A discrete time approach to option pricing

Adam Aleksander Majewski

Scuola Normale Superiore

A thesis submitted for a degree

Doctor of Philosophy

Pisa, 2015

supervisors: professor Giacomo Bormetti

professor Fulvio Corsi

professor Stefano Marmi

Contents

Acknowledgments

First and foremost I want to thank my advisors. It has been an honour to be the first Ph.D. student of Giacomo Bormetti. I have gained many skills and a lot of knowledge while working under his supervision. This dissertation would not have been possible without his invaluable guidance and training.

I would also like to thank my advisor Fulvio Corsi for his suggestions, ideas and productive discussions. Working with him was a truly inspiring and fruitful experience. Moreover I would like to thank Fulvio and his wife Giulia for wonderful hospitality in Antignano, where we could develop our research in the beautiful scenery of Tuscan seaside.

I wish to express my deepest gratitude to my advisor and director of PhD programme Stefano Marmi for his guidance. I appreciate all his contributions of time, ideas, and funding to make my Ph.D studies productive and efficient.

Additionally, I am thankful to Fabrizio Lillo for all lectures and discussions on quantitative finance. I am also grateful to my dear colleague and co-author of articles Dario Alitab for helpful comments and fruitful discussions. Moreover I am thankful to professors, colleagues and all incredible people that I met during my PhD programme at Scuola Normale Superiore in Pisa.

Last but not least I would like to thank my family for all the support.

List of publications

Published articles:

1. Witold Bołt, Adam A. Majewski and Tomasz Szarek, *An invariance principle for the law of the iterated logarithm for some Markov chains*, Studia Math. 212, pp. 41-53, (2012), (Bołt et al., 2012)

2. Adam A. Majewski, Giacomo Bormetti and Fulvio Corsi, *Smile from the Past: A General Option Pricing Framework with Multiple Volatility and Leverage Components*, Journal of Econometrics 187, pp. 521-531, (2015), (Majewski et al., 2015)

Working papers available online:

3. Giacomo Bormetti, Fulvio Corsi and Adam A. Majewski, *Term structure of variance risk premium in multi-component GARCH models*, (Bormetti et al., 2015)

4. Dario Alitab, Giacomo Bormetti, Fulvio Corsi and Adam A. Majewski, *A Jump and Smile Ride: Continuous and Jump Variance Risk Premia in Option Pricing*, (Alitab et al., 2015)

Work in progress and collaborations:

5. Fabrizio Lillo, Adam A. Majewski, and Paris Pennasi, *Causalities between herding and liquidity in FX market*

6. Dario Alitab, Giacomo Bormetti, Fulvio Corsi and Adam A. Majewski, *Term structure of jump risk premium*

Introduction

An European option is a financial contract which gives the owner the right to buy or sell an underlying asset at a strike price on a maturity date. Option pricing theory tries to understand what is a fair price for such a contract. Option price models consist of two ingredients: dynamics of the price of the underlying asset under a physical measure and a pricing mechanism which is described by a stochastic discounting. Modelling the dynamics of asset price should incorporate the well documented fact that return variance is stochastic and to improve the model's performance it is necessary to have multi-component structure in volatility. Concerning the second ingredient of option pricing model, stochastic discounting should be multi-dimensional, in particular it should take into account variance risk premium.

Stochastic volatility models were introduced to reproduce well-established stylized facts like volatility smile and negative correlation of returns and volatility. Despite many successful applications, stochastic volatility models in continuous and discrete time exhibit serious problems with fitting strike profile and term structure of implied volatility surface, especially for far in-the-money and out-of-the-money options. In order to overcome this problem volatility models should incorporate heterogeneity of agents acting in the market. Investors with different time horizons have different impact on instantaneous volatility and as a consequence a single factor of volatility, running on a single time scale, is simply not sufficient for describing the dynamics of the volatility process. This argument has been empirically confirmed (Müller et al., 1997) and has led to the development of models with multi-component volatility structure, where each component of volatility corresponds to different time scale.

The necessity of taking into account variance risk premium stems from stochastic nature of volatility. Since the future level of return variance is a source of uncertainty, it is natural to assume that investor will demand a premium for bearing that risk. Variance risk premium is equal to a compensation that a representative investor is demanding for investing in an asset with unknown future return variance and it has a huge impact on the form and the properties of the pricing kernel in the economy. Moreover, incorporating variance risk premium in the model results in the so called 'U-shape' log ratio between the risk-neutral and physical densities which corresponds to the one observed in the market data.

Due primarily to mathematical tractability, the literature on option pricing traditionally has been dominated by continuous time processes (for example Black and Scholes (1973), Merton (1976), Heston (1993) and Bates (1996)). On the other hand, models for asset dynamics under the physical measure $\mathbb{P}$ have primarily been developed in discrete time. The time-varying volatility models of the ARCH-GARCH families (Engle, 1982; Bollerslev, 1996; Glosten et al., 1993; Nelson, 1991) have led the field in estimating and predicting the volatility dynamics. Another well-established discrete time volatility modelling approach is the so called Realized Volatility (RV) approach which provides a precise nonparametric measure of daily volatility (i.e., making it observable) leading to simplicity in model estimation and superior forecasting performance. Discrete time models present the important advantage of being easily filtered and estimated even in the presence of complex dynamical features such as long memory, multiple components and asymmetric effects, which turns out to be crucial in improving volatility forecast and option pricing performances. However, in the current literature, the analytical tractability of discrete time option pricing models is guaranteed only for rather specific types of models and pricing kernels.

The goal of the thesis is to propose a very general and fully analytical option pricing framework encompassing a wide class of discrete time models featuring multiple components structure in both volatility and leverage (the mechanism producing the asymmetric impact of positive and negative past returns on future volatility) and a flexible pricing kernel with multiple risk premia. We propose a framework general enough to include either GARCH-type volatility, Realized Volatility or a combination of the two. Moreover, we apply multi-dimensional pricing kernel, taking into account various risk premia and obtaining semi-closed form solutions for option prices.

The class of processes nested within our option pricing framework are affine processes in state variables, which are log-returns, volatility and leverage components. For such a class of processes we are able to derive the moment generating function of log-returns. Moreover, applying exponential-affine stochastic discount factor, often called Esscher transform, we are able to characterise the formal change of measure, write no-arbitrage condition and moment generating function under risk-neutral measure. One of our main contributions is generalization: our framework embraces several different option pricing models considered in the literature. The other important novelty of our approach is multi-dimensionality: we are considering both multi-dimensional affine processes (multi-component structure in volatility and leverage) and multi-dimensional Esscher transform with each component being related to a premium for a different risk.

Exploiting our general framework we propose three new option pricing models with original dynamics under physical measure. We also reconsider the CGARCH model of Christoffersen et al. (2008) by applying two-dimensional pricing kernel. Since our framework guarantees existence of semi-closed form formulas for option prices, the option pricing methodology is fast and efficient in implementation. In addition, by applying family of our fully analytically models with multi-component structure in volatility on a large sample of Standard and Poor 500 Index

(S&P 500) options, we show that our models improve pricing out-of-the-money (OTM) options compared to existing benchmarks. The strength of our framework is highlighted by a proposal of a model being a combination of realized and latent volatility approaches which gives superior option pricing performance.

To provide more accurate description of financial markets one has to take into account strong discontinuities, so called jumps, which are observed even in the most liquid financial markets. Various studies has provided statistical confirmation of theirs existence and several asset pricing models allowing presence of jumps were proposed (see Maheu and McCurdy (2004), Duan et al. (2006) and Christoffersen et al. (2010) for models with jumps in returns and see Eraker et al. (2003), Eraker (2004), and Broadie et al. (2007) for jumps in volatility). In this thesis we will consider a model with jumps in volatility.

Recent financial literature has devoted much attention to the measurement of variance risk premium. Carr and Wu (2009); Bollerslev and Todorov (2011) and others provide model-free methodologies of estimating a single maturity variance risk premium. Further analysis was devoted to decomposition of variance risk premium into continuous and jump component (Du and Kapadia, 2012; Bollerslev et al., 2014). Model based measurement of variance risk premium has been proposed by Wang and Eraker (2015). Beyond all mentioned references which consider single maturity VRP, some studies were dedicated to the whole term structure of variance risk premia (Mueller et al., 2013; Ait-Sahalia et al., 2015).

In the thesis we propose a dynamic measure of the VRP implied by a multi-component GARCH model with a multi-dimensional stochastic discount factor. While most of the studies focus almost exclusively on single maturity, the use of an analytically tractable parametric model allow us to compute risk premia over different maturity recovering the whole term structure of VRP. Application of multiple components structure in volatility reproduces a realistic family of term structures of variance risk premium including the empirically observed hump-shaped curves (see Egloff et al. (2010)). Contrary to the majority of research papers which analyze variance swap market, we compute the VRP term structure extracting the information contained in the stock and option prices time series. Due to market segmentation our VRP term structure might contain different information with respect to the term structure observed in the variance swap market. The presence of relevant information contained in the VRP term structure extracted from equity and option data is confirmed by the final empirical analysis which identifies the slope of VRP as a significant predictor of future stock market returns.

The thesis is divided in five chapters. We start the first chapter with a short introduction to asset pricing theory, where we summarise the continuous and discrete time approaches to asset prices modelling and we explain how Stochastic Discount Factor arises in pricing theory. In the second chapter we introduce the general framework for dynamics under physical measure that satisfies certain affine property and possesses multi-component structure in volatility. In the end we motivate why a multi-dimensional Esscher transform is a good choice for an Stochastic Discount Factor by deriving it from Pareto optimal allocation problem.

In the third chapter we introduce new models nested in general option pricing framework. We start with an extension of HARG-RV model with heterogenous and analytically tractable leverage structure called LHARG-RV. Then we take a a CGARCH model of Christoffersen et al. (2008) and we apply a new change of measure to obtain new dynamics under risk-neutral measure. We propose a model being mixture of LHARG-RV and GARCH approach acronymed GARCH-LHARG-RV and finally we consider an extension of LHARG-RV with jumps called JLHARG-RV.

In the forth chapter we present the applied procedure of realized variance measurement. Then we describe the methodology and results of models parameters' estimation. Next we introduce the option pricing procedure and we give the results of Stochastic Discount Factors's parameters calibration on option prices. We finish the chapter with an empirical assessment of proposed option pricing models.

In the last chapter we focus on variance risk premium - we derive a formula for variance risk premium implied by k-CGARCH model. Then we justify application of CGARCH model implied measurement of variance risk premium by showing that it generates a realistic family of shapes of term structure of variance swap rate, contrary to single component volatility model. Motivated by the significance of variance risk premium in asset pricing, we propose an original and efficient methodology for estimating the time evolution of the term structure of variance risk premium and we show its predictive power in explaining stock market excess returns.

The second chapter where general option pricing framework is introduced and the Section 3.1 on LHARG model are based on Majewski et al. (2015). The part of the thesis devoted to CGARCH model with two-dimensional Esscher transform (Section 3.2) and the part concerned on variance risk premium (Chapter 5) are based on Bormetti et al. (2015). The Section 3.4 is based on Alitab et al. (2015). A generalisation of result in Bühlmann et al. (1998) to multi-dimensional case (Theorem 10 in the thesis is showing that multi-dimensional Esscher transform ensures Pareto equilibrium) and everything about GARCH-LHARG-RV model (Section 3.3 and empirical results in Chapter 4) have not been published anywhere but in this thesis.

Chapter 1

Review of asset pricing theory

1.1 Review of price dynamics models in continuous and discrete time

Asset pricing is determined by three components: probabilistic description of future states of economy, attitude towards certain risks and payoff structure. While the last one is specified in a contract,[1] the possible outcomes in economy and risk discounting have to be modelled. In this section we shortly review the history of modelling the time evolution of prices in financial markets and in the following section we describe the fundamentals of risk discounting.

The history of financial mathematics begins with the PhD thesis of Louis Bachelier (1900) titled *Théorie de la spéculation* in which he proposes to model assets price with Brownian motion. Among many original insight of Bachelier the two most striking are the first mathematical description of Brownian motion and the concept of martingale. Bachelier derived in his thesis the distribution function of Wiener process linking it mathematically with the diffusion equation and he did it 5 years before famous paper of Albert Einstein (1905) where a partial differential

[1] For example, in the case of European call option the payoff is specified by function $f(S_T) = \max(S_T - K, 0)$, where S_T is a price of the underlying asset at the time of maturity of the option T and K is called strike of an option.

equation governing Brownian motion is derived. Moreover when providing the price of a barrier option (an option which depends on whether the share price crosses a given threshold) Bachelier has already realised that it must be computed under a probability measure which we call today martingale measure, namely a measure under which the expected profit of a speculator is zero.

The major drawback of Bachelier modelling approach is that Brownian motion can generate negative values while the price of an asset cannot. For this reason Paul Samuelson (1965) proposes to replace Brownian motion with geometric Brownian motion which is a stochastic process with a log-normal distribution. The next big breakthrough in financial mathematics is the paper by Black and Scholes (1973) in which they derive the closed-form European call option price formula assuming that the asset price dynamics is given by geometric Brownian motion. Applying Itô lemma to payoff function of European call option and to dynamics of the asset price Black and Scholes obtained a stochastic differential equation describing the evolution of option's price. Then assuming that risk preferences of agents have been neutralised, the drift of the price process normalized by the numeraire has to be equal zero and this condition is written as a partial differential equation. By a transformation of variables PDE becomes a heat equation which has a well-known solution.[2]

Black-Scholes model due to its simplicity and tractability gained so much popularity that it became a market standard of quoting options. When a trader looks at her screen instead of seeing option prices she would see implied volatilities - the volatility parameter in the diffusion equation of Black-Scholes model that makes model option price match the current market option price. Obviously if Black-Scholes assumption of constant volatility would be satisfied trader should observe the same implied volatilities for all strikes and maturities. However, in today's reality[3] the market implied volatility surface is far from being flat. Indeed, plotting

[2]Originally Black and Scholes have derived the PDE describing the option price by a hedging argument.

[3]This behaviour of markets became extremely evident after the Black monday (market crash of October 19th, 1987). Before this event market implied volatility surfaces were much flatter, close to Black-Scholes World.

implied volatility against different strikes trader observes a parabolic shape, a deviation from Black-Scholes World resembling a smile when it is symmetric or a smirk otherwise. Inability of reproducing a volatility smile/smirk is considered as the main limitation of Black-Scholes model.

In the end of previous century, there have been two major approaches of introducing the smile in option pricing model developed. The first one is by allowing jumps in the dynamics of the asset price. Consequently the dynamics of a underlying asset's price is generalised from Brownian motion to a Lévy process and the distribution of log-returns admits skewness and non-zero excess kurtosis. Examples of modelling the price with jump-diffusion process are Merton (1976); Bates (1996); Geman et al. (2001); Kou (2002).[4] Second way of introducing volatility smile is by allowing a time-varying volatility in diffusion equation. Volatility can become a deterministic function of price, like it is assumed in local volatility models (Dupire, 1994) or it can be a stochastic process itself, like it is assumed in stochastic volatility models. One of the first and most celebrated stochastic volatility model is Heston (1993) where the dynamics of price follows a diffusion process with volatility following mean-reverting process called Cox-Ingersoll-Ross process (Cox et al., 1985).[5]

The majority of mentioned option pricing models belong to the family of affine processes. Roughly speaking, a stochastic process is called affine if the logarithm of characteristic function of its transition distribution is affine with respect to initial state. Mathematical properties of affine processes together with theirs financial application to option pricing, credit risk and interest rates modelling can be found in seminal papers Duffie et al. (2000) and Duffie et al. (2003). The importance of affine process in finance is twofold: it is very general family of stochastic process containing most of well-known Markov jump-diffusions processes and it allows for closed-form solutions for majority of pricing problems. In this thesis we will consider affine processes in discrete time setting.

[4] For an introduction to jump-diffusion models see Cont and Tankov (2004).

[5] For an introduction to local and stochastic volatility models see Gatheral (2011).

In the time-varying volatility models of the ARCH-GARCH families (Engle, 1982; Bollerslev, 1996; Glosten et al., 1993; Nelson, 1991), returns feature conditional heteroskedasticity which is described by an auto-regressive structure. Describing variance by recursion that facilitates maximum likelihood estimation has lead GARCH models to pioneer the field of measuring and predicting the volatility dynamics. More recently, thanks to the availability of high-frequency data, the so called Realized Volatility (RV) approach also became a prominent approach for measuring volatility.[6] RV is defined as a sum of consecutive squared intra-day returns and under the assumption that price is a L^2 semi-martingale it can be shown that neglecting microstructure noise it is a consistent estimator of quadratic variation of the price. The key advantage of RV approach is that the mentioned estimation procedure makes volatility an observable quantity which removes the need of volatility filtering and this in turn significantly simplifies estimation of the model parameters. The standard model for describing and forecasting the dynamics of RV is the Heterogeneous Autoregressive multi-components model by Corsi (2009) which together with information contained in RV measure provides superior volatility predicting performance.

The main problem of accommodating econometric models for option pricing application was lack of risk-neutralisation procedure. Relatively lately, Duan (1995) using equilibrium argument and postulating particular conditions on agent's risk preference have proposed the locally risk-neutral valuation relationship for GARCH processes. Since then we have witnessed renaissance of discrete time volatility models and many examples of GARCH based option pricing models have been proposed (Heston and Nandi (2000), Gourieroux and Monfort (2007), Christoffersen et al. (2008) and Gagliardini et al. (2011) among others). Empirical comparison suggests that GARCH models outperform continuous time stochastic volatility models (Lehar et al., 2002; Christoffersen et al., 2006). Recently, it has been shown that option pricing models based

[6]See Andersen et al. (2001b, 2003); Barndorff-Nielsen and Shephard (2001, 2002a,b, 2005); Comte and Renault (1998).

on realized volatility provide good performance (Corsi et al., 2013; Christoffersen et al., 2014; Majewski et al., 2015). In this thesis we will present a general option pricing framework encompassing both GARCH and RV based models.

Contrary to continuous-time models, volatility, in discrete time models, is readily observable from the history of asset prices by filtration procedure (GARCH models) or by precise non-parametric measurement from intra-day data (RV approach) and consequently all the parameters of discrete-time model can be easily estimated directly from the time series of observed quantities. It holds true even in the presence of complex dynamical features like long memory, multifractality, cascade and asymmetric effects. These features turn out to be crucial in option pricing and from now on we will consider only discrete time option pricing models.

1.2 Introduction to stochastic discounting

The basic function of financial market in the economy is an efficient allocation of capital. Agents can invest their wealth surplus in exchange for future stream of income. The investment decisions of agents are based on two aspects: their statistical view on the cash-flow which is described by probability law $\mathbb{P}$ and their attitude towards particular risks which can be described by stochastic discounting. The way in which agents are discounting random payoffs depends on compensation they are demanding for bearing investment uncertainty and it gives rise to an operator which associates a price to every claim.

We consider a risk-free asset with interest rate r and a risky asset with price S_t and geometric return

$$y_{t+1} = \log\left(\frac{S_{t+1}}{S_t}\right)$$

defined on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{1\leq t\leq T}, \mathbb{P})$. The state space under our consideration is generated by the risky asset price S till some horizon T, $\Omega = \mathbb{R}^T$. Let $L^2_T = L^2(\mathbb{R}^T, \mathbb{P}) = \{X :$

$\mathbb{E}^{\mathbb{P}}\left[\sum_{i=1}^{T} X_i^2\right] < \infty\}$ be a set of payoffs. It is easy to see that L_T^2 is a Hilbert space with a scalar product $\langle X, Y \rangle = \mathbb{E}^{\mathbb{P}}\left[\sum_{i=1}^{T} X_i Y_i\right]$. Any operator $\mathcal{Q}_T : L_T^2 \to \mathbb{R}$ associating a price to a payoff is called a pricing operator. In this section we will provide conditions which a reasonable pricing operator should satisfy.

Definition 1. A payoff $X \in L_T^2$ is an arbitrage opportunity if $X_j \geq 0$ for every $j \in \{1, \dots T\}$ almost surely ($\mathbb{P}(X_j \geq 0 \text{ for every } j \in \{1, \dots, T\}) = 1$) with non-zero probability of one component being positive ($\mathbb{P}(X_j > 0 \text{ for some } j \in \{1, \dots, T\}) > 0$) and has price $\mathcal{Q}_T(X) \leq 0$.

We call an operator $\mathcal{Q}_T$ positive if $\mathcal{Q}_T(X) \geq 0$ for $X \geq 0$ almost surely, where $\geq$ has to be understood componentwise. We call an operator $\mathcal{Q}_T$ strictly positive if it is positive and additionally $\mathcal{Q}_T(X) > 0$ if $\mathbb{P}(X_j > 0 \text{ for some } j \in \{1, \dots, T\}) > 0$. Linear, strictly positive pricing operators play very important role in asset pricing theory.

Theorem 2. *There is no arbitrage opportunities in the market if and only if there exists a strictly positive linear pricing operator.*

The sufficient condition for no arbitrage in the market is an immediate consequence of the definitions of arbitrage and strictly positive operator. The necessary condition is the difficult part of the proof and we will not prove it here. It becomes substantially simpler if the state space Ω is finite. In that case space of claims is equal to $\mathbb{R}^m$, where m is cardinality of Ω and using results from convex analysis one can easily construct desired operator (see Duffie (2010) or Cochrane (2005)). In the case when state space is infinite, one has to operate within topology induced by L_T^2 space. For the details of the proof see Bühlmann et al. (1998).

The existence of pricing operator guarantees existence of a particular stochastic process.

Theorem 3. *If there exists a strictly positive linear pricing operator $\mathcal{Q}_T : L_T^2 \to \mathbb{R}$ then there exists a positive payoff $\tilde{M} \in L_T^2$ such that $\mathcal{Q}_T(X) = \mathbb{E}\left[X\tilde{M}\right]$ for all $X \in L_T^2$.*

Proof. Since L_T^2 is a Hilbert space it follows immediately from Riesz representation theorem. □

An economic interpretation of $\tilde{M} = (\tilde{M}_1, \tilde{M}_2, \ldots, \tilde{M}_T)$ from Riesz representation in Theorem 3 is that it is discounting the future value of a payoff X. To disentangle the risk discounting from temporal discounting we introduce process $M_t = \mathrm{e}^{rt}\tilde{M}_t$. Let us observe an interesting property of process M_t. Introducing notation that $S^{(t)}$ is a payoff from set L^2_T with S_t on t-th component and 0 on otherwise we obtain for $t \geq 1$ that

$$S_0 = \mathcal{Q}_T(S^{(t)}) = \mathbb{E}^{\mathbb{P}}\left[S^{(t)} \cdot \tilde{M}\right] = \mathrm{e}^{-rt}\mathbb{E}^{\mathbb{P}}\left[S_t M_t\right]. \tag{1.2.1}$$

Therefore one can formally show that stochastic process $\mathrm{e}^{-rt}S_t M_t$ is a $\mathbb{P}$-martingale. In this sense process M_t is discounting risk associated with the future states of the economy and it is called a Stochastic Discount Factor (SDF). Moreover considering the price of risk-free asset we obtain that it determines the mean of SDF,

$$\mathbb{E}^{\mathbb{P}}\left[M_t\right] = \mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{rt}\tilde{M}_t\right] = \mathcal{Q}_T\left(B^{(t)}\right) = 1, \tag{1.2.2}$$

where $B^{(t)}$ is a bond with maturity t (a risk-free asset with a payoff e^{rt} at time t). Higher moments of SDF depend on risk preferences of investors.

Valuing claims by taking the time-discounted expected value of payoff under physical measure would lead to arbitrage opportunity and hence it cannot be accepted as an asset pricing methodology.[7] The failure of this approach becomes comprehensible if one acknowledges that the value of money depends not only on time but also on a state of economy. One dollar in a bad state of economy is worth more than one dollar in a good state of economy. Therefore, during pricing of an asset one should discount both time and state of the World with associated risks, which is achieved by stochastic discounting mechanism.

