estimators and Applications. *Journal of Financial and Quantitative Analysis.* 22(4) 419-138.

[99] Roussas, George, G.(2005). *An Introduction to Measure-Theoretical Probability.* Elsevier Academic Press.

[100] Sato, Ken-Iti (1999). *Lévy Processes and Infinitely Divisible Distributions.* Cambridge University Press: Cambridge, UK.

[101] Schoutens, W., Cariboni, J.(2009). *Lévy Processes in Credit Risk.* Wiley Finance.

[102] Schoutens, W. (2003). *Lévy Processes in Finance:Pricing Financial Derivatives.* John Wiley & sons. Ltd.

[103] Seneta, E. (2004). Fitting the Variance Gamma Model to Financial Data. *Journal of Applied Probability.* 41A, 77-87.

[104] Shreve, S.(2004). *Stochastic Calculus for Finance-II.* Springer Finance.

[105] Siu, T.K., Tong, H., Yang, H.(2004).On pricing Derivatives under Garch models: A dynamic Gerber-Shiu approach.*North American Actuarial Journal.* 8, 17-31.

[106] Shiryaev, A. N.(1999).*Essentials of Stochastic Finance : Facts, Models, Theory.* World Scientific.

[107] Sheikh, A.M.(1991). Transaction Data Test of S&P 100 Call Option Pricing. *Journal of Financial and Quantitative Analysis.* 26, 459-475.

[108] Stein, E., Stein, J.(1991). Stock Price Distribution with Stchastic Volatility: An Analytic Approach. *Review of Financial Studies.* 4, 727-752.

[109] Stock, J. H. (1998). Estimating Continuous-Time Processes Subject to Time Deformation.*Journal of the American Statistical Association.* 83, 77-85.

[110] Sorwar, G., Dowd, K.(2010). Estimating financial risk measures for options. *Journal of Banking & Finance.*34(8) 1982-1992.

[111] Tweedie, M.(1984). An index that distinguishes beteen some important exponential families. *In: Ghosh,J., Roy, J.(Eds). Statistics: Applications and New Directions.* In: Proc. Indian Statistical Institute Golden Jubilee International Conference. Calcutta, 579-604.

[112] Watson, G.N.(1944).*A Treatise on the Theory of Bessel Functions.* Cambridge University Press, Cambridge.

[113] Whitt, W., Abate, J.(1992).The Fourier-series method for inverting transforms of probability distributions. *Queueing Syst. Theory Appl.* Vol.2,Np 1-2, 5-88.

[114] Wiggins, J.B.(1987).Option values under stochastic volatility: theory and empirical estimates. *Journal of Financial Economics* 19, 351-377.

[115] Zolotarev, V. M. (1986). *One-Dimensional Stable Distributions.* American Mathematical Society.

[116] Zolotarev, V. M. (1966). On representation of stable laws by integrals. *Selected Translations in Mathematical Statistics and Probability.* 6, 84-88.