Existence of stochastic discount factor enables us to define a probability measure $\mathbb{Q}$ equivalent

[7] See a very nice discussion in Carr (2005), where an arising example of arbitrage strategy is provided.

to $\mathbb{P}$ (we call two measures $\mathbb{P}$ and $\mathbb{Q}$ equivalent when $\mathbb{Q}(A) = 0$ if and only if $\mathbb{P}(A) = 0$). Let us denote by $\mathbb{P}_t$ the family probability measures such that $\mathbb{P}_t = \mathbb{P}|\mathcal{F}_t$. Then one can construct a family of probability measures $\mathbb{Q}_t$ satisfying relations $\mathrm{d}\mathbb{Q}_t = M_t \mathrm{d}\mathbb{P}_t$ and $\mathbb{Q}_t|\mathcal{F}_{t-1} = \mathbb{Q}_{t-1}$.[8] Using the notation $\mathbb{P} = \mathbb{P}_T$ and $\mathbb{Q} = \mathbb{Q}_T$ we obtain

$$\mathbb{Q}(S_t \in A) = \mathbb{E}^{\mathbb{P}} \left[\chi_A \left(S_t \right) M_t \right], \tag{1.2.3}$$

where χ_A is indicator function of a set A. Using the definition of measure $\mathbb{Q}$ and the property (1.2.1) of stochastic discount factor we obtain that S_t is a $\mathbb{Q}$-martingale. For that reason measure $\mathbb{Q}$ is called risk-neutral measure (or equivalent martingale measure). Price of any claim in the market is an expectation of its payoff under risk-neutral measure:

$$\mathcal{Q}_T \left(S^{(t)} \right) = \mathbb{E}^{\mathbb{Q}} \left[S_t \right]. \tag{1.2.4}$$

All above results can be collected in the fundamental theorem of financial mathematics.

Theorem 4 (The First Fundamental Theorem of Asset Pricing)**.** *The following five statements are equivalent:*

1. *There are no arbitrage opportunities.*
2. *A strictly positive, linear pricing operator $\mathcal{Q}_T$ exists.*
3. *A stochastic discount factor exists.*
4. *There exists a process M_t such that $\mathrm{e}^{-rt} S_t M_t$ is a $\mathbb{P}$-martingale.*
5. *A risk-neutral probability measure exists.*

While the First Fundamental Theorem of Asset Pricing states the conditions for the existence of strictly positive linear pricing operator, the Second Fundamental Theorem of Asset Pricing

[8] For details see Bühlmann et al. (1996).

states conditions for the uniqueness of the pricing operator. We call a market complete if agents can construct a strategy that will generate wealth exactly equal to any claim available in the market.

Theorem 5 (The Second Fundamental Theorem of Asset Pricing)**.** *An arbitrage-free market is complete if and only if there exists a unique stochastic discount factor.*

For the proof of the theorem see Duffie (2010). An example of a complete market is Black-Scholes model. Stochastic volatility and all models considered in this thesis are incomplete. In the case of incomplete market one has to determine the pricing kernel. The form of stochastic discount factor is strictly related to the risk attitude of investors. In Section 2.2 we will see that stochastic discount factor is determined as a solution to a Pareto optimal allocation in the case of one asset and several investors. Here we present a result for a stochastic discount factor in the case of several assets and one agent.

Lets assume there is an economic agent who wants to choose a portfolio θ so that he optimizes his terminal wealth (at time t) by investing in L risky assets and one risk-free asset. The optimization problem of the agent is

$$\max_{\theta} \mathbb{E}^{\mathbb{P}} \left[u(W_t^{\theta}) \right] \tag{1.2.5}$$

with the constraint

$$\sum_{i=0}^{L} \theta_i S_t^i = W_0^{\theta}, \tag{1.2.6}$$

where W_t^{θ} is value of portfolio θ at time t and u is a utility function of the agent. The assumption of no-arbitrage condition and convexity of the function $\theta \to \mathbb{E}^{\mathbb{P}} \left[u(W_t^{\theta}) \right]$ is sufficient to ensure the existence of solution θ^* to the above optimization problem. Then the SDF is given in the following form

$$M_t = \frac{u'(W_t^{\theta^*})}{\mathbb{E}^{\mathbb{P}}\left[u'(W_t^{\theta^*})\right]}. \tag{1.2.7}$$

For instance, assuming the logarithmic utility of an investor one obtains that SDF is the inverse of a Kelly portfolio. For further details see for example Kardaras (2010).

A great advantage of discrete time models is that they offer a simple procedure to estimate pricing kernel. In continuous-time setting parameters of the model are usually fitted directly to the option prices, completely neglecting the information contained in the time series of the log-returns of the underlying and the preferences of agents. On the contrary discrete-time volatility models provide a straightforward insight how investors are pricing certain risks. For instance, assuming multi-dimensional power utility function of agents, discrete-time models provide an easy estimation procedure of variance risk aversion.

Chapter 2

General Option Pricing Framework

2.1 Modelling volatility with multiple components

Despite large success, first stochastic volatility models (Heston, 1993; Heston and Nandi, 2000) cannot price correctly options with long or short maturity and out-of-money options. As a consequence they misfit strike profile and term structure of implied volatility surface. The reason for a poor performance at those regions of moneyness and maturity is that modelling volatility by single factor of volatility, running on a single time scale, is not sufficient to describe volatility dynamics. There exist many stylised facts that cannot be explained by single-factor volatility model. The family of term structures of variance swap has more realistic shapes under model with multi factors which we discuss in Section 5.2. Moreover principal component analysis (PCA) shows the necessity of using two components to explain the dynamics of variance swap rates (Filipovic et al. (2015); Ait-Sahalia et al. (2015)), while PCA of volatility surface dynamics suggests at least two-three factors (Alexander (2001); Cont et al. (2002)). Finally, the existence of several stochastic volatility factors running on different time scales has been proven in the literature using empirical data (see for example Müller et al. (1997)).

The main purpose of introducing a multi-factor structure in volatility modelling is to account for dependencies among volatilities at different time scales. Currently, there exist two alternative approaches in the literature. The first one is to decompose the daily volatility into several factors and model the dynamics of each factor independently, as done by Christoffersen et al. (2008) or Fouque and Lorig (2011) in terms of short-run and long-run volatility components. The second approach is to define factors as an average of past volatilities over different time horizons, for instance the daily, weekly and monthly components in Corsi (2009). In this section we describe a general framework introduced in Majewski et al. (2015) which includes both approaches.

To model the dynamics of log-returns of the risky asset we define the k-dimensional vector of $\mathcal{F}_t$-measurable volatility factors $\mathrm{f}_t^{(1)}, \ldots, \mathrm{f}_t^{(k)}$ which we shortly denote as $\mathbf{f}_t$. The daily log-returns on day t are modelled by equation

$$y_t = r + \lambda\, \mathcal{L}\left(\mathbf{f}_{t-1}, \mathbf{f}_t\right) + \sqrt{\mathcal{L}\left(\mathbf{f}_{t-1}, \mathbf{f}_t\right)}\; \epsilon_t\,, \tag{2.1.1}$$

where $\mathcal{L} : \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}_+$ is a linear function of factors giving volatility at day t, r is the risk-free rate, λ is the market price of risk, and ϵ_t are i.i.d. $\mathcal{N}(0,1)$. Function $\mathcal{L}$ acts on $\mathbf{f}_t$ in the case of realized volatility models and on $\mathbf{f}_{t-1}$ if GARCH model is considered. The different domain of function $\mathcal{L}$ for those two class of models underlines one of the most important differences between realized volatility and GARCH modelling approaches. In the case of realized variance models, volatility at day t is $\mathcal{F}_t$-measurable while for GARCH models, volatility at day $t+1$ is $\mathcal{F}_t$-measurable. As a consequence a vector of variance factors for realized variance $\mathbf{f}_t$ corresponds to the level of variance factors on day t, while in the case of GARCH it corresponds to level of variance factors on the following day $t+1$. In both cases we model $\mathbf{f}_{t+1}$ as

$$\mathbf{f}_{t+1} | \mathbf{F}_t, \mathbf{L}_t \sim \mathcal{D}\left(\Theta_0, \Theta(\mathbf{F}_t, \mathbf{L}_t)\right)\,, \tag{2.1.2}$$

where $\mathcal{D}$ denotes a generic distribution depending on the vector of parameters Θ which is a

k-dimensional function of the matrices $\mathbf{F}_t = (\mathbf{f}_t, \ldots, \mathbf{f}_{t-p+1}) \in \mathbb{R}^{k\times p}$ and $\mathbf{L}_t = (\boldsymbol{\ell}_t, \ldots, \boldsymbol{\ell}_{t-q+1}) \in \mathbb{R}^{k\times q}$ for $p > 0$ and $q > 0$, respectively. We consider the case of a linear dependence of Θ on $\mathbf{F}_t$ and $\mathbf{L}_t$

$$\Theta(\mathbf{F}_t, \mathbf{L}_t) = \mathbf{d} + \sum_{i=1}^{p} \mathbf{M}_i \mathbf{f}_{t+1-i} + \sum_{j=1}^{q} \mathbf{N}_j \boldsymbol{\ell}_{t+1-j}\,, \tag{2.1.3}$$

where $\mathbf{M}_i, \mathbf{N}_j \in \mathbb{R}^{k\times k}$ for $i = 1, \ldots, p$ and $j = 1, \ldots, q$, $\mathbf{d} \in \mathbb{R}^k$, and vectors $\boldsymbol{\ell}_t$ are of the form

$$\boldsymbol{\ell}_t = \begin{bmatrix} \left(\epsilon_t - \gamma_1 \sqrt{\mathcal{L}\left(\mathbf{f}_{t-1}, \mathbf{f}_t\right)}\right)^2 \\ \vdots \\ \left(\epsilon_t - \gamma_k \sqrt{\mathcal{L}\left(\mathbf{f}_{t-1}, \mathbf{f}_t\right)}\right)^2 \end{bmatrix}. \tag{2.1.4}$$

The vector Θ_0 collects all the parameters of the distribution $\mathcal{D}$ which do not depend on the past history of the factors and of the leverage. For the distribution $\mathcal{D}$ considered in this thesis (Dirac delta and non-central Gamma distribution) the sufficient condition for the non-negativity of process reads:

$$\mathbf{d} \geq 0 \quad \mathbf{M}_i \geq 0 \ \text{ for all } i \in \{1, \ldots, p\} \quad \mathbf{N}_j \geq 0 \ \text{ for all } j \in \{1, \ldots, q\}, \tag{2.1.5}$$

where $\geq$ has to be meant as component-wise inequality.

The results presented in this thesis are derived under the general assumption

Assumption 6. *The following relation holds true*

$$\mathbb{E}\left[\mathrm{e}^{zy_{s+1}+\mathbf{b}\cdot\mathbf{f}_{s+1}+\mathbf{c}\cdot\boldsymbol{\ell}_{s+1}} \middle| \mathcal{F}_s\right] = \mathrm{e}^{\mathcal{A}(z,\mathbf{b},\mathbf{c})+\sum_{i=1}^{p}\boldsymbol{\mathcal{B}}_i(z,\mathbf{b},\mathbf{c})\cdot\mathbf{f}_{s+1-i}+\sum_{j=1}^{q}\boldsymbol{\mathcal{C}}_j(z,\mathbf{b},\mathbf{c})\cdot\boldsymbol{\ell}_{s+1-j}} \tag{2.1.6}$$

for some functions $\mathcal{A} : \mathbb{R}\times\mathbb{R}^k\times\mathbb{R}^k \to \mathbb{R}$, $\boldsymbol{\mathcal{B}}_i : \mathbb{R}\times\mathbb{R}^k\times\mathbb{R}^k \to \mathbb{R}^k$, *and* $\boldsymbol{\mathcal{C}}_j : \mathbb{R}\times\mathbb{R}^k\times\mathbb{R}^k \to \mathbb{R}^k$, *where* $\mathbf{b}, \mathbf{c} \in \mathbb{R}^k$ *and* $\cdot$ *stands for the scalar product in* $\mathbb{R}^k$.

Our framework is suited to include both GARCH-like models and realized volatility models or combination of two. As far as the former class is concerned, we encompass the family of multiple component GARCH models with parabolic leverage pioneered in Heston and Nandi (2000) and later extended to the two Component GARCH (CGARCH) by Christoffersen et al. (2008). For instance, the latter model corresponds to the following dynamics

$$\begin{aligned} y_{t+1} &= r + \lambda h_{t+1} + \sqrt{h_{t+1}}\epsilon_{t+1}\,, \\ h_{t+1} &= q_{t+1} + \beta_h \left(h_t - q_t\right) + \alpha_h \left(\epsilon_t^2 - 1 - 2\gamma_h \epsilon_t \sqrt{h_t}\right)\,, \\ q_{t+1} &= \omega + \beta_q q_t + \alpha_q \left(\epsilon_t^2 - 1 - 2\gamma_q \epsilon_t \sqrt{h_t}\right)\,. \end{aligned} \tag{2.1.7}$$

Setting $k = 2$, we define $\mathrm{f}_t^{(1)} = h_{t+1} - q_{t+1}$ and $\mathrm{f}_t^{(2)} = q_{t+1}$ and rewrite the model as

$$\begin{bmatrix} \mathrm{f}_{t+1}^{(1)} \\ \mathrm{f}_{t+1}^{(2)} \end{bmatrix} = \begin{bmatrix} -\alpha_h \\ \omega - \alpha_q \end{bmatrix} + \begin{bmatrix} \beta_h - \alpha_h \gamma_h^2 & -\alpha_h \gamma_h^2 \\ -\alpha_q \gamma_q^2 & \beta_q - \alpha_q \gamma_q^2 \end{bmatrix} \begin{bmatrix} \mathrm{f}_t^{(1)} \\ \mathrm{f}_t^{(2)} \end{bmatrix} + \begin{bmatrix} \alpha_h & 0 \\ 0 & \alpha_2 \end{bmatrix} \begin{bmatrix} \left(\epsilon_t - \gamma_h \sqrt{\mathcal{L}\left(\mathbf{f}_t\right)}\right)^2 \\ \left(\epsilon_t - \gamma_q \sqrt{\mathcal{L}\left(\mathbf{f}_t\right)}\right)^2 \end{bmatrix}\,, \tag{2.1.8}$$

where $\mathcal{L}\left(\mathbf{f}_t, \mathbf{f}_{t+1}\right) = \mathrm{f}_t^{(1)} + \mathrm{f}_t^{(2)} = h_t$. If we now specify for $\mathcal{D}$ in eq. (2.1.2) the form of a Dirac delta distribution, define $\mathbf{d} = \left(-\alpha_h,\ \omega - \alpha_q\right)^t$, and identify the matrices $\mathbf{M}_1$ and $\mathbf{N}_1$ in a natural way from the right term side of eq. (2.1.8), the model by Christoffersen et al. (2008) fits the general formula (2.1.2). It is worth mentioning that for the CGARCH model it is not possible to ensure the non-negative definiteness of both h_t and q_t for all t (condition (2.1.5) is not satisfied). Nonetheless, for realistic values of the parameters the probability of obtaining negative volatility factors is extremely low, and this drawback is largely compensated for by the effectiveness of the model in capturing real time series empirical features. Since all models proposed in the thesis are subject to the issue of positivity, we discuss it in greater detail in Section 4.2.

The second example that we discuss is the class of realized volatility models known as Autoregressive Gamma Processes (ARG) introduced in Gourieroux and Jasiak (2006), to whom the Heterogeneous Autoregressive Gamma (HARG) model presented in Corsi et al. (2013) belongs. The process RV_t is an ARG(p) if and only if its conditional distribution given $(\mathrm{RV}_{t-1}, \ldots, \mathrm{RV}_{t-p})$

is a noncentred gamma distribution $\bar{\gamma}(\delta, \sum_{i=1}^{p} \beta_i \mathrm{RV}_{t-i}, \theta)$, where δ is the shape, $\sum_{i=1}^{p} \beta_i \mathrm{RV}_{t-i}$ the non-centrality, and θ the scale. Then, the model described by eq.s (2.1.2)-(2.1.3) reduces to an ARG(p) if we fix $k = 1$, $\mathrm{f}_t = \mathrm{RV}_t$, $\mathcal{D}\left(\Theta_0, \Theta(\mathbf{F}_{t-1})\right) = \bar{\gamma}\left(\delta, \Theta(\mathbf{F}_{t-1}), \theta\right)$ with

$$\Theta_0 = (\delta, \theta)^t \ , \quad \text{and} \quad \Theta(\mathbf{F}_{t-1}) = \sum_{i=1}^{p} \beta_i \mathrm{f}_{t-i} \, .$$

In Chapter 2 we will introduce new models belonging to our general framework. They are extension of HARG-RV model with heterogenous and analytically tractable leverage structure called LHARG-RV, extension of LHARG-RV with jumps called JLHARG-RV and the mixture of LHARG-RV and GARCH model. Moreover we will reconsider CGARCH model of Christoffersen et al. (2008) by applying new change of measure to obtain new dynamics under risk-neutral measure.

Recently, Christoffersen et al. (2014) have proposed an alternative option pricing model nesting GARCH and realized volatilities models called General Affine Realized Volatility (GARV). Even though both approaches are very general, provide closed-form solutions and allow for multi component structure, they do not coincide. The main difference is in the addition of new source of randomness related to the realized volatility. While Christoffersen et al. (2014) are adding new innovation process in the realized volatility dynamics, we are introducing a transition distribution (in the examples given in this thesis we consider a non-central gamma distribution). As a consequence our approach nests HARG-RV model whereas GARV do not.

The general framework defined by eq.s (2.1.1)-(2.1.4) combined with the assumption (2.1.6) allows us to completely characterise the MGF of the log-returns under the physical measure.

Theorem 7. *If the dynamics of the underlying price satisfies Assumption 6 then the moment generating function of log-returns under the physical measure $\mathbb{P}$ is given by recursive relations in terms of functions $\mathcal{A}$, $\mathcal{B}_i$, $\mathcal{C}_j$, where $i \in \{1, \ldots, p\}$ and $j \in \{1, \ldots, q\}$.*

Proof. For the proof and the recursive relations see Appendix A.1. □

2.2 Stochastic discounting with multi-dimensional Esscher transform

The standard problem in asset pricing theory is how one can identify the stochastic discount factor which gives an economically consistent and justifiable price for a contingent claim. Theorem 4, often called the first fundamental theorem of asset pricing, states that it exists if the market does not allow arbitrage opportunities. The second fundamental of asset pricing states that uniqueness of stochastic discount factor is equivalent to completeness of the market. Since the market considered in our framework are generally incomplete there are many stochastic discount factors and the choice of a suitable pricing operator is arbitrary.

In our general framework we will apply multi-dimensional exponential-affine stochastic discount factor given by formula

$$M_t = \prod_{s=0}^{t-1} M_{s,s+1} \tag{2.2.1}$$

where

$$M_{s,s+1} = \frac{\mathrm{e}^{-\boldsymbol{\nu}_{\mathbf{f}} \cdot \mathbf{f}_{s+1} - \nu_y y_{s+1}}}{\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{-\boldsymbol{\nu}_{\mathbf{f}} \cdot \mathbf{f}_{s+1} - \nu_y y_{s+1}} \middle| \mathcal{F}_s\right]}, \tag{2.2.2}$$

with $\boldsymbol{\nu}_{\mathbf{f}} \in \mathbb{R}^k$. The one-dimensional and unconditional version of transform of probability measure given by (2.2.2) was originally introduced to actuarial science in seminal work of Esscher (1932) where random variables were independent and it was used to approximate the distribution of aggregate claims. Extensive application of Esscher transform to derivative pricing took off with an original paper by Gerber and Shiu (1993) where they extended Esscher's idea to Lévy processes framework. One-dimensional version of conditional Esscher transfrom described by equations (2.2.1)-(2.2.2) has been introduced in a beautiful paper by Bühlmann et al. (1996).

Advantages of using multi-dimensional conditional Esscher transform are threefold. First, it is easy to write the constraints that the parameters of the Esscher transform have to satisfy in order to be a stochastic discount factor in our framework.

Theorem 8 (No arbitrage restriction)**.** *If the dynamics of the underlying price satisfies Assumption 6 then the Esscher transform (2.2.2) is a stochastic discount factor if, and only if the following relations are satisfied*

$$\begin{aligned} \mathcal{A}(1-\nu_y,-\boldsymbol{\nu}_{\mathbf{f}},\mathbf{0}) &= r+\mathcal{A}(-\nu_y,-\boldsymbol{\nu}_{\mathbf{f}},\mathbf{0}) \\ \boldsymbol{\mathcal{B}}_i(1-\nu_y,-\boldsymbol{\nu}_{\mathbf{f}},\mathbf{0}) &= \boldsymbol{\mathcal{B}}_i(-\nu_y,-\boldsymbol{\nu}_{\mathbf{f}},\mathbf{0}) \quad \textit{for } i=1,\dots,p \\ \boldsymbol{\mathcal{C}}_j(1-\nu_y,-\boldsymbol{\nu}_{\mathbf{f}},\mathbf{0}) &= \boldsymbol{\mathcal{C}}_j(-\nu_y,-\boldsymbol{\nu}_{\mathbf{f}},\mathbf{0}) \quad \textit{for } j=1,\dots,q. \end{aligned} \tag{2.2.3}$$

.

Proof. From Theorem 4 we know that there exists an SDF in the market if price process is $\mathbb{Q}$-martingale or equivalently price process multiplied by stochastic discount factor is $\mathbb{P}$-martingale. The last condition can be read as

$$\mathbb{E}^{\mathbb{P}}\left[M_{s,s+1}\mathrm{e}^{y_{s+1}}|\mathcal{F}_s\right]=\mathrm{e}^r \quad \text{for} \in \{0,1,\dots,T-1\}. \tag{2.2.4}$$

Firstly, let us rewrite Esscher transform as

$$\begin{aligned} M_{s,s+1} &= \frac{\mathrm{e}^{-\boldsymbol{\nu}_1\cdot\mathbf{f}_{s+1}-\nu_2 y_{s+1}}}{\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{-\boldsymbol{\nu}_1\cdot\mathbf{f}_{s+1}-\nu_2 y_{s+1}}|\mathcal{F}_s\right]} \\ &= \exp\begin{pmatrix} -\mathcal{A}(-\nu_y,-\boldsymbol{\nu}_{\mathbf{f}},\mathbf{0})-\sum_{i=1}^{p}\boldsymbol{\mathcal{B}}_i(-\nu_y,-\boldsymbol{\nu}_{\mathbf{f}},\mathbf{0})\cdot\mathbf{f}_{s+1-i} \\ -\sum_{i=1}^{q}\boldsymbol{\mathcal{C}}_i(-\nu_y,-\boldsymbol{\nu}_{\mathbf{f}},\mathbf{0})\cdot\boldsymbol{\ell}_{s+1-i}-\boldsymbol{\nu}_{\mathbf{f}}\cdot\mathbf{f}_{s+1}-\nu_y y_{s+1} \end{pmatrix}, \end{aligned} \tag{2.2.5}$$

where $\boldsymbol{\nu}_{\mathbf{f}}\in\mathbb{R}^k$ and functions $\mathcal{A}$, $\boldsymbol{\mathcal{B}}_i$ and $\boldsymbol{\mathcal{C}}_j$ are defined in (2.1.6). Finally, the condition (2.2.4) reads

$$\begin{aligned}&\mathbb{E}^{\mathbb{P}}\left[\exp\left(-\boldsymbol{\nu}_{\boldsymbol{f}}\cdot\mathbf{f}_{s+1}+(1-\nu_y)\,y_{s+1}\right)|\mathcal{F}_s\right]\\&=\exp\left(r+\mathcal{A}(-\nu_y,-\boldsymbol{\nu}_{\mathbf{f}},\mathbf{0})+\sum_{i=1}^{p}\boldsymbol{\mathcal{B}}_i(-\nu_y,-\boldsymbol{\nu}_{\mathbf{f}},\mathbf{0})\cdot\mathbf{f}_{s+1-i}+\sum_{j=1}^{q}\boldsymbol{\mathcal{C}}_j(-\nu_y,-\boldsymbol{\nu}_{\mathbf{f}},\mathbf{0})\cdot\boldsymbol{\ell}_{s+1-j}\right).\end{aligned} \tag{2.2.6}$$

Using once again the relation (2.1.6) we obtain conditions for Esscher transform (2.2.2) to be an SDF. Following Theorem 4 conditions (2.2.3) can be viewed as no-arbitrage conditions. □

Second, SDF (2.2.2) guarantees analytic expression for moment generating function under risk-neutral measure which allows us to write a semi-closed formula for option price.

Theorem 9. *If the dynamics of the underlying price satisfies Assumption 6 and the SDF is given by (2.2.2) then the moment generating function of log-returns under the risk-neutral measure $\mathbb{Q}$ is given by recursive relations in terms of functions $\mathcal{A}$, $\mathcal{B}_i$, $\mathcal{C}_j$, where $i \in \{1, \ldots, p\}$ and $j \in \{1, \ldots, q\}$.*

Proof. For the proof and the recursive relations see Appendix A.1. □

Third, Esscher transform has a strong economic foundation. Lets assume that there are N agents in the economy and the total volume of shares of asset with price S_t is equal V. Lets define the wealth income in the economy as follows $W_{t+1} = V(\ln S_{t+1} - \ln S_t)$. One of the basic question in economic theory is which allocation of wealth among agents

$$\left(W_{t+1}^{(1)}, W_{t+1}^{(2)}, \ldots, W_{t+1}^{(N)}\right) \tag{2.2.7}$$

is optimal. Allocation (2.2.7) can be equivalently expressed in terms of number of shares owned by each agent $(V_1, V_2, \ldots, V_N)$. Let us denote a price of payoff W_{t+1} at time t by $\mathcal{Q}_{t,t+1}(W_{t+1})$. In this thesis we will consider Pareto optimal allocation and following Bühlmann et al. (1998) it can be obtained as Price equilibrium, i.e. at time t two conditions has to be satisfied: for each agent j we have that

$$\mathbb{E}\left[u_j\left(W_{t+1}^{(j)} - \mathcal{Q}_{t,t+1}\left(W_{t+1}^{(j)}\right) - \mathbf{m}_j\mathbf{f}_{t+1}\right)|\mathcal{F}_t\right] \tag{2.2.8}$$

achieves maximum among all possible random variables $W_{t+1}^{(j)} \in L_2$ and the allocation (2.2.7) has to satisfy

$$W_{t+1} = \sum_{j=1}^{N} W_{t+1}^{(j)}. \tag{2.2.9}$$

Condition (2.2.8) means that agents want to maximise their expected utility of profit from the investment $\left(W_{t+1}^{(j)} - \mathcal{Q}_{t,t+1}\left(W_{t+1}^{(j)}\right)\right)$ corrected by the variance factors with penalty coefficients $\mathbf{m}_j$. We will assume that agents have power utility:

$$u_j(x) = \frac{1}{\gamma_j}\left(1 - \mathrm{e}^{-\gamma_j x}\right) \quad \text{for } j = 1, \dots, N, \tag{2.2.10}$$

with risk-aversion parameters given by $\gamma_j > 0$. The above problem can be restated as: which pricing operator $\mathcal{Q}_{t,t+1}$ is satisfying Pareto optimal allocation? We will derive now Esscher transform (2.2.2) from problem described by (2.2.7)-(2.2.10).

Theorem 10. *Pricing operator $\mathcal{Q}_{t,t+1}$ given by Esscher transform (2.2.2) is a solution to Pareto optimal allocation problem described by (2.2.7)-(2.2.10).*

Proof. Given the concavity of function u, we obtain maximum in (2.2.8) if and only if for every j the first order condition is satisfied

$$\frac{\partial}{\partial V_j}\mathbb{E}\left[u_j\left(W_{t+1}^{(j)} - \mathcal{Q}_{t,t+1}\left(W_{t+1}^{(j)}\right) - \mathbf{m}_j\mathbf{f}_{t+1}\right)|\mathcal{F}_t\right] = 0. \tag{2.2.11}$$

Since conditional expectation is a linear operator we can move differential operator inside and we are going to obtain that

$$\begin{aligned} &\mathbb{E}\left[u_j'\left(W_{t+1}^{(j)}-\mathcal{Q}_{t,t+1}\left(W_{t+1}^{(j)}\right)-\mathbf{m}_j\mathbf{f}_{t+1}\right)y_{t+1}|\mathcal{F}_t\right]= \\ &=\mathbb{E}\left[u_j'\left(W_{t+1}^{(j)}-\mathcal{Q}_{t,t+1}\left(W_{t+1}^{(j)}\right)-\mathbf{m}_j\mathbf{f}_{t+1}\right)|\mathcal{F}_t\right]\mathcal{Q}_{t,t+1}\left(y_{t+1}\right)\end{aligned} \tag{2.2.12}$$

From equation (2.2.12) we observe that the candidate $M_{t,t+1}$ in Riesz representation has to satisfy relation

$$u_j'\left(W_{t+1}^{(j)}-\mathcal{Q}_{t,t+1}\left(W_{t+1}^{(j)}\right)-\mathbf{m}_j\mathbf{f}_{t+1}\right)=\mathbb{E}\left[u_j'\left(W_{t+1}^{(j)}-\mathcal{Q}_{t,t+1}\left(W_{t+1}^{(j)}\right)-\mathbf{m}_j\mathbf{f}_{t+1}\right)|\mathcal{F}_t\right]M_{t,t+1} \tag{2.2.13}$$

for every $j\in\{1,2,\dots,N\}$. Since $u_j'(x)=\mathrm{e}^{-\gamma x}$ and $\mathcal{Q}_{t,t+1}\left(W_{t+1}^{(j)}\right)$ is $\mathcal{F}_t$-measurable, we rewrite (2.2.13) as

$$\mathrm{e}^{-\gamma_j W_{t+1}^{(j)}+\gamma_j\mathbf{m}_j\mathbf{f}_{t+1}}=\mathbb{E}\left[\mathrm{e}^{-\gamma_j W_{t+1}^{(j)}+\gamma_j\mathbf{m}_j\mathbf{f}_{t+1}}|\mathcal{F}_t\right]M_{t,t+1} \tag{2.2.14}$$

for $j=1,2,\dots,N$ and then taking logarithm we obtain

$$-W_{t+1}^{(j)}+\mathbf{m}_j\mathbf{f}_{t+1}=\frac{1}{\gamma_j}\ln\mathbb{E}\left[\mathrm{e}^{-\gamma_j W_{t+1}^{(j)}+\gamma_j\mathbf{m}_j\mathbf{f}_{t+1}}|\mathcal{F}_t\right]+\frac{1}{\gamma_j}\ln M_{t,t+1}. \tag{2.2.15}$$

We take a sum of (2.2.15) over $j=1,2,\dots,N$ and using notation $\mathbf{m}=\frac{1}{V}\sum_{j=1}^N\mathbf{m}_j$ and $\frac{1}{\gamma}=\frac{1}{V}\sum_{j=1}^N\frac{1}{\gamma_j}$ we obtain

$$-\frac{\gamma}{V}W_{t+1}+\gamma\mathbf{m}\mathbf{f}_{t+1}=\frac{\gamma}{V}A_t+\ln M_{t,t+1} \tag{2.2.16}$$

where A_t is some $\mathcal{F}_t$-measurable random variable. From condition (2.2.4) we obtain the form of A_t and the stochastic discount factor has the following form

$$M_{t,t+1}=\frac{\mathrm{e}^{-\gamma y_{t+1}+\gamma\mathbf{m}\mathbf{f}_{t+1}}}{\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{-\gamma y_{t+1}+\gamma\mathbf{m}\mathbf{f}_{t+1}}|\mathcal{F}_t\right]} \tag{2.2.17}$$

□

If we compare (2.2.17) with (2.2.2) we obtain an interpretation of the parameters of the Esscher transform. The parameter ν_y is equal to γ i.e. it is aggregated risk aversions of agents multiplied by the volume of assets. The parameter $\boldsymbol{\nu}_{\mathbf{f}}$ is equal to $-\gamma\mathbf{m}$ i.e. it is $-\gamma$ multiplied by aggregated variance penalisation of agents divided by the volume of assets. Obviously the above considerations are only true if the preferences of all agents in the economy are given by power utility.

Chapter 3

Particular Models

3.1 Heterogeneous Autoregressive Gamma model of realized volatility with Heterogeneous parabolic Leverage (LHARG-RV)

3.1.1 Realized volatility and log-returns dynamics

Continuous-time stochastic volatility models are the most famous way of obtaining heavy-tailed log-returns in financial mathematics literature. Log-returns dynamics is described by SDE

$$\mathrm{d}Y(t) = \left(r + \lambda\sigma^2(t)\right)\mathrm{d}t + \sigma(t)\mathrm{d}W(t),$$

where r is the risk-free rate, λ is the market price of risk, $W(t)$ is a Brownian motion and $\sigma(t)$ is a stochastic process describing the volatility of log-returns. Ané and Geman (2000) show that Y can be seen as Brownian motion with a changed time

$$Y(t) = rt + \lambda\mathrm{IV}(t) + W\left(\mathrm{IV}(t)\right),$$

where $\mathrm{IV}(t)$ is integrated variance

$$\mathrm{IV}(t) = \int_0^t \sigma^2(s)\mathrm{d}s.$$

In stochastic analysis literature this feature is often described as $\mathrm{IV}(t)$ process being a random time change for all continuous diffusion processes. Identifying integrated variance with a measure of market activity provide us the financial interpretation of this result: rescaling log-return process by market activity restores Brownian motion in calendar time.

In the case of continuous-time diffusions integrated variance is equal to quadratic variation defined as follows

$$\mathrm{QV}(t) = \lim_{||P_n|| \to 0} \sum_{i=1}^{n} (Y_{t_i} - Y_{t_{i-1}})^2,$$

where P_n stands for an n-element partition of interval $[0, t]$ and the mesh of the partition is the length of the longest subinterval ($||P_n|| = \max\{|t_i - t_{i-1}| : i = 1, \ldots, n\}$). Then

$$Y(t)|\mathrm{QV}(t) \sim \mathcal{N}\left(rt + \lambda \mathrm{QV}(t), \mathrm{QV}(t)\right).$$

Even if volatility $\sigma(t)$ and quadratic variation are unobservable processes, there exists a reliable proxy of QV. Let us denote by QV_t quadratic variation at day t and by RV_t realized variance at day t

$$\mathrm{RV}_t = \sum_{i=1}^{M} y_{t,i}^2 \tag{3.1.1}$$

where $y_{t,i}$ are intra-day log-returns

$$y_{t,i} = Y\left(t - 1 + \frac{i}{M}\right) - Y\left(t - 1 + \frac{i-1}{M}\right), \quad \text{for } i = 1, \ldots, M.$$

Then RV_t is a consistent (as $M \to \infty$) estimator of QV_t. Precision or rate of convergence estimation of quadratic variation with realized variance has been verified in several studies (see

Barndorff-Nielsen and Shephard (2002b) among others). Description of RV measurement is described in Section 4.1.

The basic idea of realized variance goes back to Merton (1980) who showed that the integrated variance of Brownian motion can be approximated by RV_t in (3.1.1). Realized Variance as a measure of volatility was proposed by Andersen et al. (2001b) who have generalised the result of Merton (1980) to semi-martingales. This model-free (nonparametric) measure makes volatility an observable quantity, which can have several applications. First statistical properties of volatility can be tested directly and much simpler than in the case when volatility is latent. Second it can be applied to forecast future level of volatility with high accuracy. Third information contained in realized variance might be very useful in pricing financial derivatives.

Andersen et al. (2001a) and Andersen et al. (2003) show that log-returns standardised by realized variance are in the first approximation normally distributed.[1] Therefore we assume that log-returns has the following dynamics on daily scale

$$y_t = r + \lambda \mathrm{RV}_t + \sqrt{\mathrm{RV}_t}\epsilon_t. \tag{3.1.2}$$

where ϵ_t are i.i.d. with standard normal random distribution. The dynamics of log-returns like in equation (3.1.2) have been already assumed in Corsi et al. (2013).

3.1.2 Motivation and basic idea of Heterogeneous Autoregressive processes

Modern volatility models aim to incorporate three stylized fact: long memory in the volatility, multifractality and volatility cascade. Intuitively, process is perceived to have a long memory feature if its autocorrelation remains significant for several months. Long memory property of

[1]When the dynamics of logreturns includes jumps this approximation is not true anymore. In Section 3.4 we are going to consider a model taking into account jumps in realized variance.

the realized volatility has been widely accepted since the seminal analyses by Ding et al. (1993), Andersen et al. (2001b) and Andersen et al. (2003). Despite broad recognition of long memory property there is still a lot of ambiguity about statistical tests verifying its existence. Formally, time series RV_t is said to have long memory property if

$$\sum_{k=1}^{\infty} \mathrm{Cov}\left(\mathrm{RV}_t, \mathrm{RV}_{t+k}\right) = \infty. \tag{3.1.3}$$

Since we do not have infinite time series of realized variance, the above definition is useless in practical applications and it is replaced by testing if volatility has power law decay

$$\mathrm{Cov}\left(\mathrm{RV}_t, \mathrm{RV}_{t+k}\right) \sim C/k^{\gamma}, \tag{3.1.4}$$

where $\gamma < 1$. Since it is not possible to estimate the asymptotic behaviour of the covariance function in model-free setting, usually it is done under some specification. For example, the most classical statistical test for long memory assumes ARFIMA dynamics of realized volatility which allows us to compute explicitly the variance of integrated realized volatility

$$V(\Delta) = \mathrm{Var}\left[\sum_{i=1}^{\Delta} \sqrt{\mathrm{RV}_i}\right] = \Delta^{2-\gamma}, \tag{3.1.5}$$

where γ is a decay of autocovariance function in (3.1.4). If we obtain that $V(\Delta)$ behaves like $\Delta^{2-\gamma}$ the test concludes that there is a long memory with parameter γ. However, it turns out that there exist processes satisfying the above long memory test which have autocorrelation function without power law decay (see Corsi (2009), LeBaron (2001), Gatheral et al. (2014)). For this reason, econometricians continue to debate whether market volatility process is a real long memory process or it just resembles one.

Similar ambiguity arises when multifractality property is considered. We say that RV_t is multifractal if

$$\mathbb{E}\left[|\mathrm{RV}_t|^q\right] = ct^{\zeta(q)} \tag{3.1.6}$$

where c is a constant, $q > 0$ is the order of the moment and $\zeta(q)$ is a nonlinear (concave) function. The evidence of multifractality in financial data has been observed by Ding et al. (1993), Fisher et al. (1997) and Calvet and Fisher (2002) among others. One can formally show that only multiplicative process have multifractal property. However, multifractal feature can be observed in the data generated from a process which is not really multifractal (see LeBaron (2001) and Corsi (2009)). Both long memory and multi-fractality can be detected falsely, if the aggregation level in the data is not large enough compared to the lowest frequency component of the model. These ambiguities about multifractatility and long memory raises doubts about applying sophisticated multiplicative models that are difficult to identify and estimate, and it suggests to employ simpler additive process that can exhibit the demanded properties.

The most prominent example of an additive processes generating time series with long memory and multifractal property is Heterogeneous Autoregressive process (HAR) introduced to financial literature by Corsi (2009). The basic idea of HAR processes stems from "Heterogeneous Market Hypothesis" by Müller et al. (1997) which aims to explain positive correlation between market volatility and market presence. In the classical, homogeneous market framework with all market agents identical, the more agents are active in the market, the faster should price converge reducing the volatility. On the contrary, in the heterogeneous setting, agents try to execute their transaction at different prices creating volatility.

While heterogeneity of agents may be due to difference in theirs' beliefs, endowments, degree of information, risk profiles and so on, the HAR model originates from the assumption that agents have different investment horizon. Participants of financial markets can be characterised by different trading frequency: agents with very high trading frequency (dealers and high frequency traders) are actors with low trading frequency (central banks and pension funds). Members of each group perceive and react to events on financial markets with different trading frequency so that theirs contribution to overall market volatility can be described by different components of volatility.

HAR process is characterised by the different impact that past realized variances aggregated on a daily, weekly and monthly basis have on today's realized variance. Lagged terms are collected in three different non-overlapping factors: $\mathrm{RV}_t^{(d)}$ (short-term volatility factor), $\mathrm{RV}_t^{(w)}$ (medium-term volatility factor), and $\mathrm{RV}_t^{(m)}$ (long-term volatility factor).

$$\mathrm{RV}_{t+1} = d + \beta_d \mathrm{RV}_t^{(d)} + \beta_w \mathrm{RV}_t^{(w)} + \beta_m \mathrm{RV}_t^{(m)} + er_{t+1} \tag{3.1.7}$$

where d is some constant, er_{t+1} is an error term, the source of randomness in the RV's dynamics and

$$\mathrm{RV}_t^{(d)} = \mathrm{RV}_t, \quad \mathrm{RV}_t^{(w)} = \frac{1}{4}\sum_{i=1}^{4} \mathrm{RV}_{t-i}, \quad \text{and} \quad \mathrm{RV}_t^{(m)} = \frac{1}{17}\sum_{i=5}^{21} \mathrm{RV}_{t-i}.$$

If we describe the source of randomness by non-central gamma distribution we obtain HARG process

$$\mathrm{RV}_{t+1}|\mathcal{F}_t \sim \bar{\gamma}(\delta, \Theta(\mathbf{RV}_t), \theta) \tag{3.1.8}$$

where δ and θ are shape and scale parameters, respectively, and location of the distribution is given by

$$\Theta(\mathbf{RV}_t) = \beta_d \mathrm{RV}_t^{(d)} + \beta_w \mathrm{RV}_t^{(w)} + \beta_m \mathrm{RV}_t^{(m)}. \tag{3.1.9}$$

HAR and HARG process reproduce volatility cascade effect - volatility over longer time intervals has stronger influence on those at shorter time intervals than conversely. The asymmetric propagation of volatility have been empirically confirmed by Müller et al. (1997) and Zumbach and Lynch (2001). The economic interpretation of this effect is that while short-term traders react to long-term volatility levels, long-term traders are not affected by short-term volatility levels.

3.1.3 Realized variance dynamics

An extension of the HARG-RV with a daily binary Leverage component (HARGL) was applied to option pricing by Corsi et al. (2013). The first main drawback of HARGL model is lack of closed-form solutions for option prices (pricing needs to be done via heavy Monte Carlo simulation). Another drawback of HARGL model is its too simple and unrealistic form of leverage - the importance of a heterogeneous structure for leverage is stressed by Corsi and Renò (2012). Thus we develop an Autoregressive Gamma model with Heterogeneous parabolic Leverage, and we name it the LHARG-RV model.

LHARG-RV belongs to the family of models described by (2.1.1)-(2.1.4) setting $k = 1$ and $\mathrm{f}_t = \mathrm{RV}_t$. Realized variance at time $t+1$ conditioned on information at day t is sampled from a non-centred gamma distribution

$$\mathrm{RV}_{t+1}|\mathcal{F}_t \sim \bar{\gamma}(\delta, \Theta(\mathbf{RV}_t, \mathbf{L}_t), \theta) \tag{3.1.10}$$

with

$$\Theta(\mathbf{RV}_t, \mathbf{L}_t) = d + \beta_d \mathrm{RV}_t^{(d)} + \beta_w \mathrm{RV}_t^{(w)} + \beta_m \mathrm{RV}_t^{(m)} + \alpha_d \ell_t^{(d)} + \alpha_w \ell_t^{(w)} + \alpha_m \ell_t^{(m)} . \tag{3.1.11}$$

In the previous equation $d \in \mathbb{R}$ is a constant and the quantities

$$\begin{aligned} \mathrm{RV}_t^{(d)} &= \mathrm{RV}_t, & \ell_t^{(d)} &= \left(\epsilon_t - \gamma\sqrt{\mathrm{RV}_t}\right)^2, \\ \mathrm{RV}_t^{(w)} &= \tfrac{1}{4}\textstyle\sum_{i=1}^{4} \mathrm{RV}_{t-i}, & \ell_t^{(w)} &= \tfrac{1}{4}\textstyle\sum_{i=1}^{4} \left(\epsilon_{t-i} - \gamma\sqrt{\mathrm{RV}_{t-i}}\right)^2, \\ \mathrm{RV}_t^{(m)} &= \tfrac{1}{17}\textstyle\sum_{i=5}^{21} \mathrm{RV}_{t-i}, & \ell_t^{(m)} &= \tfrac{1}{17}\textstyle\sum_{i=5}^{21} \left(\epsilon_{t-i} - \gamma\sqrt{\mathrm{RV}_{t-i}}\right)^2, \end{aligned} \tag{3.1.12}$$

correspond to the heterogeneous components associated with the short-term (daily), medium-term (weekly), and long-term (monthly) volatility and leverage factors, on the left and right columns respectively. The structure of leverage is analogous to the one in Heston and Nandi

(2000), and it is based on asymmetric influence of shock: large positive idiosyncratic component ϵ_t has a smaller impact on RV_{t+1} than large negative ϵ_t. As consequence the log-returns and variance process are negatively correlated:

$$\begin{aligned} Cov_{t-1}(y_t, \mathrm{RV}_{t+1}) &= -2\theta\alpha_d\gamma\mathbb{E}\left[\mathrm{RV}_t|\mathcal{F}_{t-1}\right] \\ &= -2\theta^2\alpha_d\gamma\left(\delta + \Theta(\mathbf{RV}_{t-1}, \mathbf{L}_{t-1})\right) . \end{aligned} \tag{3.1.13}$$

In order to adjust eq. (3.1.11) to our framework we rewrite $\Theta(\mathbf{RV}_t, \mathbf{L}_t)$ as

$$d + \sum_{i=1}^{22}\beta_i \mathrm{RV}_{t+1-i} + \sum_{j=1}^{22}\alpha_j\left(\epsilon_{t+1-j} - \gamma\sqrt{\mathrm{RV}_{t+1-j}}\right)^2 , \tag{3.1.14}$$

with

$$\beta_i = \begin{cases} \beta_d & \text{for } i = 1 \\ \beta_w/4 & \text{for } 2 \leq i \leq 5 \\ \beta_m/17 & \text{for } 6 \leq i \leq 22 \end{cases} \qquad \alpha_j = \begin{cases} \alpha_d & \text{for } j = 1 \\ \alpha_w/4 & \text{for } 2 \leq j \leq 5 \\ \alpha_m/17 & \text{for } 6 \leq j \leq 22 \end{cases} . \tag{3.1.15}$$

Crucial advantage of LHARG process is affinity, namely it satisfies Assumption 6.

Proposition 11. *For LHARG process the following relation holds true*

$$\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{zy_s + b\mathrm{RV}_s + c\ell_s}|\mathcal{F}_{s-1}\right] = \exp\left[\mathcal{A}(z,b,c) + \sum_{i=1}^{p}\mathcal{B}_i(z,b,c)\mathrm{RV}_{s-i} + \sum_{j=1}^{q}\mathcal{C}_i(z,b,c)\ell_{s-j}\right] , \tag{3.1.16}$$

where

$$\begin{aligned} \mathcal{A}(z,b,c) &= zr - \frac{1}{2}\ln(1-2c) - \delta\mathcal{W}(x,\theta) + d\mathcal{V}(x,\theta) , \\ \mathcal{B}_i(z,b,c) &= \mathcal{V}(x,\theta)\beta_i , \\ \mathcal{C}_j(z,b,c) &= \mathcal{V}(x,\theta)\alpha_j . \end{aligned} \tag{3.1.17}$$

The functions $\mathcal{V}$, $\mathcal{W}$ are defined as follows

$$\mathcal{V}(x,\theta) = \frac{\theta x}{1-\theta x}\,, \qquad \mathcal{W}(x,\theta) = \ln\left(1 - x\theta\right)\,,$$

and

$$x(z,b,c) = z\lambda + b + \frac{\frac{1}{2}z^2 + \gamma^2 c - 2c\gamma z}{1-2c}\,.$$

Proof: See Appendix A.1.

Then, the MGF for LHARG process reads

Proposition 12. *Under $\mathbb{P}$, the MGF for LHARG model has the following form*

$$\varphi^{\mathbb{P}}(t,T,z) = \mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{zy_{t,T}}|\mathcal{F}_t\right] = \exp\left(\mathrm{a}_t + \sum_{i=1}^{p}\mathrm{b}_{t,i}\mathrm{RV}_{t+1-i} + \sum_{j=1}^{q}\mathrm{c}_{t,j}\ell_{t+1-j}\right) \tag{3.1.18}$$

where

$$\begin{aligned}
\mathrm{a}_s &= \mathrm{a}_{s+1} + zr - \frac{1}{2}\ln(1-2\mathrm{c}_{s+1,1}) - \delta\mathcal{W}(\mathrm{x}_{s+1},\theta) + d\mathcal{V}(\mathrm{x}_{s+1},\theta)\\
\mathrm{b}_{s,i} &= \begin{cases} \mathrm{b}_{s+1,i+1} + \mathcal{V}(\mathrm{x}_{s+1},\theta)\beta_i & \textit{for } 1 \le i \le p-1\\ \mathcal{V}(\mathrm{x}_{s+1},\theta)\beta_i & \textit{for } i = p \end{cases}\\
\mathrm{c}_{s,j} &= \begin{cases} \mathrm{c}_{s+1,j+1} + \mathcal{V}(\mathrm{x}_{s+1},\theta)\alpha_j & \textit{for } 1 \le j \le q-1\\ \mathcal{V}(\mathrm{x}_{s+1},\theta)\alpha_j & \textit{for } j = q \end{cases}
\end{aligned} \tag{3.1.19}$$

with

$$\mathrm{x}_{s+1} = z\lambda + \mathrm{b}_{s+1,1} + \frac{\frac{1}{2}z^2 + \gamma^2\mathrm{c}_{s+1,1} - 2\mathrm{c}_{s+1,1}\gamma z}{1-2\mathrm{c}_{s+1,1}}\,.$$

and the terminal conditions read $\mathrm{a}_T = \mathrm{b}_{T,i} = \mathrm{c}_{T,j} = 0$ for $i = 1,\ldots,p$ and $j = 1,\ldots,q$.

Proof: It follows immediately by plugging expressions for $\mathcal{A}$, $\mathcal{B}_i$ and $\mathcal{C}_j$ (3.1.17) into recursive relations from Theorem 7.

Following the reasoning in Appendix F in Gourieroux and Jasiak (2006) one can derive the stationarity condition for RV_t process:

$$\theta\left(\beta_d + \beta_w + \beta_m + \gamma^2\left(\alpha_d + \alpha_w + \alpha_m\right)\right) < 1. \tag{3.1.20}$$

3.1.4 Risk-neutral dynamics

To preserve analytical tractability of the model under martingale measure we proceed a risk-neutralisation via Esscher transform suggested in Section (2.2), whose high flexibility allows to incorporate multiple factor-dependent risk-premia. For LHARG process the proposed transform takes the following form

$$M_{s,s+1} = \frac{\mathrm{e}^{-\nu_r \mathrm{RV}_{s+1} - \nu_y y_{s+1}}}{\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{-\nu_r \mathrm{RV}_{s+1} - \nu_y y_{s+1}} \middle| \mathcal{F}_s\right]}, \tag{3.1.21}$$

Esscher transform (3.1.21) has to satisfy no-arbitrage condition in order to be an SDF for LHARG model. The no-arbitrage condition for LHARG is a consequence of Theorem 8.

Proposition 13. *Esscher transform specified as in (3.1.21) is an SDF for LHARG model defined by eq.s (3.1.2) and (3.1.10)-(3.1.12) if, and only if*

$$\nu_y = \lambda + \frac{1}{2}. \tag{3.1.22}$$

Proof. The no-arbitrage condition follows from formulae (3.1.17) and relations (2.2.3) noticing that it is sufficient to impose

$$x(1 - \nu_2, -\nu_1, 0) = x(-\nu_2, -\nu_1, 0).$$

□

Proposition 14. *Under the risk-neutral measure $\mathbb{Q}$ the MGF for LHARG has the form*

$$\varphi^{\mathbb{Q}}(t,T,z) = \exp\left(\mathrm{a}_t^* + \sum_{i=1}^{p} \mathrm{b}_{t,i}^* \mathrm{RV}_{t+1-i} + \sum_{j=1}^{q} \mathrm{c}_{t,j}^* \ell_{t+1-j}\right),$$

where a_t^*, $\mathrm{b}_{t,i}^*$ *and* $\mathrm{c}_{t,j}^*$ *are given by recursive relations.*

Proof. It follows immediately by plugging expressions for $\mathcal{A}$, $\mathcal{B}_i$ and $\mathcal{C}_j$ (3.1.17) into recursive relations from Theorem 9. For recursive relations see (A.2.5). □

To derive the price of vanilla options, for example, it is sufficient to know the MGF under the risk-neutral measure $\mathbb{Q}$ which has been given in Proposition 14. However, for exotic instruments it is essential to know the log-return dynamics under $\mathbb{Q}$. The comparison of the physical and risk-neutral MGFs provides us the one-to-one mapping among the parameters which transforms the dynamics under $\mathbb{Q}$ into the dynamics under $\mathbb{P}$.

Proposition 15. *Under the risk-neutral measure* $\mathbb{Q}$ *the realized variance still follows a LHARG process with parameters*

$$\begin{aligned}
&\beta_d^* = \tfrac{1}{1-\theta \mathrm{y}^*}\beta_d\,, \quad \beta_w^* = \tfrac{1}{1-\theta \mathrm{y}^*}\beta_w\,, \quad \beta_m^* = \tfrac{1}{1-\theta \mathrm{y}^*}\beta_m\,,\\
&\alpha_d^* = \tfrac{1}{1-\theta \mathrm{y}^*}\alpha_d\,, \quad \alpha_w^* = \tfrac{1}{1-\theta \mathrm{y}^*}\alpha_w\,, \quad \alpha_m^* = \tfrac{1}{1-\theta \mathrm{y}^*}\alpha_m\,,\\
&\theta^* = \tfrac{1}{1-\theta \mathrm{y}^*}\theta\,, \qquad \delta^* = \delta\,, \qquad \gamma^* = \gamma + \lambda + \tfrac{1}{2}\,,\\
&d^* = \tfrac{1}{1-\theta \mathrm{y}^*}d\,,
\end{aligned} \tag{3.1.23}$$

where $\mathrm{y}^* = -\lambda^2/2 - \nu_1 + \frac{1}{8}$.

Proof: See Appendix A.2.2.

From the previous results we can write the simplified risk-neutral MGF which allows us to reduce the computational burden when computing the backward recurrences.

Corollary 16. *Under* $\mathbb{Q}$, *the MGF for the LHARG model has the same form as in (3.1.18)-(3.1.19) with equity risk premium* $\lambda^* = -0.5$ *and* d^*, δ^*, θ^*, γ^*, α_l^*, β_l^* *for* $l = d, w, m$ *as in (3.1.23).*

3.1.5 Particular cases

We now discuss two special cases of the model presented in the previous section. The first instance is the HARG model with Parabolic Leverage (P-LHARG) that we obtain setting $d = 0$ in (3.1.11), while the second model is a LHARG with zero-mean leverage (ZM-LHARG). The shape of the leverage in the latter has been inspired by the model of Christoffersen et al. (2008) but in the present context it is enriched by a heterogeneous structure

$$\begin{aligned}
\bar{\ell}_t^{(d)} &= \epsilon_t^2 - 1 - 2\epsilon_t \gamma \sqrt{\mathrm{RV}_t}\,, \\
\bar{\ell}_t^{(w)} &= \frac{1}{4}\sum_{i=1}^{4}\left(\epsilon_{t-i}^2 - 1 - 2\epsilon_{t-i}\gamma\sqrt{\mathrm{RV}_{t-i}}\right)\,, \\
\bar{\ell}_t^{(m)} &= \frac{1}{17}\sum_{i=5}^{21}\left(\epsilon_{t-i}^2 - 1 - 2\epsilon_{t-i}\gamma\sqrt{\mathrm{RV}_{t-i}}\right)\,.
\end{aligned}$$

In this case the expected value of leverage components is equal zero ($\mathbb{E}[\ell_t^{(k)}] = 0$ for $k = d, w, m$) and the linear $\Theta(\mathbf{RV}_t, \mathbf{L}_t)$ reads

$$\beta_d \mathrm{RV}_t^{(d)} + \beta_w \mathrm{RV}_t^{(w)} + \beta_m \mathrm{RV}_t^{(m)} + \alpha_d \bar{\ell}_t^{(d)} + \alpha_w \bar{\ell}_t^{(w)} + \alpha_m \bar{\ell}_t^{(m)}\,, \tag{3.1.24}$$

which can be reduced to the form (3.1.11) setting $d = -(\alpha_d + \alpha_w + \alpha_m)$, $\beta_l = \beta_l - \alpha_l \gamma^2$ for $l = d, w, m$. As will be more clear in the following section, the introduction of the less constrained leverage allows the process to explain a larger fraction of the skewness and kurtosis observed in real data. However, similarly to what has been discussed in Section 2.1 about Christoffersen et al. (2008), it is no more guaranteed that the non centrality parameter of the gamma distribution is positive definite. Nonetheless, in the Section 4.2 we will provide numerical evidence of the effectiveness of our analytical results in describing a regularised version of this model.

3.2 Multi-component GARCH models (k-CGARCH)

3.2.1 The model

The starting point of our considerations in this section is a GARCH option pricing model proposed by Heston and Nandi (2000), in which latent variance of log-returns is described by NGARCH model of Engle and Ng (1993):

$$h_{t+1} = d + mh_t + n\left(\epsilon_t - \gamma\sqrt{h_t}\right)^2 . \tag{3.2.1}$$

ARCH-GARCH models have been proven to be a good volatility predictors, hence it is natural to consider theirs application to option pricing. Heston and Nandi (2000) derive the closed-form solution for the price of a European call option. While Heston and Nandi (2000) use a single lag model, authors are suggesting to extend it with multiple lags to improve the pricing of long-term options. Another possible extension for the purpose of more accurate pricing long time to maturity options is to add long-run component which give rise to CGARCH proposed by Christoffersen et al. (2008). Volatility modelling with short-run and long-run components enables one to account for dependencies among volatilities at different time-scales. One can generalise this model to k component structure. In this section we introduce a class of GARCH models with k components and multiple lags which we label as k-CGARCH(p, q).

We define the k-dimensional vector of variance factors $\mathrm{h}_t^{(1)}, \ldots, \mathrm{h}_t^{(k)}$ which we shortly denote as $\mathbf{h}_t$. The variance on day t is defined as a sum of variance factors $\mathcal{S}\left(\mathbf{h}_t\right) = \mathrm{h}_t^{(1)} + \ldots + \mathrm{h}_t^{(k)}$ and the daily log-returns on day $t+1$ are modelled by equation

$$y_{t+1} = r + \lambda\, \mathcal{S}\left(\mathbf{h}_{t+1}\right) + \sqrt{\mathcal{S}\left(\mathbf{h}_{t+1}\right)}\, \epsilon_{t+1}\,, \tag{3.2.2}$$

where r is the risk-free rate, λ is the market price of risk, and ϵ_t are i.i.d. $\mathcal{N}\left(0, 1\right)$. We model

$\mathbf{h}_{t+1}$ as

$$\mathbf{h}_{t+1} = \mathbf{d} + \sum_{i=1}^{p} \mathbf{M}_i \mathbf{h}_{t+1-i} + \sum_{j=1}^{q} \mathbf{N}_j \boldsymbol{\ell}_{t+1-j}\,, \tag{3.2.3}$$

where $\mathbf{M}_i, \mathbf{N}_j \in \mathbb{R}^{k\times k}$ for $i = 1, \dots, p$ and $j = 1, \dots, q$, $\mathbf{d} \in \mathbb{R}^k$, and the vectors representing leverage effect $\boldsymbol{\ell}_{t-j}$ are of the form

$$\boldsymbol{\ell}_{t+1-j} = \begin{bmatrix} \left(\epsilon_{t+1-j} - \gamma_1 \sqrt{\mathcal{S}\left(\mathbf{h}_{t+1-j}\right)}\right)^2 \\ \vdots \\ \left(\epsilon_{t+1-j} - \gamma_k \sqrt{\mathcal{S}\left(\mathbf{h}_{t+1-j}\right)}\right)^2 \end{bmatrix}. \tag{3.2.4}$$

We prove that the family of k-CGARCH(p, q) processes satisfies the affine property (it satisfies Assumption 6).

Proposition 17. *There exist functions $\mathcal{A}, \boldsymbol{\mathcal{B}}_i, \boldsymbol{\mathcal{C}}_j$, $i \in \{1, \dots, p\}$ and $j \in \{2, \dots, q\}$ such that the following relation for the k-CGARCH(p, q) process is satisfied*

$$\mathbb{E}\left[\mathrm{e}^{zy_{s+1}+\mathbf{b}\cdot\mathbf{h}_{s+2}+\mathbf{c}\cdot\boldsymbol{\ell}_{s+1}} \middle| \mathcal{F}_s\right] = \mathrm{e}^{\mathcal{A}(z,\mathbf{b},\mathbf{c})+\sum_{i=1}^{p}\boldsymbol{\mathcal{B}}_i(z,\mathbf{b},\mathbf{c})\cdot\mathbf{h}_{s+2-i}+\sum_{j=2}^{q}\boldsymbol{\mathcal{C}}_j(z,\mathbf{b},\mathbf{c})\cdot\boldsymbol{\ell}_{s+1-j}}. \tag{3.2.5}$$

Proof: See Appendix A.3.1.

For k-CGARCH(p, q) processes the moment generating function is available in a closed form:

Proposition 18. *Under the physical measure $\mathbb{P}$ the MGF of the log-returns $y_{t,T} = \log(S_T/S_t)$ conditional on the information available at time t is of the form*

$$\varphi^{\mathbb{P}}(t, T, z) = \mathrm{e}^{\mathrm{a}_t+\sum_{i=1}^{p}\mathbf{b}_{t,i}\cdot\mathbf{h}_{t+2-i}+\sum_{j=2}^{q}\mathbf{c}_{t,j}\cdot\boldsymbol{\ell}_{t+1-j}}\,, \tag{3.2.6}$$

where a_t, $\mathbf{b}_{t,i}$ and $\mathbf{c}_{t,j}$ are given by recursive relations.

Proof. We take the form of functions $\mathcal{A}$, $\mathcal{B}_i$, $\mathcal{C}_j$ derived in Proposition 17 and apply them to Theorem 7. For the form of coefficients a_t, $\mathbf{b}_{t,i}$ and $\mathbf{c}_{t,j}$ see Appendix A.3.2. □

3.2.2 Risk-neutral dynamics

Change of measure is performed by applying two-dimensional Esscher transform[2]

$$M_{s,s+1} = \frac{\mathrm{e}^{-\nu_h \mathcal{S}(\mathbf{h}_{s+2}) - \nu_y y_{s+1}}}{\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{-\nu_h \mathcal{S}(\mathbf{h}_{s+2}) - \nu_y y_{s+1}} \middle| \mathcal{F}_s\right]}. \tag{3.2.7}$$

The no-arbitrage restriction can be formulated in the terms of the relation between risk-premia.

Proposition 19. *The Esscher transform (3.2.7) is an SDF for k-CGARCH(p, q) model if, and only if*

$$\nu_y = \lambda + \frac{1}{2} + 2\nu_h \sum_{i=1}^{k} \sum_{j=1}^{k} n_{i,j}(\gamma_j + \lambda), \tag{3.2.8}$$

where $n_{i,j}$ are elements of matrix $\mathbf{N}_1$.

Proof. See Appendix A.3.3.

□

From relation (3.2.8) one can see that in the case of one-dimensional pricing kernel ($\nu_h = 0$) employed in Heston and Nandi (2000) the equity risk premium parameter ν_y equals equity premium plus one half. Therefore all parameters of the option pricing model are fixed on the level of estimation from log-returns time series. In the case of two-dimensional pricing kernel, ν_h remains a free parameter that has to be calibrated on the option data time series. This allows the model to reconcile the time series properties of stock returns with option prices.

The knowledge of functions $\mathcal{A}$, $\mathcal{B}_i$, $\mathcal{C}_j$ for $i = 1, \ldots, p$ and $j = 1, \ldots, q$ allows to write the conditional moment generating function (see Theorem 9) needed to price vanilla contingent claims. However, the computation of the VRP requires the knowledge of the complete dynamics under measure $\mathbb{Q}$. Moreover the derivation risk-neutral dynamics reduces the computational burden in option pricing (likewise in the case of LHARG model).

[2] Theoretically one could propose the multi-dimensional, factor-dependent pricing kernel, but as it will become clear later (see discussion in Appendix A.3.4) it would rise issue of identification problem.

Proposition 20. *Under the risk-neutral measure $\mathbb{Q}$, obtained with SDF given by (3.2.7), the dynamics of log-returns for k-CGARCH(p, q) model is still governed by equations (3.2.2)-(3.2.3) with parameters*

$$\begin{aligned}
\lambda^* &= -1/2, \\
\mathbf{d}^* &= \frac{\mathbf{d}}{1+2\nu_h \sum_{i=1}^k \sum_{j=1}^k n_{i,j}}, \\
\mathbf{M_i}^* &= \mathbf{M_i} \text{ for } 1 \leq i \leq p \\
\mathbf{N_i}^* &= \frac{\mathbf{N_i}}{\left(1+2\nu_h \sum_{i=1}^k \sum_{j=1}^k n_{i,j}\right)^2} \text{ for } 1 \leq i \leq q, \\
\gamma_l^* &= \gamma_l + \nu_y + 2\nu_h \sum_{i=1}^k \sum_{j=1}^k n_{i,j}(\gamma_l - \gamma_j) \text{ for } 1 \leq l \leq k.
\end{aligned} \tag{3.2.9}$$

The relation between the dynamics of the process under physical and risk-neutral measure is described by equations:

$$y_t = r - \frac{1}{2}\mathcal{S}\left(\mathbf{h}_t^*\right) + \epsilon_t^* \sqrt{\mathcal{S}\left(\mathbf{h}_t^*\right)}, \tag{3.2.10}$$

$$\mathbf{h}_t^* = \frac{\mathbf{h}_t}{1+2\nu_h \sum_{i=1}^k \sum_{j=1}^k n_{i,j}}. \tag{3.2.11}$$

Proof. See Appendix A.3.4. □

Given the dynamics under $\mathbb{Q}$, the risk-neutral moment generating function is a straightforward consequence of Proposition 18.

Corollary 21. *Under $\mathbb{Q}$, the MGF for the k-CGARCH(p, q) model has the same form as in (3.2.6) with parameters of the process λ^*, $\mathbf{d}^*$, $\mathbf{M}^*$, $\mathbf{N}^*$, $\boldsymbol{\gamma}^*$ as in (3.2.9).*

Equation (3.2.11) provides a clear interpretation of the risk-premia appearing in the SDF (3.2.7). We first observe that reducing the dimensionality of the Esscher transform to one (by setting $\nu_h = 0$) implies that the volatility process under the two measures remains the same. When ν_h

is nonzero then the risk-neutral and physical volatilities differ and their ratio reads

$$\xi = \frac{\sqrt{\mathcal{S}\left(\mathbf{h^*}_t\right)}}{\sqrt{\mathcal{S}\left(\mathbf{h}_t\right)}} = \left(1 + 2\nu_h \sum_{i=1}^{k}\sum_{j=1}^{k} n_{i,j}\right)^{-1/2} . \tag{3.2.12}$$

It is worth noticing that our specification of the pricing kernel implies a constant volatility ratio, which is mainly determined by the volatility risk premium ν_h.

3.2.3 The log-ratio of the risk-neutral and physical densities

Early option pricing literature (for example Rubinstein (1976) and Brennan (1979)) implicitly assumes the existence of a monotonic relation between the risk-neutral and physical densities log-ratio and market returns. However, in recent empirical studies it has been shown that the ratio has a parabolic shape with a positive smile (see Bakshi et al. (2010)). As pointed out by Christoffersen et al. (2013), a premium for the variance risk explains a number of puzzles concerning the level and movement of implied option variance compared with observed time series variance. The key feature of their modelling approach is that, although the pricing kernel is monotonic on both returns and variance, the projection of the pricing kernel onto the stock price return alone is U-shaped. The strong option smile associated to this non-monotonic relation can be quantified looking at the natural logarithm of the ratio of the risk-neutral and physical conditional densities – f and f^*, respectively – implied by the model

$$\ln\left(f^*(y_t|\mathcal{L}\left(\mathbf{f}_t\right))/f(y_t|\mathcal{L}\left(\mathbf{f}_t\right))\right) . \tag{3.2.13}$$

The parabolic shape of the log-ratio (3.2.13) for SDF (3.2.7) and multi-component GARCH models readily follows noticing that $y_t|\mathcal{L}\left(\mathbf{f}_t\right) \sim \mathcal{N}(r + \lambda\mathcal{L}\left(\mathbf{f}_t\right), \mathcal{L}\left(\mathbf{f}_t\right))$ under measure $\mathbb{P}$ and $y_t|\mathcal{L}\left(\mathbf{f^*}_t\right) \sim \mathcal{N}(r - \frac{1}{2}\mathcal{L}\left(\mathbf{f^*}_t\right), \mathcal{L}\left(\mathbf{f^*}_t\right))$ under measure $\mathbb{Q}$. Knowing that $\mathcal{L}\left(\mathbf{f^*}_t\right) = \xi^2\mathcal{L}\left(\mathbf{f}_t\right)$ we obtain the following corollary (for details see Appendix A.3.5).

Corollary 22. *The logarithm of the ratio of the risk-neutral and physical conditional densities* *is a quadratic function of the log-return*

$$\ln\left(\frac{f^*(y_t|\mathcal{L}(\mathbf{f^*}_t))}{f(y_t|\mathcal{L}(\mathbf{f}_t))}\right) = \frac{\xi^2-1}{2\mathcal{L}(\mathbf{f}_t)\xi^2}(y_t-r)^2 - \left(\lambda+\frac{1}{2}\right)(y_t-r) + \frac{4\lambda^2-\xi^2}{8}\mathcal{L}(\mathbf{f}_t) - \ln\xi. \quad (3.2.14)$$

From Corollary 22 we can infer the importance of the ratio ξ. When $\xi = 1$ ($\nu_h = 0$), then the log-ratio (3.2.13) becomes a linear decreasing function of log-returns. Whereas for values greater than 1 ($\nu_h < 0$) the relation (3.2.14) becomes U-shaped, consistently with empirical observations.

3.3 Combination of latent and realized volatility (GARCH-LHARG-RV)

3.3.1 The model

Our general framework allows us to incorporate a model being a combination of realized volatility and latent volatility. Measure of RV applied in Section 3.1 is a very precise measure of continuous part of volatility and its dynamics can be modelled accurately by LHARG process. Though it does not take into account volatility due to jumps and to overnight effect. In this section we will add to RV modelled with LHARG, a parallel factor of latent volatility which we will model with GARCH process and we label the complete model GARCH-LHARG-RV.

GARCH-LHARG-RV is described by general framework (2.1.1)-(2.1.4) setting $k = 2$, $\mathbf{f}_t = (h_{t+1}, \mathrm{RV}_t)$ and $\mathcal{L}(\mathbf{f}_{t-1}, \mathbf{f}_t) = \mathrm{f}_{t-1}^{(1)} + \mathrm{f}_t^{(2)} = h_t + \mathrm{RV}_t$. Thus, log-returns evolve according to the equation

$$y_{t+1} = r + \lambda(\mathrm{RV}_{t+1} + h_{t+1}) + \sqrt{\mathrm{RV}_{t+1} + h_{t+1}}\epsilon_{t+1}\,. \quad (3.3.1)$$

To model dynamics of latent variance h we adopt non-linear GARCH model of Heston and

Nandi (2000), but in our case leverage depends on both components of volatility:

$$h_{t+1} = \omega + \beta_h h_t + \alpha_h \left(\epsilon_t - \gamma_h \sqrt{\mathrm{RV}_t + h_t}\right)^2. \tag{3.3.2}$$

Analogously dynamics of RV is described by LHARG process with leverage component depending on both volatility components:

$$\mathrm{RV}_{t+1}|\mathcal{F}_t \sim \bar{\gamma}(\delta, \Theta(\mathbf{RV}_t, \mathbf{L}_t), \theta) \tag{3.3.3}$$

with

$$\Theta(\mathbf{RV}_t, \mathbf{L}_t) = \beta_d \mathrm{RV}_t^{(d)} + \beta_w \mathrm{RV}_t^{(w)} + \beta_m \mathrm{RV}_t^{(m)} + \alpha_d \bar{\ell}_t^{(d)} + \alpha_w \bar{\ell}_t^{(w)} + \alpha_m \bar{\ell}_t^{(m)}. \tag{3.3.4}$$

In previous equation $d \in \mathbb{R}$ is a constant and the quantities

$$\begin{aligned}
\mathrm{RV}_t^{(d)} &= \mathrm{RV}_t, & \bar{\ell}_t^{(d)} &= \epsilon_t^2 - 1 - 2\gamma\epsilon_t\sqrt{\mathrm{RV}_t + h_t},\\
\mathrm{RV}_t^{(w)} &= \tfrac{1}{4}\textstyle\sum_{i=1}^{4} \mathrm{RV}_{t-i}, & \bar{\ell}_t^{(w)} &= \tfrac{1}{4}\textstyle\sum_{i=1}^{4}\left(\epsilon_{t-i}^2 - 1 - \gamma\epsilon_{t-i}\sqrt{\mathrm{RV}_{t-i} + h_{t-i}}\right),\\
\mathrm{RV}_t^{(m)} &= \tfrac{1}{17}\textstyle\sum_{i=5}^{21} \mathrm{RV}_{t-i}, & \bar{\ell}_t^{(m)} &= \tfrac{1}{17}\textstyle\sum_{i=5}^{21}\left(\epsilon_{t-i}^2 - 1 - 2\gamma\epsilon_{t-i}\sqrt{\mathrm{RV}_{t-i} + h_{t-i}}\right),
\end{aligned} \tag{3.3.5}$$

correspond to the heterogeneous components associated to the short-term (daily), medium-term (weekly), and long-term (monthly) volatility and leverage factors, on the left and right column respectively. In this thesis we consider only zero-mean leverage case ($\mathbb{E}[\ell_t^{(k)}] = 0$ for $k = d, w, m$).[3]

It can be shown that that GARCH-LHARG-RV model satisfies Assumption 6. Then, the MGF can be obtained easily from Theorem 7.

[3]Similarly to ZM-LHARG and CGARCH processes, positivity of $\Theta(\mathbf{RV}_t, \mathbf{L}_t)$ and of volatility is no more guaranteed, but it can be justified by approximation analysis in Section 4.2.

Proposition 23. *Under the physical measure $\mathbb{P}$ the MGF for GARCH-LHARG has the form*

$$\varphi^{\mathbb{P}}(t,T,z) = \exp\left(\mathrm{a}_t + \mathrm{b}_t^h h_{t+1} + \sum_{i=1}^{22} \mathrm{b}_{t,i}^r \mathrm{RV}_{t+1-i} + \sum_{j=1}^{22} \mathrm{c}_{t,j} \ell_{t+1-j}\right),$$

where a_t, b_t^h, $\mathrm{b}_{t,i}^r$, $\mathrm{c}_{t,j}$ *are given by recursive relations and* $\ell_t = \left(\epsilon_t - \gamma\sqrt{\mathrm{RV}_t + h_t}\right)^2$.

Proof. See Appendix A.4.1. □

3.3.2 Risk-neutralisation

To derive the pricing measure $\mathbb{Q}$ we apply three-dimensional Esscher transform

$$M_{s,s+1} = \frac{\mathrm{e}^{-\nu_r \mathrm{RV}_{s+1} - \nu_h h_{s+2} - \nu_y y_{s+1}}}{\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{-\nu_r \mathrm{RV}_{s+1} - \nu_h h_{s+2} - \nu_y y_{s+1}} \middle| \mathcal{F}_s\right]}, \tag{3.3.6}$$

where $\nu_r, \nu_h, \nu_y \in \mathbb{R}$ are parameters of the transform. The main advantage of the above change of measure is that it clearly identifies the sources of risk and explicitly compensate them with separated risk-premia. The parameter ν_r corresponds to risk related with continuous part of realized variance and ν_h to the remaining, latent part of daily volatility.

The derivation of the no-arbitrage condition for GARCH-LHARG readily follows from the Proposition 8.

Proposition 24. *Esscher transform (3.3.6) is an SDF in a setting described by equations (3.3.1)-(3.3.5) if, and only if*

$$\nu_y = \lambda + \frac{1}{2} + 2\nu_h \alpha_h (\gamma_h + \lambda). \tag{3.3.7}$$

Proof. The no-arbitrage condition follows from formulae (A.4.7) and relations (2.2.3). □

Proposition 25. *Under the risk-neutral measure $\mathbb{Q}$ the MGF for GARCH-LHARG-RV has* *the form*

$$\varphi^{\mathbb{Q}}(t,T,z) = \exp\left(\mathrm{a}_t^* + \sum_{i=1}^{p} \mathrm{b}_{t,i}^* \mathrm{RV}_{t+1-i} + \sum_{j=1}^{q} \mathrm{c}_{t,j}^* \ell_{t+1-j}\right),$$

where a_t^*, $\mathrm{b}_{t,i}^*$ *and* $\mathrm{c}_{t,j}^*$ *are given by recursive relations.*

Proof. It follows immediately by plugging expressions for $\mathcal{A}$, $\mathcal{B}_i$ and $\mathcal{C}_j$ (A.4.7) into recursive relations from Theorem 9. For recursive relations see Appendix A.4.2. □

3.4 Jump component of realized variance (JLHARG-RV)

3.4.1 The model

In this section we employ the measurement of jump component of RV instead of modelling latent volatility component h_t with GARCH process like we did in Section 3.3. Under the assumption of the continuity of price process, quadratic variation of log-price is equal to integrated variance, like we had in Section 3.1. In fully generality, in the presence of jumps, the total quadratic variation of a log-price process has another component - squared jump variation. In this section we take into account both components of quadratic variation and we assume the following dynamics of log-returns

$$y_t = r + \lambda\left(\mathrm{RV}_t^c + \mathrm{RV}_t^j\right) + \sqrt{\mathrm{RV}_t^c + \mathrm{RV}_t^j}\,\epsilon_t. \tag{3.4.1}$$

where r is the risk-free rate, λ is the market price of risk, ϵ_t are i.i.d. with standard normal random distribution, RV_t^c is continuous component of RV and RV_t^j is jump component of RV (details on the RV measure employed in the implementation of the model are given in Section 4.2).

Dynamics proposed in equation (3.4.1) may be justified by the empirical studies of Andersen et al. (2001a), who find that the distributions of daily equity returns standardized by the corresponding RV is approximately Gaussian and Andersen et al. (2010) who investigate the deviation from normality ascribed to a jump component in the price process. The latter results indicate that the discontinuous component has a minor impact on the distributional properties, since the jump-adjusted standardized series are not systematically closer to the Gaussian than the $y_t/\sqrt{RV_t}$ standardized returns.[4] This is especially true for time series generated from futures contracts on the S&P500 Index, which are recognized in Andersen et al. (2010) to suffer from minimal microstructure distortion and low liquidity effects. As can be seen from the density plots of Figure 3.1, we observe the same feature for the S&P500 Futures in our sampling period. The two-sample Kolmogorov-Smirnov test between the RV standardized and jump-adjusted series indicates that the two distributions cannot be distinguished. If any, by judging on the value of the kurtosis of 3.64 for the jump-adjusted distribution and 3.06 for the RV standardized, we conclude that the latter is closer to a normal distribution than the former one.

The dynamics of the realized volatility components is given by sampling at time $t+1$ a new realisation from two distributions conditionally independent given the information at time t. The continuous part of RV depends on past realisations of RV^c and on past realisations of leverage term ℓ_t which is a quadratic function of the total realized variance thus including the contributions from the jumps

$$\bar{\ell}_t = \epsilon_t^2 - 1 - 2\gamma\epsilon_t\sqrt{\mathrm{RV}_t^c + \mathrm{RV}_t^j}. \tag{3.4.2}$$

We introdcue notation $\mathbf{RV^c}_t = (\mathrm{RV}_{t-21}^c, \dots, \mathrm{RV}_t^c)$ and $\mathbf{L}_t = (\bar{\ell}_{t-21}, \dots, \bar{\ell}_t)$. Then the continuous

[4] "*Perhaps surprisingly, the results indicate that neither of the jump-adjusted standardized series are systematically closer to Gaussian than the non-adjusted realized volatility standardized returns. [...] One reason is that jumps largely self-standardize: a large jump tends to inflate the (absolute) value of both the return (numerator) and the realized volatility (denominator) of standardized returns, so the impact is muted.*" - Andersen et al. (2010)

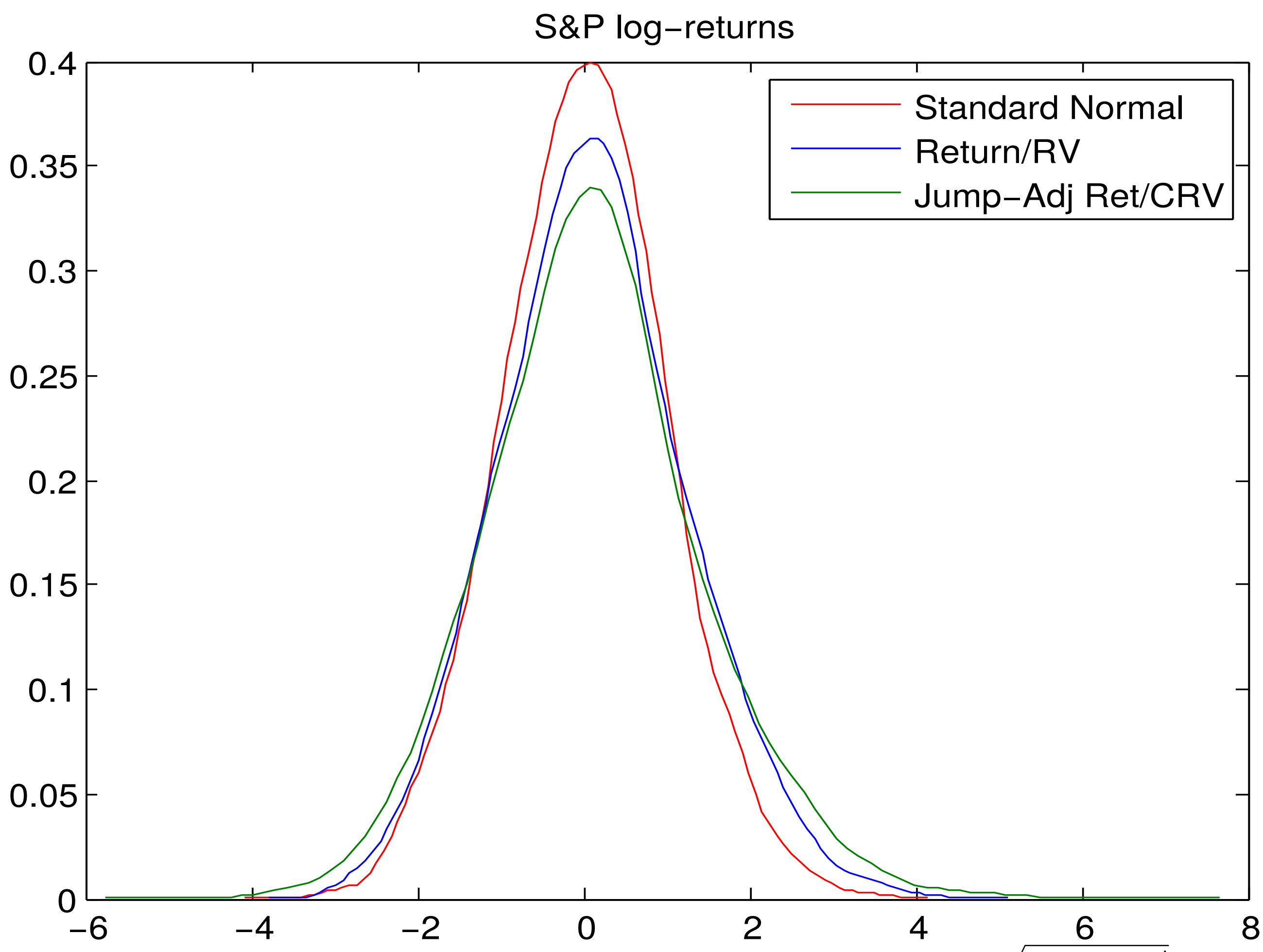

Figure 3.1: Histogram of returns rescaled by total realized volatility ($y_t/(\sqrt{\mathrm{RV}^c + \mathrm{RV}^j})$), histogram of returns purified from jumps rescaled by continuous component of realized volatility ($\tilde{y}_t/(\sqrt{\mathrm{RV}^c})$, where $\tilde{y}_t$ are returns without jumps) and standard normal distributition.

component of RV is drawn from a non-central gamma distribution

$$\mathrm{RV}^c_{t+1}|\mathcal{F}_t \sim \bar{\gamma}(\delta, \Theta(\mathbf{RV^c}_t, \mathbf{L}_t), \theta)\,, \tag{3.4.3}$$

where δ is the shape parameter, θ is the scale and the non-centrality is given by

$$\Theta(\mathbf{RV^c}_t, \mathbf{L}_t) = \beta_d \mathrm{RV}^{c\,(d)}_t + \beta_w \mathrm{RV}^{c\,(w)}_t + \beta_m \mathrm{RV}^{c\,(m)}_t + \alpha_d \bar{\ell}^{(d)}_t + \alpha_w \bar{\ell}^{(w)}_t + \alpha_m \bar{\ell}^{(m)}_t\,. \tag{3.4.4}$$

where $\beta_i \in \mathbb{R}^+$, $\alpha_i \in \mathbb{R}^+$ are constant and the quantities

$$\begin{array}{ll} \mathrm{RV}^{c\,(d)}_t = \mathrm{RV}^c_t, & \bar{\ell}^{(d)}_t = \bar{\ell}_t\,, \\ \mathrm{RV}^{c\,(w)}_t = \frac{1}{4}\sum_{i=1}^{4} \mathrm{RV}^c_{t-i}, & \bar{\ell}^{(w)}_t = \frac{1}{4}\sum_{i=1}^{4} \bar{\ell}_{t-i}\,, \\ \mathrm{RV}^{c\,(m)}_t = \frac{1}{17}\sum_{i=5}^{21} \mathrm{RV}^c_{t-i}, & \bar{\ell}^{(m)}_t = \frac{1}{17}\sum_{i=5}^{21} \bar{\ell}_{t-i} \end{array} \tag{3.4.5}$$

represent the heterogeneous components corresponding to the short-term or daily (d), medium-term or weekly (w) and long-term or monthly (m) realized variance and leverage terms, respectively on the left and right columns above. In this thesis we consider only zero-mean leverage version of JLHARG-RV model ($\mathbb{E}[\ell^{(k)}_t] = 0$ for $k = d, w, m$).[5] A positive leverage version of JLHARG model is discussed in Alitab et al. (2015).

The jump component of the realized variance is instead modelled as a compound Poisson process with intensity $\tilde{\Theta}$ and sizes sampled from a gamma distribution with shape $\tilde{\delta}$ and scale $\tilde{\theta}$

$$\mathrm{RV}^j_{t+1}|\mathcal{F}_t \sim \sum_{i=1}^{n_{t+1}} Y_i \quad \text{with} \quad n_{t+1} \sim \mathcal{P}(\tilde{\Theta}) \quad \text{and} \quad Y_i \text{ i.i.d.} \sim \gamma(\tilde{\delta}, \tilde{\theta}). \tag{3.4.6}$$

Equations (3.4.1)-(3.4.6) completely characterise the dynamics of log-returns by Autoregressive Gamma model in Realized Volatility with heterogeneous leverage and jumps and we acronym

[5]Similarly to ZM-LHARG and CGARCH processes, positivity of $\Theta(\mathbf{RV}_t, \mathbf{L}_t)$ and of volatility is no more guaranteed, but it can be justified by approximation analysis in Section 4.2.

it JLHARG-RV model. The crucial advantage of JLHARG process is that it is affine, namely Assumption 6 is satisfied for some functions $\mathcal{A}$, $\mathcal{B}_i$ and $\mathcal{C}_j$. Knowing the form of the functions one can prove the following

Proposition 26. *Under $\mathbb{P}$, the MGF of the log-return $y_{t,T} = \sum_{k=t+1}^{T} y_k$ for JLHARG model has the following form*

$$\phi^{\mathbb{P}}(t,T,z) = \mathbb{E}^{\mathbb{P}}\left[e^{zy_{t,T}}|\mathcal{F}_t\right] = \exp\left(\mathrm{a}_t + \sum_{i=1}^{p} \mathrm{b}_{t,i}\mathrm{RV}^c_{t+1-i} + \sum_{i=1}^{q} \mathrm{c}_{t,i}\ell_{t+1-i}\right) \tag{3.4.7}$$

where a_t, $\mathrm{b}_{t,i}$ and $\mathrm{c}_{t,i}$ are given by recursive relations.

Proof. See Appendix A.5.1. □

3.4.2 Risk-neutralisation

To proceed change of measure we apply Esscher transform whose high flexibility allows to incorporate multiple factor-dependent risk-premia:

$$M_{s,s+1} = \frac{e^{-\nu_r \mathrm{RV}^c_{s+1} - \nu_j \mathrm{RV}^j_{s+1} - \nu_y y_{s+1}}}{\mathbb{E}^{\mathbb{P}}\left[e^{-\nu_r \mathrm{RV}^c_{s+1} - \nu_j \mathrm{RV}^j_{s+1} - \nu_y y_{s+1}}|\mathcal{F}_s\right]}. \tag{3.4.8}$$

Specifically, it allows to take into account both variance risk premia components: continuous (ν_r) and jump (ν_j), in addition to the standard equity premium (ν_y). Esscher transform (3.4.8) has to satisfy no-arbitrage condition in order to be an SDF.

Proposition 27. *The Esscher transform (3.4.8) is an SDF for JLHARG model if and only if*

$$\nu_y = \lambda + \frac{1}{2}.$$

Proof. The no-arbitrage condition follows from formulae (A.5.13) and relations (2.2.3). □

An advantage of SDF (3.4.8) is that under risk-neutral measure the dynamics of the log-returns is still given by JLHARG process with mapped parameters. Moreover we are able to provide a one-to-one mapping of parameters from $\mathbb{P}$ to $\mathbb{Q}$ dynamics.

Proposition 28. *Under risk-neutral measure $\mathbb{Q}$ the realized variance follows a JLHARG process with parameters*

$$\begin{aligned}
\beta_d^* &= \frac{\beta_d}{1-\theta \mathrm{y}^{c*}},\ \beta_w^* = \frac{\beta_w}{1-\theta \mathrm{y}^{c*}},\ \beta_m^* = \frac{\beta_m}{1-\theta \mathrm{y}^{c*}},\\
\alpha_d^* &= \frac{\alpha_d}{1-\theta \mathrm{y}^{c*}},\ \alpha_w^* = \frac{\alpha_w}{1-\theta \mathrm{y}^{c*}},\ \alpha_m^* = \frac{\alpha_m}{1-\theta \mathrm{y}^{c*}},\\
\theta^* &= \frac{\theta}{1-\theta \mathrm{y}^{c*}},\ \delta^* = \delta,\ \gamma^* = \gamma + \lambda + \frac{1}{2},\\
d^* &= \frac{d}{1-\theta \mathrm{y}^{c*}}\\
\tilde{\Theta}^* &= \frac{\tilde{\Theta}}{\left(1-\tilde{\theta} \mathrm{y}^{j*}\right)^{\tilde{\delta}}},\ \tilde{\delta}^* = \tilde{\delta},\ \tilde{\theta}^* = \frac{\tilde{\theta}}{1-\tilde{\theta} \mathrm{y}^{j*}},
\end{aligned} \tag{3.4.9}$$

where $\mathrm{y}^{c*} = -\lambda^2/2 - \nu_r + \frac{1}{8}$ *and* $\mathrm{y}^{j*} = -\lambda^2/2 - \nu_j + \frac{1}{8}$.

Proof: Appendix A.5.2.

Knowing the dynamics of process under $\mathbb{Q}$, moment generating function under risk-neutral measure is a straightforward consequence of Proposition 26.

Corollary 29. *Under $\mathbb{Q}$ the MGF of the JLHARG model is formally the same as in Proposition 26 with equity risk premium parameter $\lambda^* = -0.5$ and d^*, δ^*, θ^*, $\tilde{\Theta}^*$, $\tilde{\delta}^*$, $\tilde{\theta}^*$ γ^*, α_l^*, β_l^* for $l = d, w, m$ as in (3.4.9).*

We point out that the risk premia parameters ν_r and ν_j need to be calibrated on option data. All the parameters governing the dynamic of the process under $\mathbb{Q}$, can be explicitly computed through the set of equation (3.4.9) from those estimated under $\mathbb{P}$ once (ν_r, ν_j) has been calibrated.

Chapter 4

Option Pricing

4.1 Estimation of realized variance

In this section we describe the measurement of realized variance employed in RV models (LHARG-RV, GARCH-LHARG-RV, JLHARG-RV). For this family of stochastic volatility models, we employ the RV computed from tick-by-tick data for the S&P 500 Futures, from January 1, 1990 to December 31, 2007. The choice of future contracts is for the sake of their high liquidity - while the S&P 500 index is not exchanged directly, futures are traded extensively. Moreover time series generated from futures contracts on the S&P 500 Index suffer from minimal microstructure noise.[1]

Our RV measurement procedure is based on estimating total quadratic variation of log-prices using the Two-Scale estimator introduced by Zhang et al. (2005) (with a fast scale of two ticks and a slower one of 20 ticks). Such a proxy of total quadratic variation includes jumps in both returns and volatility. To identify the jump component we apply two step procedure. In the first step we proceed the Threshold Bipower variation method with a significance level of 99% by Corsi et al. (2010) which detects the spikes in RV time series and we remove it from

[1] See Andersen et al. (2010).

RV time series. In the second step we remove the most extreme observations in the remaining RV series, seemingly due to volatility jumps, employing a threshold-based jumps detection method: we set a four standard deviations threshold computed on a rolling window of 200 days.

The purified RV series are our proxy of the integrated variance ($\mathrm{IV_t}$). In models LHARG-RV and GARCH-LHARG-RV we use it as a measure of RV and we label it RV_t. We label the same quantity as RV_t^c (continuous component of RV) in JLHARG-RV model. The difference between total quadratic variation and continuous component of RV is called jump component of RV and it is labeled as RV_t^j. On Figure 4.1 we plot the time series of continuous and jump components of RV.

Both RV measures (RV_t^c and RV_t^j) are proxies of volatility during the trading period, i.e., from open to close. As a result, they neglects the contribution coming from overnight returns. To overcome this problem for models without GARCH component (LHARG-RV and JLHARG-RV), we rescale our RV estimator to match the unconditional mean of the squared daily (i.e., close-to-close) returns. In the case of GARCH-LHARG-RV model overnight effect is captured by GARCH component.

4.2 Estimation of models under physical probability measure

We choose the FED Fund rate as proxy for the risk-free rate r in all considered models. For empirical assessment we will apply two-component GARCH model by Christoffersen et al. (2008) with dynamics (2.1.7) and we will apply to it a new SDF proposed in (3.2.7). The estimation of parameters under physical measure of CGARCH models is done by maximum likelihood estimation (MLE) used by Bollerslev (1996) and others.

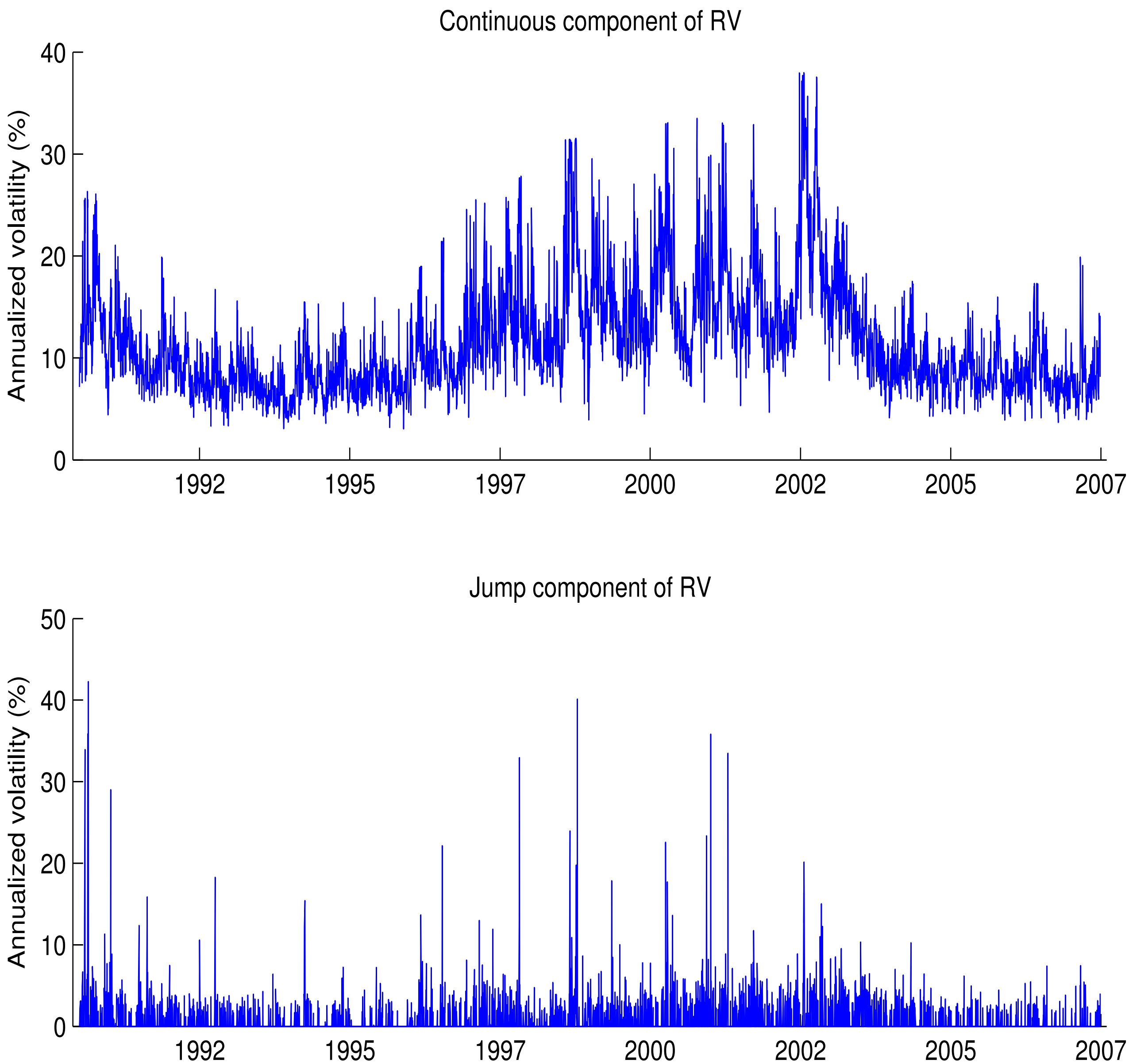

Figure 4.1: Time series of RV^c and RV^j.

The estimation of the parameters in LHARG-RV model is greatly simplified by the use of Realized Volatility, which avoids any filtering procedure related to latent volatility processes. Firstly we determine the market price of risk λ in equation (3.1.2) regressing the centred and normalised log-return on the realized volatility. This regression is performed by rewriting the equation (3.1.2) as

$$\frac{y_{t+1} - r}{\sqrt{\mathrm{RV}_{t+1}}} = \lambda\sqrt{\mathrm{RV}_{t+1}} + \epsilon_{t+1}, \tag{4.2.1}$$

The use of an RV proxy for the unobservable volatility allows us simply to employ a Maximum Likelihood Estimator (MLE) on historical data. Arguing as in Gourieroux and Jasiak (2006), the conditional transition density for the LHARG-RV family is available in closed-form, and so the log-likelihood reads

$$l_t^T(\delta, \theta, d, \beta_d, \beta_w, \beta_m, \alpha_d, \alpha_w, \alpha_m, \gamma) = \\ -\sum_{t=1}^{T}\left(\frac{\mathrm{RV}_t}{\theta} + \Theta\left(\mathbf{RV}_{t-1}, \mathbf{L}_{t-1}\right)\right) + \sum_{t=1}^{T}\log\left(\sum_{k=1}^{\infty}\frac{\mathrm{RV}_t^{\delta+k-1}}{\theta^{\delta+k}\Gamma(\delta+k)}\frac{\Theta\left(\mathbf{RV}_{t-1}, \mathbf{L}_{t-1}\right)^k}{k!}\right)$$

where $\Theta\left(\mathbf{RV}_{t-1}, \mathbf{L}_{t-1}\right)$ is given in eq. (3.1.11). To implement the MLE, we truncate the infinite sum on the right hand side to the $90th$ order as done in Corsi et al. (2013).

In the case of JLHARG-RV model we have two time series for the RV components $(\mathrm{RV}_t^{(c)}, \mathrm{RV}_t^{(j)})$ and again proceed the estimation via Maximum Likelihood Estimator. According to the model specified in equation (3.4.3) and (3.4.6), the log-likelihood functions for the continuous and jump RV components, respectively $l^c_{t,T}$ and $l^j_{t,T}$, are given by the following series-expansions

$$l^c_{t,T}\left(\delta,\theta,d,\beta_d,\beta_w,\beta_m,\alpha_d,\alpha_w,\alpha_m,\gamma\right) = -\sum_{t=1}^{T}\left(\frac{\mathrm{RV}^c_t}{\theta} + \Theta\left(\mathbf{RV}_{t-1},\mathbf{L}_{t-1}\right)\right) + \sum_{t=1}^{T}\log\left(\sum_{k=1}^{\infty}\frac{\left(\mathrm{RV}^c_t\right)^{\delta+k-1}}{\theta^{\delta+k}\Gamma\left(\delta+k\right)}\frac{\Theta\left(\mathbf{RV}_{t-1},\mathbf{L}_{t-1}\right)^k}{k!}\right) \tag{4.2.2}$$

$$l^j_{t,T}\left(\tilde{\delta},\tilde{\theta},\tilde{\Theta}\right) = -\sum_{t=1}^{T}\left(\frac{\mathrm{RV}^j_t}{\tilde{\theta}} + \tilde{\Theta}\right) + \sum_{t=1}^{T}\log\left(\sum_{k=1}^{\infty}\frac{\left(\mathrm{RV}^j_t\right)^{k\tilde{\delta}-1}}{\theta^{k\tilde{\delta}}\Gamma\left(k\tilde{\delta}\right)}\frac{\tilde{\Theta}^k}{k!}\right). \tag{4.2.3}$$

We truncate the infinite sum on the right hand side in both log-likelihoods to the $90th$ order similarly to LHARG-RV estimation.

In the case of GARCH-LHARG-RV model, in the first step we apply MLE to estimate λ and parameters of GARCH part of the model given the returns and time series of continuous realized volatility $\mathrm{RV_t}$. Next, having the parameters of latent variance process dynamics, we filter out the time series of h_t and finally, we apply MLE to estimate the parameters of the LHARG-RV part of the model.

For the sake of completeness we also estimate HARGL-RV presented in Corsi et al. (2013). THe dynamics of HARGL differs from LHARG by the non-central parameter in gamma distribution which in case of HARGL is equal

$$\Theta(\mathbf{RV}_t) = \beta_d\mathrm{RV}^{(d)}_t + \beta_w\mathrm{RV}^{(w)}_t + \beta_m\mathrm{RV}^{(m)}_t + \alpha_d I_{(y_t<0)}\mathrm{RV}^{(d)}_t\,, \tag{4.2.4}$$

where $I_{(y_t<0)}$ takes value one if the log-return at date t is negative and takes value zero otherwise. In this way we lose analytical tractability of the model and heterogeneity of leverage component. The estimation procedure for the HARGL model can be found in Corsi et al. (2013).

	Model					
Param.	HARGL	P-LHARG	ZM-LHARG	JLHARG	CGARCH	GARCH-LHARG
λ	2.005	2.005	2.005	2.005	2.9392	2.2
	(1.489)	(1.489)	(1.489)	(1.489)	(1.5614)	(1.504)
θ	1.116e-005	1.068e-005	1.117e-005	9.357e-06	-	6.9802e-06
	(9.864e-008)	(9.466e-008)	(9.484e-008)	(8.3e-08)		(6.1416e-08)
δ	1.395	1.243	1.78	1.880	-	1.7568
	(0.04646)	(0.0482)	(0.04319)	(2.8e-02)		(0.034487)
β_d	2.993e+004	2.429e+004	3.382e+004	3.939e+04	-	5e+004
	(1037)	(439.4)	(180.1)	(6.2e+02)		(9.6e+003)
β_w	2.796e+004	2.317e+004	2.542e+004	3.028e+04	-	4.55e+004
	(1247)	(1199)	(225)	(2.8e+02)		(2.2e+003)
β_m	1.132e+004	1.322e+004	1.338e+004	1.689e+04	-	2.18e+004
	(897)	(1690)	(142.7)	(1.3e+02)		(253)
α_d	1.389e+004	0.2376	0.3991	0.4338	-	0.4129
	(1235)	(0.00113)	(0.007164)	(7.3e-03)		(0.046074)
α_w	-	0.1194	0.3446	0.410	-	0.41801
		(0.002058)	(0.01162)	(2.1e-02)		(0.13826)
α_m	-	3.85e-006	0.4034	0.519	-	0.25521
		(3.649e-006)	(0.02082)	(7.5e-02)		(0.68535)
γ	-	223.7	134.8	125.4	-	126.54
		(5.122)	(9.525)	(6.8)		(13.702)
$\tilde{\theta}$	-	-	-	4.70e-05	-	-
				(3.0e-06)		
$\tilde{\delta}$	-	-	-	1.152	-	-
				(2.5e-02)		
$\tilde{\Theta}$	-	-	-	0.2994	-	-
				(8.9e-03)		
ω	-	-	-	-	1.2667e-006	0
					(1.8699e-007)	(0)
α_h	-	-	-	-	1.49e-006	2.8999e-06
					(6.5849e-007)	(6.6497e-07)
β_h	-	-	-	-	0.49505	0.647
					(0.061499)	(0.06215)
γ_h	-	-	-	-	425.59	237.46
					(169.0385)	(36.62)
α_q	-	-	-	-	2.4502e-006	-
					(2.8226e-007)	
β_q	-	-	-	-	0.9861	-
					(0.0020844)	
γ_q	-	-	-	-	87.824	-
					(15.0899)	
Log-lik.	-25344	-25279	-25234	-24476	12473	23175

Table 4.1: Maximum likelihood estimates, robust standard errors, and models' performance.

The variance process in the majority of models considered in the thesis (ZM-LHARG, JL-HARG, GARCH-LHARG, CGARCH) is not always well-defined. For example, while we can ensure that P-LHARG model satisfies condition (2.1.5), for the ZM-LHARG model the relation (3.1.24) cannot be prevented from obtaining negative values. Since the ZM-LHARG is worth considering, we provide some numerical evidence supporting the analytical MGF as a reliable approximation of the MGF computed by simulation. We compare an extensive Monte Carlo (MC) simulation of the ZM-LHARG dynamics where the non centrality parameter is artificially bounded from below (by zero) with the analytical MGF computed according to Proposition 12. As the probability of obtaining a negative value for the non centrality of the gamma distribution is small (given the parameter values in Table 4.1), we can assess that the analytical MGF is a good approximation of the unknown MGF of the regularised ZM-LHARG. We fix the number of MC to 0.5×10^6 and consider six relevant maturities, one day ($T = 1$), one week ($T = 5$), one month ($T = 22$), one quarter ($T = 63$), six months ($T = 126$), and one year ($T = 256$). In the left column from top to bottom of Figure 4.2 we plot the MGF, the real and imaginary parts of the characteristic function under the physical measure, respectively, while in the right column we show the same quantities under the risk-neutral measure. The lines correspond to the analytical MGFs while the MC expectations are represented by points whose size is larger than the associated error bars. The quality of the agreement is extremely high. Moreover, the MC estimate of the probability associated with the event $\Theta(\mathbf{RV}_{t-1}, \mathbf{L}_{t-1}) < 0$ is 2×10^{-5} under $\mathbb{P}$, and 3×10^{-6} under $\mathbb{Q}$, confirming once more the reliability of the approximation. Similar analysis can be proceeded for ZM-JLHARG, GARCH-LHARG or CGARCH processes.

4.3 Option pricing methodology and calibration of SDF

Our option data set contains European out-of-the-money (OTM) options on S&P 500 index for each Wednesday from January 1, 1996 to December 31, 2004. We first apply a standard filter removing options with maturity less than 10 days or more than 365 days, implied volatility larger than 70% and prices less than 0.05$ (see Barone-Adesi et al. (2008) and Corsi et al.

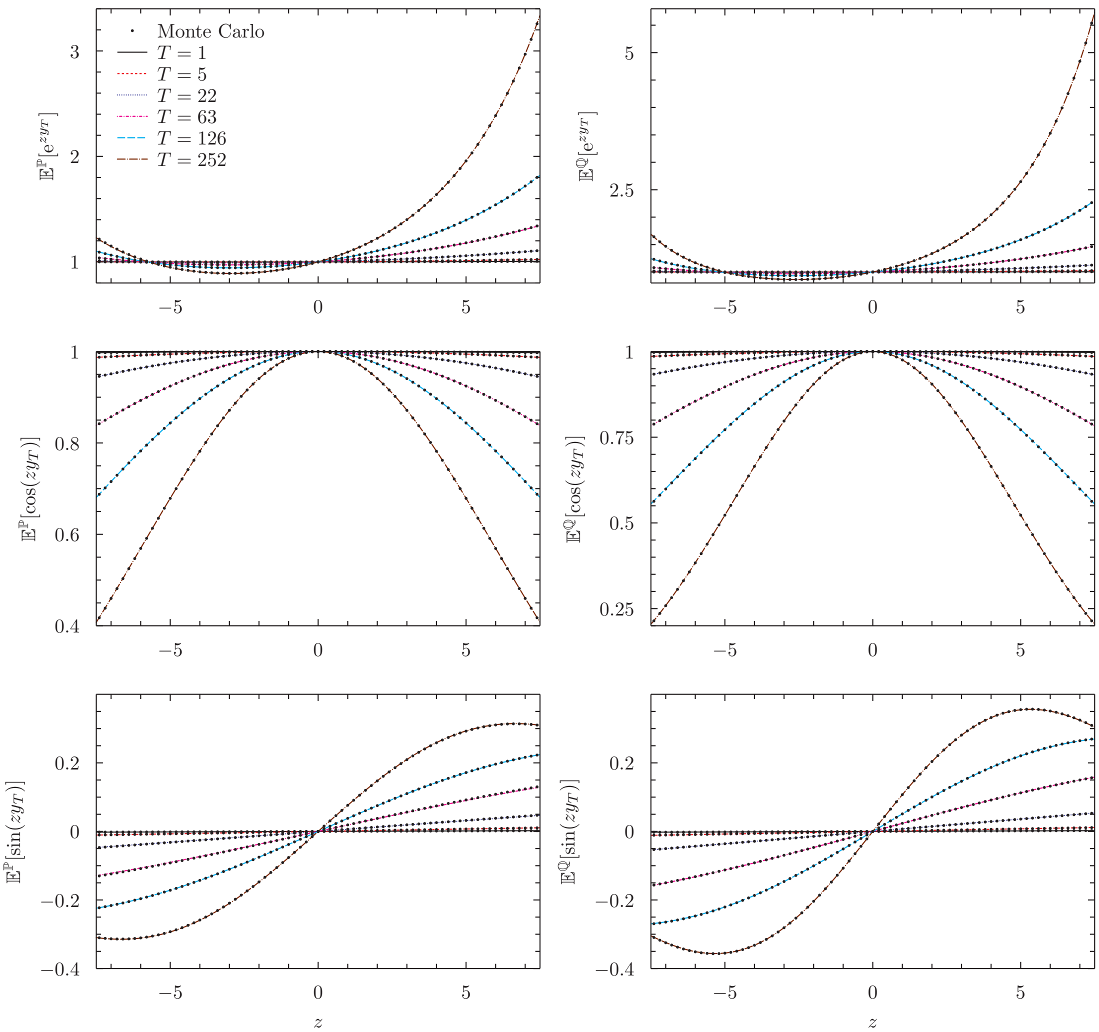

Figure 4.2: Left column, from top to bottom: MGF, real and imaginary parts of the characteristic function of the ZM-LHARG process under the physical measure $\mathbb{P}$. Right column, from top to bottom: MGF, real and imaginary parts of the Characteristic Function of the ZM-LHARG process under the risk-neutral measure $\mathbb{Q}$. The lines correspond to different maturities $T = 1, 5, 22, 63, 126, 252$, while points to Monte Carlo expected values; Monte Carlo error bars are smaller than the point size.

(2013)). Using K/S_t as definition of moneyness, we filter out deep OTM options with moneyness larger than 1.3 for call options and less than 0.7 for put options. This choice yields a total number of 46066 observations. For our purposes, put options are identified as Deep OTM (DOTM) if their moneyness is between $0.7 \leq m \leq 0.9$ and OTM if $0.9 < m \leq 0.98$. On the other hand, call options are said to be DOTM if $1.1 < m \leq 1.3$ and OTM if $1.02 < m \leq 1.1$. Options are called at-the-money (ATM) if $0.98 < m \leq 1.02$. As far as the time to maturity τ is concerned, we identify options as short maturity ($\tau \leq 50$ days), short-medium maturity ($50 < \tau \leq 90$ days), long-medium maturity ($90 < \tau \leq 160$ days) and long maturity ($\tau > 160$ days).

Proposed Esscher transform (2.2.2) has a vector of free parameters. In all considered models suiting our general framework ν_y is constrained by no-arbitrage condition, but vector $\boldsymbol{\nu}_\mathbf{f}$ has to be calibrated on option data (for example ν_r in case of LHARG-RV model and ν_h in case of CGARCH model). For the calibration procedure, we adopt a method based on an unconditional optimisation made by minimising the distance between the market implied and the model implied volatility surface. For this reason, we divide our dataset in different intervals of moneyness and maturity obtaining a 5x4 moneyness-maturity grid. For each subset of the grid we compute the unconditional mean of the market implied volatility of the options within the subset. In this way, we obtain a 20-points-grid representation of the implied volatility surface as shown in Table 4.2. Finally, the calibration of the variance risk-premia is obtained by computing the same grid for the model implied volatility and finding the optimal values for the $\boldsymbol{\nu}_\mathbf{f}$ which minimise the distance between the two grids, i.e. The objective function $f_{\text{obj}}(\boldsymbol{\nu}_\mathbf{f})$ is defined as

$$f_{\text{obj}}(\boldsymbol{\nu}_\mathbf{f}) = \sqrt{\sum_{i=1}^{5} \sum_{j=1}^{4} \left(\text{IV}_{ij}^{\text{mod}}\left(\boldsymbol{\nu}_\mathbf{f}\right) - \text{IV}_{ij}^{\text{mkt}}\right)^2},$$

and represents the distance between the two matrices relative to the model implied volatility surface and the market one, whose elements are $\text{IV}_{ij}^{\text{mod}}\left(\boldsymbol{\nu}_\mathbf{f}\right)$ and $\text{IV}_{ij}^{\text{mkt}}$, respectively. In Table

4.2 we report variance-risk parameters of SDF calibrated on option data.

Model	Parameter ν_r	ν_h	ν_j
HARGL	-3119	-	-
P-LHARG	-3069	-	-
ZM-LHARG	-3375	-	-
JLHARG	-4442	-	-1033
CGARCH	-	-46437	-
GARCH-LHARG	1195	-7798	-

Table 4.2: Variance-risk parameters of SDF calibrated on option data.

There can be three classes of numerical methods for option pricing distinguished: Monte Carlo simulations, numerical scheme for solving the pricing PDE and Fourier transform techniques. Since characteristic function is available for price processes in models considered in the thesis, we compute price of European options by means of Fourier inversion methods which turn out to be a very effective tool. In order to implement the option pricing scheme numerically, we use the COS method which is based on Fourier-cosine expansions and it was introduced by Fang and Oosterlee (2008).

We can summarise the option pricing procedure in four steps: (i) estimation of the parameters under the physical measure $\mathbb{P}$; (ii) unconditional calibration of parameter vector $\boldsymbol{\nu}_f$; (iii) mapping of parameters of the model estimated under $\mathbb{P}$ into parameters under $\mathbb{Q}$; (iv) numerical computation of option prices through COS method using the MGF recursive formulas with parameters under $\mathbb{Q}$.

4.4 Option pricing performance

In this section we present empirical results for option pricing with models presented in the thesis. We compare CGARCH model with the SDF proposed in our general framework with the original change of measure proposed by Christoffersen et al. (2008) which is based on making equal the second moment of the process under both measures. As a result the dynamics of the process is described under $\mathbb{Q}$ by (2.1.7) with parameters λ^*, ω, α_1, β_1^*, γ_1^*, α_2, β_2^*, γ_2^*, where $\lambda^* = 1/2$ and

$$\begin{aligned}\beta_1^* &= \beta_1 + \alpha_1 \left(\gamma_1^{*2} - \gamma_1^2\right) + \alpha_2 \left(\gamma_2^{*2} - \gamma_2^2\right),\\ \beta_2^* &= \beta_2 + \alpha_1 \left(\gamma_1^{*2} - \gamma_1^2\right) + \alpha_2 \left(\gamma_2^{*2} - \gamma_2^2\right),\\ \gamma_i^* &= \gamma_i + \lambda + \frac{1}{2} \quad \text{for} \quad i = 1, 2.\end{aligned} \tag{4.4.1}$$

Note that by construction this change of measure does not require calibration of any parameter under martingale measure and it gives the risk-neutral variance equal to physical one with the ratio ξ between volatilities under measure $\mathbb{Q}$ and $\mathbb{P}$ equal to 1.[2] Whereas applying change of measure by SDF (3.2.7) give us

$$\xi = \frac{1}{\sqrt{1 + 2\nu_1 \left(\alpha_1 + \alpha_2\right)}}. \tag{4.4.2}$$

We also compare RV models with the HARGL-RV presented in Corsi et al. (2013). Since the functional form of the leverage of the latter model is not consistent with the current general framework, closed-form formulae for the MGF and for option pricing are not available. Thus, we resort to numerical methodologies such as extensive Monte Carlo scenario generation.

As a measure of the option pricing performance we use the percentage Implied Volatility Root

[2] Notice that risk-neutral dynamics proposed by Christoffersen et al. (2008) is not equivalent to dynamics obtained by applying SDF (3.2.7) with $\nu_1 = 0$. Even if in both cases risk-neutral variance is equal to physical variance, particular components of variance have different dynamics.

Mean Square Error ($RMSE_{IV}$) put forward by Renault (1997) and computed as

$$RMSE_{IV} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(IV_i^{mkt} - IV_i^{mod}\right)^2} \times 100\,,$$

where N is the number of options, IV^{mkt} and IV^{mod} represent the market and model implied volatilities, respectively. An alternative performance measure corresponds to the Price Root Mean Square Error ($RMSE_P$) defined in a similar way as $RMSE_{IV}$ but with implied volatilities replaced by relative prices. We employ the $RMSE_{IV}$ measure since it tends to put more weight on OTM options, while the $RMSE_P$ emphasises the importance of ATM options.

Implied Volatility RMSE

	Moneyness	
Model	$0.9 < m < 1.1$	$0.8 < m < 1.2$
HARGL	4.067	6.428
P-LHARG	3.664	5.026
ZM-LHARG	3.539	4.732
GARCH-LHARG	3.385	4.349
CGARCH 1D	5.543	6.877
CGARCH 2D	4.426	5.384

Table 4.3: Global option pricing performance on S&P500 out-of-the-money options from January 1, 1996 to December 31, 2004, computed with the CRV measure estimated from 1990 to 2007. We use the maximum likelihood parameter estimates from Table 4.1. GARCH 1D stands for model with original change of measure (4.4.1) and GARCH 2D stands for model with change of measure by applying SDF (3.2.7) .

The result of our empirical analysis is that both LHARG models outperform competing RV-based stochastic volatility model (HARGL). Table 4.3 shows that P-LHARG and ZM-LHARG outperforms HARGL by about 4% and 7%, respectively, in range of moneyness $0.9 < m < 1.1$ and by about 17% and 22%, respectively, in range of moneyness $0.8 < m < 1.2$. Moreover, ZM-LHARG improves P-LHARG by about 3% and 6%, in range of moneyness $0.9 < m < 1.1$

and $0.8 < m < 1.2$, respectively. Adding latent volatility factor in GARCH-LHARG-RV gives improvement of about 4% in narrow and 8% in wide region of moneyness. Table 4.3 underlines also the importance of variance risk premium in option pricing via the example of CGARCH model. The change of measure in CGARCH proposed by Christoffersen et al. (2008) does not take into account variance risk premium and it underperforms the model with change of measure done by applying SDF (3.2.7). Introducing a variance risk premium in SDF gives us immediately an improvement of 22.1% for the wider range of moneyness. Both CGARCH model exhibit worse performance than RV models. This can be explained by the importance of information contained in realized variance for option pricing. The detailed analysis in Tables 4.5 and 4.6 shows that the discussed improvements are basically independent of the region of moneyness or maturity.

Comparing JLHARG (and also GARCH-LHARG) performance with ZM-LHARG, when RV in the latter model is rescaled to adjust the mean level of volatility is unfair. By rescaling RV we are pretending to take into account jumps in volatility in ZM-LHARG even if the model is not designed to do this. Therefore in Table 4.4 we compare JLHARG and GARCH-LHARG with ZM-LHARG where RV is rescaled just by over-night effect for JLHARG and ZM-LHARG models (in the case of GARCH-LHARG any rescaling in not needed as GARCH component is taking into account both jumps and over-night effect in volatility). We can observe that adding jumps in volatility in JLHARG-RV gives improvement of about 4% and 14%, in range of moneyness $0.9 < m < 1.1$ and $0.8 < m < 1.2$, respectively. Moreover GARCH-LHARG outperforms JLHARG by 40% and 29%, in range of moneyness $0.9 < m < 1.1$ and $0.8 < m < 1.2$, respectively. This huge outperformance of GARCH-LHARG over JLHARG can be explained in two ways. First the impact of jumps on volatility is not big compared with over-night effect and other latent factors of volatility. Second JLHARG lacks memory in modelling jumps in volatility which implies too little persistence.

Implied Volatility RMSE

Model	Moneyness $0.9 < m < 1.1$	$0.8 < m < 1.2$
ZM-LHARG	5.690	7.186
JLHARG	5.454	6.159
GARCH-LHARG	3.385	4.349

Table 4.4: Comparison of option pricing performance for ZM-LHARG, JLHARG and GARCH-LHARG on S&P500 out-of-the-money options from January 1, 1996 to December 31, 2004.

The difference in option pricing performance can be explained by looking at term structure of skewness and kurtosis. On Figure 4.3 we compare the skewness and excess kurtosis for models using RV time series. Good option pricing performance of GARCH-LHARG model corresponds with the largest negative skewness and the highest excess kurtosis. On Figure 4.4 we compare the skewness and excess kurtosis for CGARCH under measure $\mathbb{P}$ and measure $\mathbb{Q}$ resulting from a one-dimensional and two-dimensional pricing kernel - green and red lines, respectively. We can clearly see that the process under the risk-neutral measure corresponding to a SDF which includes compensation for variance risk exhibits the largest negative skewness and the highest excess kurtosis.

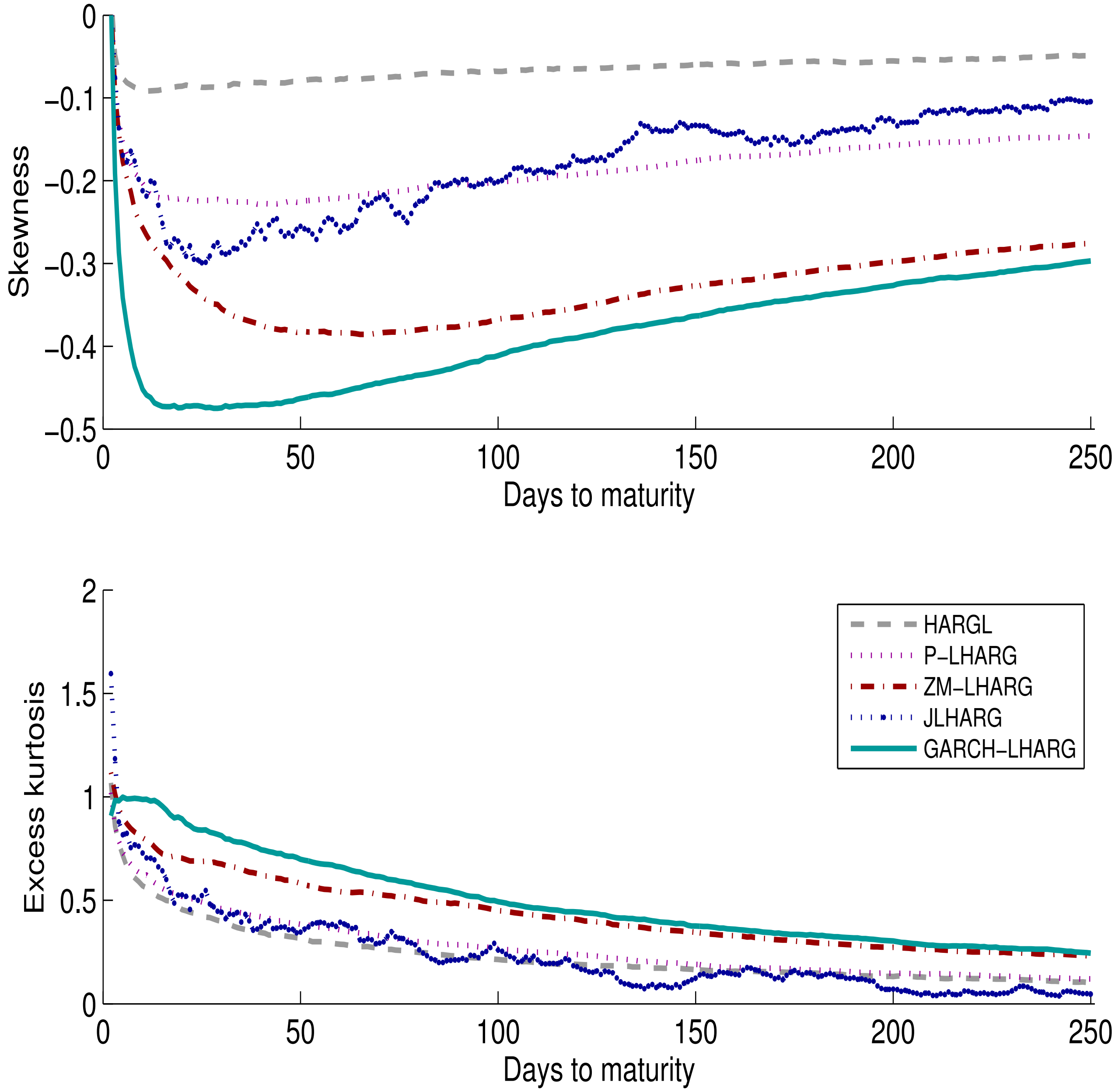

Figure 4.3: Skewness and excess kurtosis under physical measure for processes: HARGL, P-LHARG, ZM-LHARG, JLHARG and GARCH-LHARG. Since process HARGL does not have MGF, we compute skewness and kurtosis by means of Monte Carlo simulation.

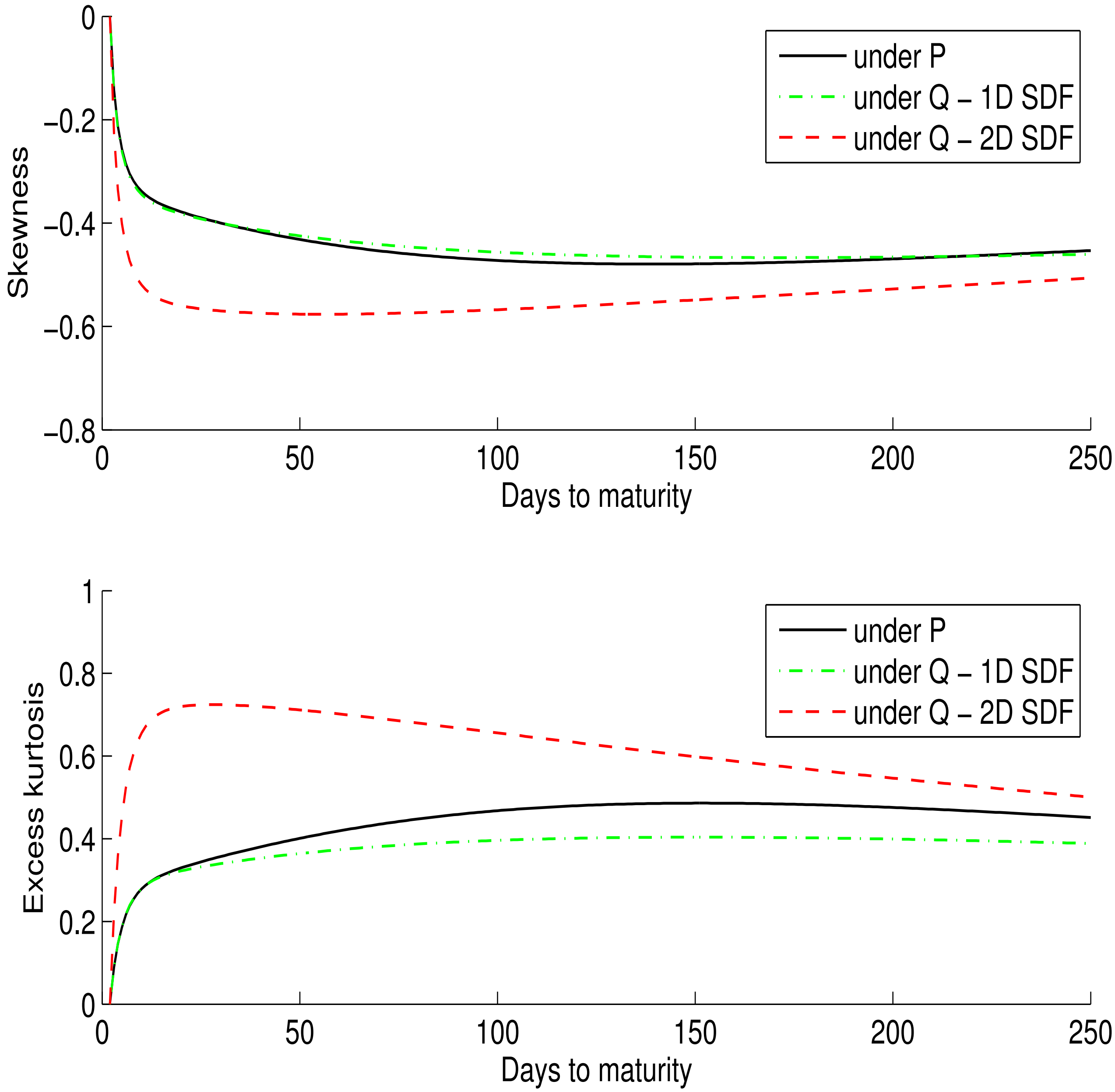

Figure 4.4: Skewness and excess kurtosis of CGARCH process under physical and risk-neutral measure, computed as in Christoffersen et al. (2008) (1D SDF) and by means of the two-dimensional SDF (3.2.7). Skewness and kurtosis are computed using third and forth order derivatives of MGF.

	Maturity			
Moneyness	$\tau \leq 50$	$50 < \tau \leq 90$	$90 < \tau \leq 160$	$160 < \tau$
Panel A	HARGL Implied Volatility RMSE			
$0.8 \leq m \leq 0.9$	16.806	8.802	7.325	5.951
$0.9 < m \leq 0.98$	5.778	4.533	4.143	4.118
$0.98 < m \leq 1.02$	2.771	2.922	3.134	3.591
$1.02 < m \leq 1.1$	3.436	3.387	3.403	3.591
$1.1 < m \leq 1.2$	4.696	3.982	3.902	3.838
Panel B	P-LHARG Implied Volatility RMSE			
$0.8 \leq m \leq 0.9$	11.160	7.579	6.390	5.368
$0.9 < m \leq 0.98$	4.991	4.080	3.841	3.951
$0.98 < m \leq 1.02$	2.650	2.849	3.077	3.546
$1.02 < m \leq 1.1$	2.984	2.975	3.123	3.421
$1.1 < m \leq 1.2$	4.322	3.095	3.137	3.359
Panel C	ZM-LHARG Implied Volatility RMSE			
$0.8 \leq m \leq 0.9$	10.458	6.807	5.741	4.975
$0.9 < m \leq 0.98$	4.708	3.837	3.658	3.877
$0.98 < m \leq 1.02$	2.650	2.878	3.083	3.574
$1.02 < m \leq 1.1$	2.941	2.899	3.036	3.381
$1.1 < m \leq 1.2$	4.244	2.769	2.748	3.115
Panel D	JLHARG Implied Volatility RMSE			
$0.8 \leq m \leq 0.9$	11.228	7.356	5.732	4.720
$0.9 < m \leq 0.98$	6.345	5.169	4.454	4.455
$0.98 < m \leq 1.02$	5.782	5.073	2.872	3.261
$1.02 < m \leq 1.1$	5.820	5.328	4.799	4.875
$1.1 < m \leq 1.2$	6.791	4.968	4.663	4.932

Table 4.5: Detailed option pricing performance on S&P500 out-of-the-money options from January 1, 1996 to December 31, 2004, computed with the CRV measure estimated from 1990 to 2007. We use the maximum likelihood parameter estimates from Table 4.1.

	Maturity			
Moneyness	$\tau \leq 50$	$50 < \tau \leq 90$	$90 < \tau \leq 160$	$160 < \tau$
Panel E	GARCH-LHARG-RV Implied Volatility RMSE			
$0.8 \leq m \leq 0.9$	9.257	5.776	4.794	4.178
$0.9 < m \leq 0.98$	4.029	3.313	3.322	3.667
$0.98 < m \leq 1.02$	2.680	2.917	3.210	3.801
$1.02 < m \leq 1.1$	3.030	3.170	3.465	3.889
$1.1 < m \leq 1.2$	4.732	2.945	3.197	3.765
Panel F	CGARCH 1D Implied Volatility RMSE			
$0.8 \leq m \leq 0.9$	13.570	9.921	8.964	8.155
$0.9 < m \leq 0.98$	7.536	6.744	6.652	6.911
$0.98 < m \leq 1.02$	4.495	4.975	5.382	6.202
$1.02 < m \leq 1.1$	3.521	2.828	2.719	3.243
$1.1 < m \leq 1.2$	7.098	2.942	2.331	2.405
Panel G	CGARCH 2D Implied Volatility RMSE			
$0.8 \leq m \leq 0.9$	10.742	6.502	4.849	3.907
$0.9 < m \leq 0.98$	5.763	4.079	3.500	3.828
$0.98 < m \leq 1.02$	3.658	3.409	3.743	4.963
$1.02 < m \leq 1.1$	3.164	3.430	5.124	7.507
$1.1 < m \leq 1.2$	6.357	2.644	4.533	7.438

Table 4.6: Continuation of Table 4.5.

Chapter 5

Variance Risk Premium

5.1 k-CGARCH implied variance risk premium

The equity risk premium represents the additional profit that investors demand from investing in an asset with uncertain future price level. It is formally defined as

$$\mathrm{ERP}(t,T) = \frac{1}{T}\left(\mathbb{E}_t^{\mathbb{P}}\left[\frac{S_{t+T}-S_t}{S_t}\right] - \mathbb{E}_t^{\mathbb{Q}}\left[\frac{S_{t+T}-S_t}{S_t}\right]\right). \tag{5.1.1}$$

Since the measure $\mathbb{Q}$ is chosen so that the process $\exp(-rt)\,S_t$ is a $\mathbb{Q}$-martingale one can easily see that the second expectation in (5.1.1) is equal to $\exp(rT)-1$, regardless of which model is used to describe the dynamics of the asset's price S_t. Therefore, the equity risk premium is a sole property of the measure $\mathbb{P}$. Since under our modelling assumptions the properties of the physical measure are fixed conditionally on $\mathcal{F}_t$ we conclude that investors require a stable compensation for the uncertainty about future price levels.

On the other hand variance risk premium essentially depends on both measures $\mathbb{P}$ and $\mathbb{Q}$. Following Bollerslev and Todorov (2011) we define

$$\mathrm{VRP}(t,T) = \frac{1}{T}\left(\mathbb{E}_t^{\mathbb{P}}\left[\mathrm{QV}_{t,t+T}\right] - \mathbb{E}_t^{\mathbb{Q}}\left[\mathrm{QV}_{t,t+T}\right]\right), \tag{5.1.2}$$

where $\mathrm{QV}_{t,t+T}$ stands for the quadratic variation of the asset price over time $[t, t+T]$. Underestimating variance risk premium was one of the reasons for fail of Long-Term Capital Management in 1997 (see Lowenstein (2000)). During 1998 implied volatility was relatively high, around 19% with realized volatility at the level of 16%. Big spread between the two volatilities can be translated into high fee for issuing options. However, higher initial profits from going short with volatility are usually associated by a higher risk of temporal changes of returns variance. LTCM was neglecting that risk and was writing options to the extent in which it has been responsible for a fourth of the overall market and it gained a nickname *the Central Bank of Volatility.* This strategy had exposed the hedge fund to huge losses when the realized variance raised.

After the LTCM collapse in 1998, variance swap contracts with payoff being the difference between realized variance and predefined strike, started to be traded OTC extensively. Buyers of such contracts are usually mutual funds and portfolio managers who need insurance against rising volatility. On the other hand variance swap contracts can be used for speculative reasons, as they provide investors with pure volatility exposure. Consequently there has been an increasing interest in pricing those contracts and in understanding the nature of the difference between volatilities under risk-neutral measure and physical one.

Variance risk premium, the spread between conditional expectation of variance under physical and risk-neutral measure, is not only interesting for the sake of understanding the features of financial contracts like variance swap rate and volatility derivatives in general, but also from the economic point of view. Variance risk premium quantifies the reward that investor demand for bearing risk related with unknown future level of variance. It has been well-documented that variance risk premium is significantly non-zero and it is time-varying with nontrivial dynamics (Carr and Wu (2009); Bakshi et al. (2010)). Since lack of information about distribution of future realized volatility under physical measure, a complete model-free dynamic measure of variance risk premium is unfeasible.

In the case of GARCH processes the proxy of the daily quadratic variation is $\mathcal{L}(\mathbf{f}_t)$ under the physical measure and $\mathcal{L}(\mathbf{f^*}_{t+k})$ under the risk-neutral measure. Consistently, the variance risk premium in the GARCH modelling framework is equal to

$$\mathrm{VRP}(t,T) = \frac{1}{T}\left(\mathbb{E}_t^{\mathbb{P}}\left[\sum_{k=1}^{T}\mathcal{L}(\mathbf{f}_{t+k})\right] - \mathbb{E}_t^{\mathbb{Q}}\left[\sum_{k=1}^{T}\mathcal{L}(\mathbf{f^*}_{t+k})\right]\right). \tag{5.1.3}$$

Straightforward computations show that

$$\mathbb{E}_t^{\mathbb{P}}[\mathbf{f}_{t+T}] = \tilde{\mathbf{d}} + \tilde{\mathbf{M}}\tilde{\mathbf{d}} + \tilde{\mathbf{M}}^2\tilde{\mathbf{d}} + \ldots + \tilde{\mathbf{M}}^{T-2}\tilde{\mathbf{d}} + \tilde{\mathbf{M}}^{T-1}\mathbf{f}_{t+1}, \tag{5.1.4}$$

where

$$\tilde{\mathbf{d}} = \mathbf{d} + \begin{bmatrix} n_1^{(d)} \\ \vdots \\ n_k^{(d)} \end{bmatrix} \quad \text{and} \quad \tilde{\mathbf{M}} = \mathbf{M} + \begin{bmatrix} n_1^{(m)} & \ldots & n_1^{(m)} \\ \vdots & & \vdots \\ n_k^{(m)} & \ldots & n_k^{(m)} \end{bmatrix}, \tag{5.1.5}$$

with $n_i^{(d)} = n_{i,1} + \ldots + n_{i,k}$ and $n_i^{(m)} = n_{i,1}\gamma_1^2 + \ldots + n_{i,k}\gamma_k^2$. An analogous formula can be derived for $\mathbb{E}_t^{\mathbb{Q}}\left[\mathbf{f}^*_{t+T}\right]$ with $\tilde{\mathbf{d}}^*$ and $\tilde{\mathbf{M}}^*$. From this we immediately obtain

Proposition 30. *The variance risk premium for the k-CGARCH(1,1) model with pricing kernel (3.2.7) is given by*

$$\begin{aligned} &VRP(t,T) = \\ &= \frac{1}{T}\mathcal{L}\left((T-1)\tilde{\mathbf{d}} + (T-2)\tilde{\mathbf{M}}\tilde{\mathbf{d}} + \ldots + \tilde{\mathbf{M}}^{T-2}\tilde{\mathbf{d}} + \left(\mathbf{I} + \tilde{\mathbf{M}} + \ldots + \tilde{\mathbf{M}}^{T-1}\right)\mathbf{f}_{t+1}\right) \\ &- \frac{1}{T}\mathcal{L}\left((T-1)\tilde{\mathbf{d}}^* + (T-2)\tilde{\mathbf{M}}^*\tilde{\mathbf{d}}^* + \ldots + \tilde{\mathbf{M}}^{*T-2}\tilde{\mathbf{d}}^* + \left(\mathbf{I} + \tilde{\mathbf{M}}^* + \ldots + \tilde{\mathbf{M}}^{*T-1}\right)\mathbf{f^*}_{t+1}\right). \end{aligned} \tag{5.1.6}$$

For technical reasons in this thesis we consider only CGARCH-implied variance risk premium (VRP computations are simplified in this case). Moreover, in the next section we show that

CGARCH implies a realistic family of term structures if variance swap rates.

5.2 Hump-shaped term structure of variance swap rate

The second component of the right-hand-side of equation (5.1.2) is termed Variance Swap Rate (VSR). The aim of this section is to show that the family of CGARCH-implied VSR term structures is richer and more realistic than in the case of single component models. The VSR term structure at day t is a function of T defined as follows

$$VSR_t(T) = \frac{1}{T}\mathbb{E}_t^{\mathbb{Q}}\left[\sum_{k=1}^{T}\mathcal{L}\left(\mathbf{f}^*_{t+k}\right)\right]. \tag{5.2.1}$$

Stated differently, the variance swap rate is an average of expected variances under measure $\mathbb{Q}$ over next T days. For one-component volatility models belonging to our framework, we show that the VSR term structure is always a monotonic function. From (5.1.4) we see that

$$\mathbb{E}_t^{\mathbb{Q}}\left[\mathcal{L}\left(\mathbf{f}^*_{t+T+1}\right)\right] - \mathbb{E}_t^{\mathbb{Q}}\left[\mathcal{L}\left(\mathbf{f}^*_{t+T}\right)\right] = \mathcal{L}\left(\tilde{\mathbf{M}}^{*\,T-1}\left(\tilde{\mathbf{d}}^* + \left(\tilde{\mathbf{M}}^* - \mathbf{I}\right)\mathbf{f}^*_{t+1}\right)\right). \tag{5.2.2}$$

For a single factor model $\mathcal{L}\left(\mathbf{f}^*_t\right) = f^*_t$ and all the quantities in (5.2.2) are scalars, and we conclude that $\{\mathbb{E}_t^{\mathbb{Q}}\left[\mathcal{L}\left(\mathbf{f}^*_{t+T}\right)\right]\}_{T\geq 1}$ (and consequently $\{VSR_t(T)\}_{T\geq 1}$) is an increasing sequence if

$$f^*_{t+1} < \frac{\tilde{d}^*}{1-\tilde{M}^*}, \tag{5.2.3}$$

decreasing if the opposite inequality holds, and constant if the relation (5.2.3) becomes an equality. Recognising that $\tilde{d}^*/(1-\tilde{M}^*)$ is the variance unconditional mean, we have that at every day t the VSR term structure is either a constant function of T – if the risk-neutral variance f^*_{t+1} is equal to the unconditional level – or a strictly monotonic function of T otherwise. This is a serious limitation of single factor volatility models. In reality the variety of shapes assumed by the term structure is richer and it includes so called hump-shaped term structures. We now show that two-factor volatility models actually feature humps in the VSR term structure.

Using (5.1.4) one can rewrite (5.2.1) as

$$VSR_t(T) = \frac{1}{T}\mathcal{L}\left((T-1)\tilde{\mathbf{d}}^* + (T-2)\tilde{\mathbf{M}}^*\tilde{\mathbf{d}}^* + \ldots + \tilde{\mathbf{M}}^{*T-2}\tilde{\mathbf{d}}^* + \left(\mathbf{I} + \tilde{\mathbf{M}}^* + \ldots + \tilde{\mathbf{M}}^{*T-1}\right)\mathbf{f}^*_{t+1}\right), \tag{5.2.4}$$

where $\tilde{\mathbf{d}}^*$ and $\tilde{\mathbf{M}}^*$ are defined in (5.1.5). Employing formula (5.2.4) we draw in the top panel of Figure 5.1 the term structure for the CGARCH model. We clearly recognise a hump-shaped term-structure – firstly it increases and then becomes a decreasing function of the maturity.

The decomposition in the middle panel of Figure 5.1 explains how the hump in term structure can be obtained (we draw the two components of variance swap rate which can be obtained from formula (5.2.4)). Let us consider a day when the long-term component is relatively high and decreasing, while the short-term component is negative but increasing. Moreover, for small maturities the tangent of the short-term component is larger in absolute terms than the tangent of the long-term component, and this relation is reversed for longer maturities. Then, as a consequence of the superposition of two components, the first derivative of the complete curve switches sign and features a hump-shaped VSR term structure.

The analysis of term structures induced by GARCH models supports our choice of two-component GARCH models when considering the variance risk premium. Since the variance swap rate is a component of the variance risk premium, it is necessary to employ multi-component volatility models to correctly capture the information content carried by the VSR over different maturities. On the bottom panel of Figure 5.1 we can see that the hump from the VSR term structure transfers to the VRP term structure. In section 5.4 we test stock return predictability exploiting the information content of the term structure of variance risk premium.

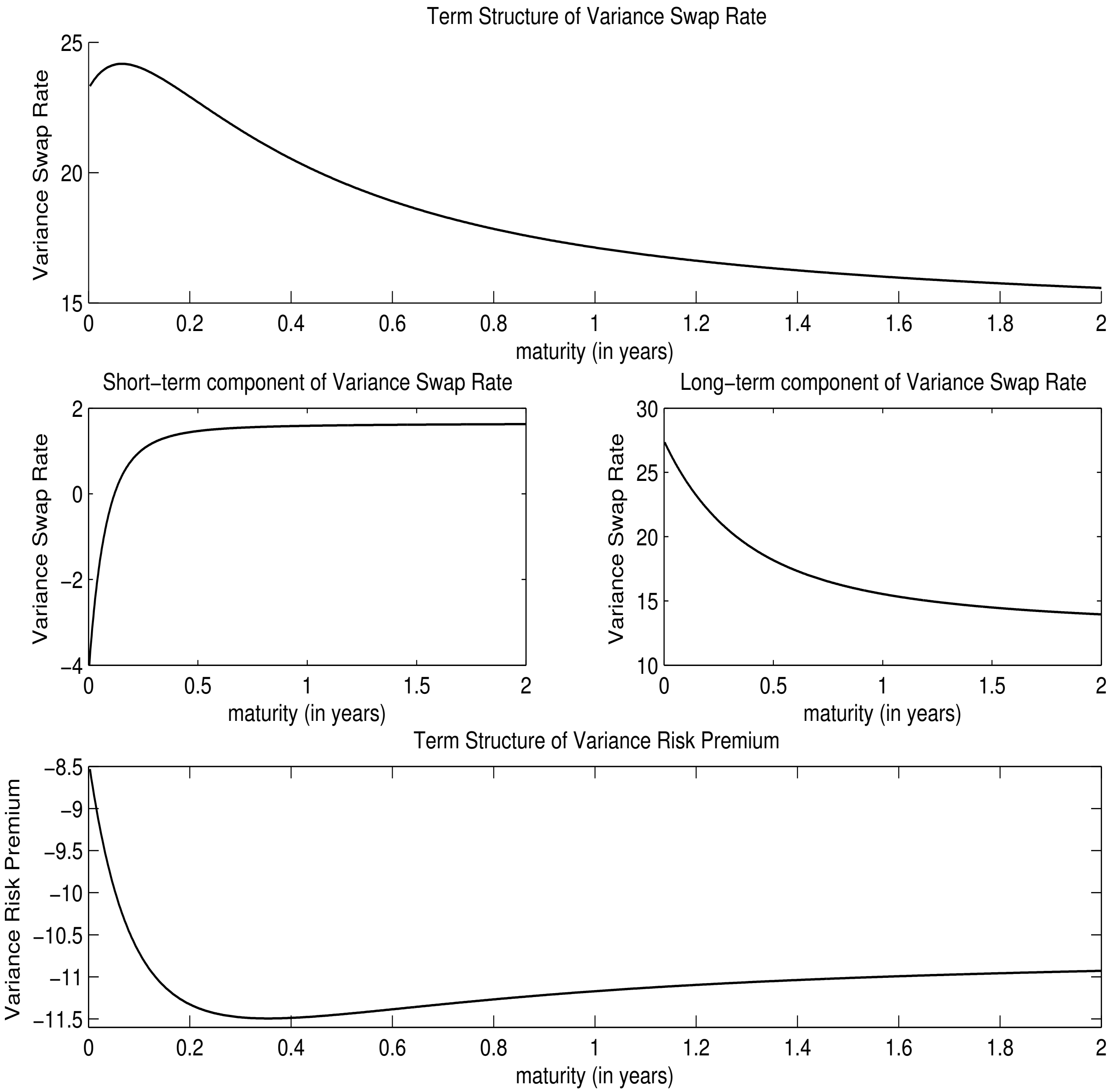

Figure 5.1: Term structure of annualised variance swap rate and variance risk premium obtained from CGARCH with initial annual variance set equal to 10.21%, short-run component of variance equal to −1.76%, and long-run component equal to 11.97%. Parameters of the model are given in Table 4.1 with $\nu_f = -46437$.

5.3 Dynamic measurement of CGARCH-implied variance risk premium

Recently, Wang and Eraker (2015) have proposed a nonlinear diffusion model based measurement of variance risk premium fitting it directly to VIX and VIX derivatives data. Moreover Ait-Sahalia et al. (2015) calibrate the term structure of VRP implied by jump-diffusion model on variance swap market data and they investigate investors' willingness to ensure against volatility risk and its relation with various economic indicators. Since variance swap contracts and other volatility derivatives are mainly traded by hedge funds and the objectives of investment in the variance swap rate market are different from the stock market, one cannot in general expect that the model-implied curve conditioned on information from option and stock market data will coincide with the variance swap rate curve observed in the market. Therefore in our study we propose a dynamic measure of the variance risk premium implied by the nonlinear two component GARCH model fitted to stock and option dataset. Due to market segmentation our term structure of VRP will contain different information than term structure seen in variance swap market and we claim that our methodology of extracting term structure of VRP provides a significant predictor of stock market excess returns.

Due to the market incompleteness there are infinite SDFs consistent with no-arbitrage condition. In CGARCH model each SDF is identified by a free parameter ν_h which represents an aggregate attitude of investors towards the uncertainty about future level of volatility. If we assume that agents are maximising their power utility function given by (2.2.8) then ν_h is equal to $m\nu_y$ where m represents how much investors are penalising the input in utility for high volatility (we remind that ν_y is fixed by no-arbitrage condition). Since investors active in the market might change from one week to another, we relax the assumption that m and consequently ν_h is constant over the entire period. Instead we assume that the variance premium ν_h follows a stochastic dynamics. Specifically, we assume that $\nu_h^{(t)}$ is a $\mathbb{P}$-martingale. Since we do not price the risk associated with the randomness of ν_h, we replace its future values with the best predictor, $\hat{\nu}_h^{(t)} = \mathbb{E}\left[\nu_h^{(s)}|\mathcal{F}_t\right]$ for all $s \geq t$. Then, on each Wednesday t market applies a different pricing kernel according to the formula

$$M_{s,s+1}^{(t)} = \frac{\mathrm{e}^{-\hat{\nu}_h^{(t)}\mathcal{L}(\mathbf{f}_{s+2})-\nu_y y_{s+1}}}{\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{-\hat{\nu}_h^{(t)}\mathcal{L}(\mathbf{f}_{s+2})-\nu_y y_{s+1}}|\mathcal{F}_s\right]} \quad \text{for } s \geq t. \tag{5.3.1}$$

In order to extract the temporary information on variance risk premium from the option data we apply a dynamic calibration procedure. We calibrate parameter $\hat{\nu}_h^{(t)}$ of SDF (5.3.1) each week to obtain different mapping of the parameters of the model estimated under $\mathbb{P}$ into the parameters under $\mathbb{Q}_t$. The whole VRP dynamic measurement procedure can be described in four steps. The first one is done once per all Wednesdays:

1. estimation under the physical measure $\mathbb{P}$,

The steps from second to the forth are repeated every Wednesday t on data from option market on that Wednesday:

2. We measure the value of SDF parameter $\hat{\nu}_h^{(t)}$ doing calibration conditioned on Wednesday t on out-of-money put and call European options traded on Wednesday t,

3. Mapping of the parameters of the model estimated under $\mathbb{P}$ into the parameters under $\mathbb{Q}_t$ using the current SDF parameter $\hat{\nu}_h^{(t)}$ on that week and transformation (3.1.23).

4. Having extracted information from implied volatility surface we compute GARCH-implied variance risk premium measure:

$$\mathrm{VRP}^G(t,T) = \frac{1}{T}\left(\mathbb{E}_t^{\mathbb{P}}\left[\sum_{k=1}^{T}\mathcal{S}\left(\mathbf{h}_{t+k}\right)\right] - \mathbb{E}_t^{\mathbb{Q}_t}\left[\sum_{k=1}^{T}\mathcal{S}\left(\mathbf{h}_{t+k}^*\right)\right]\right), \tag{5.3.2}$$

   where $\mathbb{Q}_t$ is a time varying risk-neutral measure obtained in point 3 of the algorithm.

We would like here to emphasis that we do not consider described above algorithm as a dynamic model for option pricing (because as such it would be inconsistent in the sense of Richter and

Teichmann (2014)) but as a posteriori procedure to extract from the market data the information on model-implied variance risk premium.

In this section we analyse the dynamics of the variance risk premium in the special case when T is equal to 22 days. The choice of T is motivated by the availability of a market benchmark for the variance risk premium. In fact, the second component in the right hand side of equation (5.3.2) – the conditional expected value of annualised variance over the next month under the martingale measure – is the variance swap rate $\mathrm{VSR}_t(22)$ (see formula (5.2.4)). Then, following Carr and Wu (2006) and Bollerslev et al. (2009) we approximate it by the square of the VIX Index. Under the assumption that the stock price dynamics does not admit jumps, both quantities are equal [1]. On the top panel of Figure 5.2 we present VIX^2 and the value of $\mathrm{VSR}_t(22)$ computed using the CGARCH model and the pricing kernel (3.2.7) [2].

To measure the market view on realized variance over next 22 days under physical measure we follow the approach by Bollerslev and Todorov (2011) based on reduced-form time series models. For tractability reasons we apply the HAR-RV modelling approach of Corsi (2009) and we assume the model-implied conditional expectations as a good proxy for market-implied expectations. On the bottom panel of Figure 5.2 we compare the expected variance over next 22 days under objective measure obtained with the CGARCH and HAR-RV models.

To sum up the proxy for market variance risk premium over next month can be read as:

$$\mathrm{VRP}^M(t, 22) = \frac{1}{22} \sum_{k=1}^{22} \mathbb{E}_t^{\mathbb{P}} \left[RV_{t+k} \right] - \mathrm{VIX}_t^2, \tag{5.3.3}$$

where estimate of $\mathbb{E}_t^{\mathbb{P}} \left[RV_{t+k} \right]$ is based on the application of the HAR model to the Realized

[1] See Filipović (2013) and Ait-Sahalia et al. (2015).

[2] Time series of VIX Index has been downloaded from http://finance.yahoo.com.

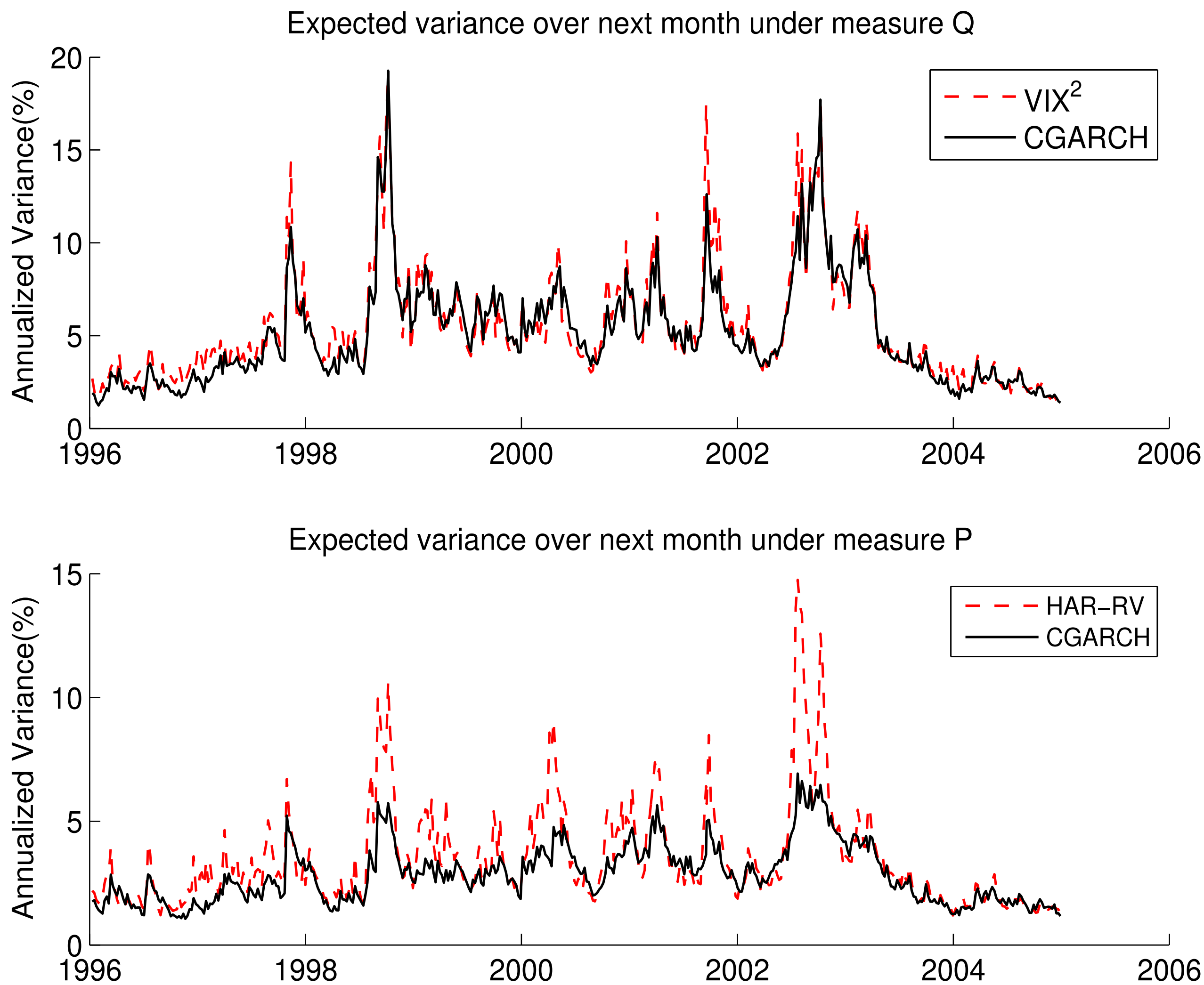

Figure 5.2: Expected variance over next month from 1996 to 2005. Top panel: comparison between squared VIX Index and expected variance over next month under risk-neutral measure resulting from CGARCH. Bottom panel: expected variance over next month under physical measure resulting from HAR-RV and CGARCH models.

Volatility process RV_t [3]. On Figure 5.3, we present the variance risk premium computed for different models and different assumptions about the parameter ν_f. From the first panel we clearly see the reason why ν_f is termed variance premium: When ν_f is equal to zero then also the variance risk premium is approximately 0, whereas for ν_f statically calibrated to option data we obtain a substantial negative premium (around -2 percent). This latter case also demonstrates that the variance risk premium is not constant over time but undergoes sizable fluctuations which can be ascribed to variations of the conditioning volatility factors. The same panel shows that this effect is exacerbated moving to the dynamic calibration on a weekly basis. The second panel compares the variance risk premium implied by the CGARCH model calibrated every Wednesday with the variance risk premium measured in the market using formula (5.3.3). In Figure 5.4 we plot the time evolution of the bucket $\text{VRP}^G(t, 22)$ and of the slope $\Delta\text{VRP}^G(t)$ of the CGARCH-implied VRP term structure.

5.4 Predictability of excess returns in the stock market with variance risk premium

The quality assessment of our variance risk premium dynamic measure includes its ability to explain the stock market returns. The predictive relationship between the variance risk premium and future stock returns has been acknowledged in several studies (Bollerslev et al. (2009); Drechsler and Yaron (2011); Du and Kapadia (2012); Camponovo et al. (2014); Bollerslev et al. (2014) among others). We test if information contained in the proposed CGARCH-implied VRP term structure can improve predictability of future long-run stock returns with respect to well-established predictors of stock returns. These include the Cyclically Adjusted Price Earning ratio (CAPE_t) of Schiller (2000) and the slope of the Treasury yield curve, also called Term Spread (TMSP_t), defined as the difference between the ten-year T-bond and the three-

[3] The measure of realized volatility is based on the methodology proposed by Corsi et al. (2013).

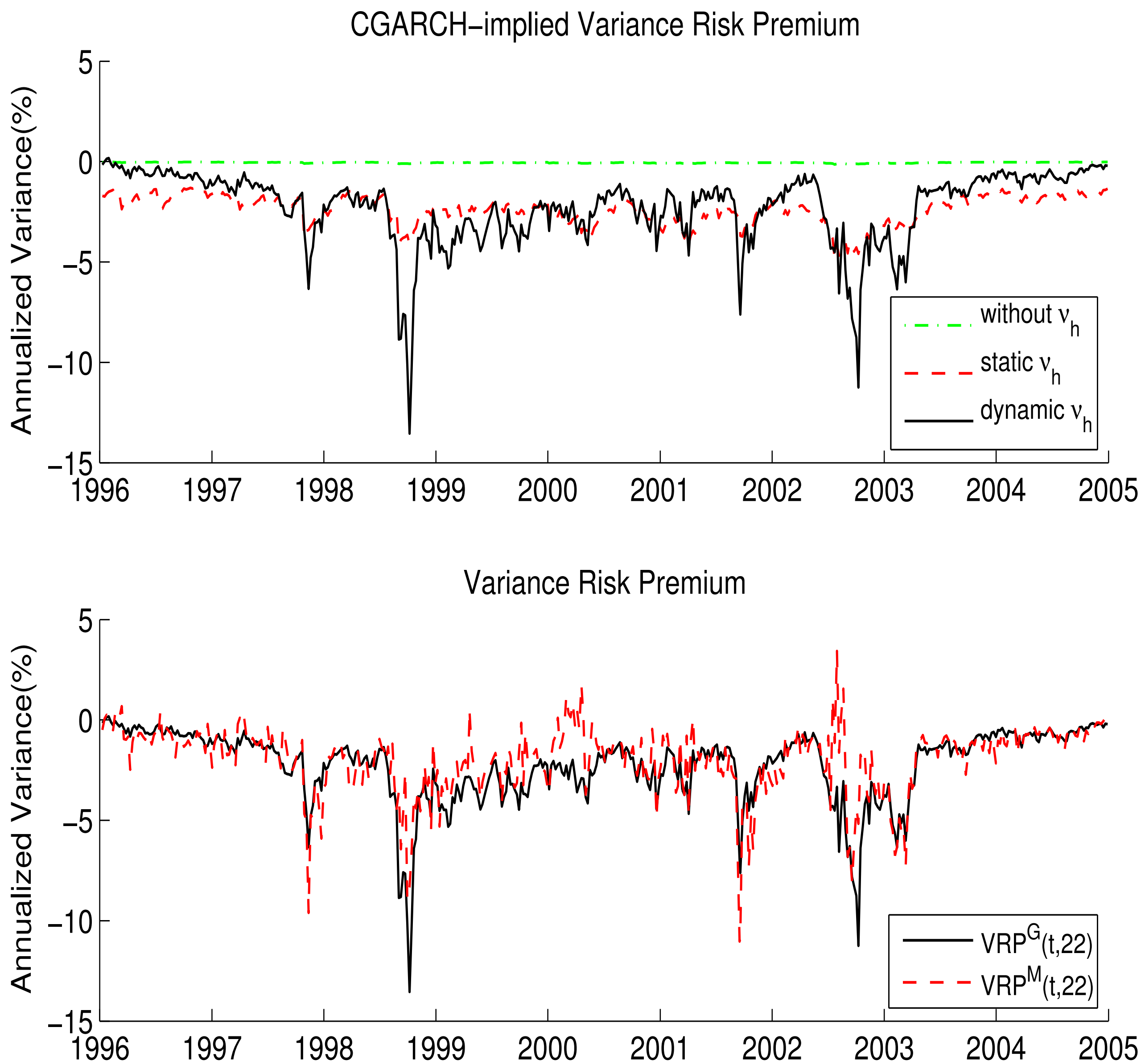

Figure 5.3: Variance risk premium with $T = 22$ from 1996 to 2005. Top panel: CGARCH-implied variance risk premium obtained with different ν_f in SDF (3.2.7): zero, constant – calibrated on the whole period of option data set – and dynamically calibrated every Wednesday. Bottom panel: comparison between $\mathrm{VRP}^G(t,22)$ (model-implied with dynamic calibration) and $\mathrm{VRP}^M(t,22)$ (market variance risk premium).

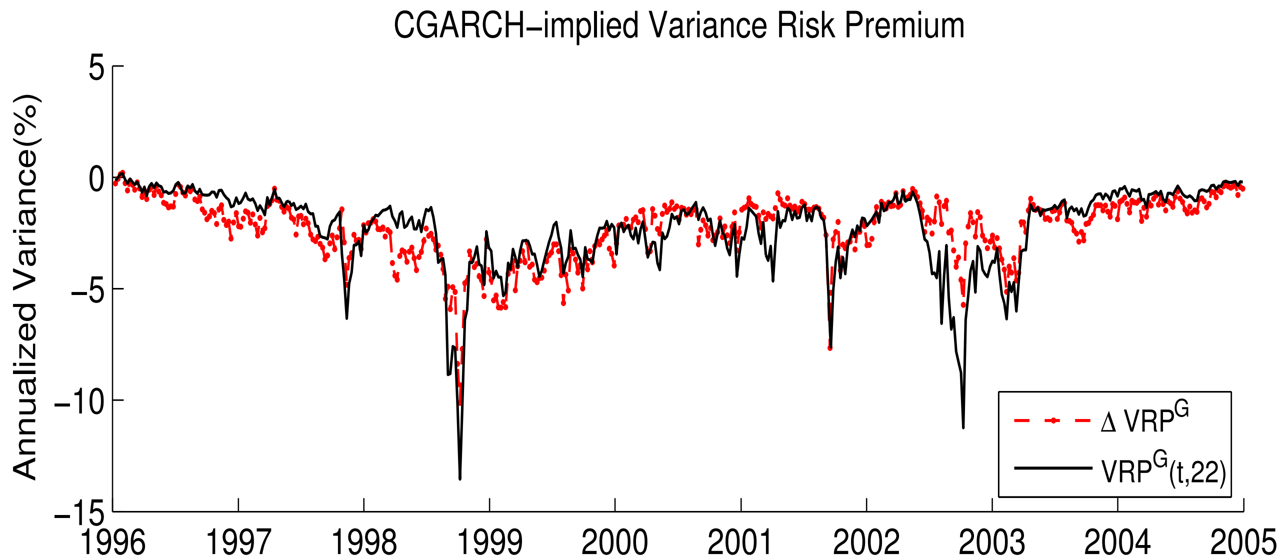

Figure 5.4: Time evolution from 1996 to 2005 of the monthly model-implied VRP, and of the difference between annual and monthly model-implied VRP (ΔVRP^G).

month T-bill yields. We also compare the predictive power of CGARCH-implied VRP with the standard market measure of the variance risk premium (5.3.3). In Table 5.1 we present the basic statistics and correlations for the considered predictors.

To test the predictive power of CGARCH-implied VRP we perform a regression of excess log-returns aggregated at different investment horizons (three months and one year) over our variance risk premium measures (5.1.6) and the other competitive predictors in various configurations. Specifically, the regression is performed with respect to the initial point of the VRP term structure $\text{VRP}^G(t, 22)$ and its slope $\Delta\text{VRP}^G(t)$, defined as the difference between $\text{VRP}^G(t, 252)$ and $\text{VRP}^G(t, 22)$. For the monthly variance risk premium $\text{VRP}^G(t, 22)$ we take the one from the last Wednesday of the month. For all regression coefficients we report Newey-West t-statistics corrected for heteroskedasticity and serial correlation effects. Finally, to assess the significance of coefficients in regressions with overlapping returns and persistent predictors

	ER_t	CAPE_t	TMSP_t	$\text{VRP}^M(t,22)$	$\text{VRP}^G(t,22)$	$\Delta\text{VRP}^G(t,252)$
mean	3.391	30.650	1.565	-1.959	-2.252	-2.253
st dev	44.172	6.540	1.186	1.638	1.670	1.363
skewness	-0.608	0.744	0.159	-1.029	-1.648	-1.275
kurtosis	4.373	2.215	1.878	6.355	7.428	5.681
				correlation		
ER_t	1.000	-0.124	-0.049	-0.189	-0.232	-0.212
CAPE_t		1.000	-0.719	-0.074	-0.219	-0.470
TMSP_t			1.000	0.140	0.253	0.400
$\text{VRP}^M(t,22)$				1.000	0.744	0.691
$\text{VRP}^G(t,22)$					1.000	0.778
$\Delta\text{VRP}^G(t,252)$						1.000

Table 5.1: Basic statistics of excess returns, CAPE, TMSP and VRP. ER_t stands for excess monthly log-returns and $\Delta\text{VRP}^G(t,252) = \text{VRP}^G(t,252) - \text{VRP}^G(t,22)$.

we report t-statistics computed following Hodrick (1992).

The results of the regressions are presented in Tables 5.2 and 5.3, from which several important observations can be drawn. First, the variance risk premium computed with the procedure proposed in Section 5.3 improves predictability of stock returns when compared to the market benchmark. Second, the forecasting power of the variance risk premium is stronger over short horizons, where it performs better than economic fundamentals. Finally we observe that enriching the VRP term structure with the information for large T (like $\text{VRP}^G(t,252)$) increases the adjusted R^2 of the linear regression. Given the relative scarcity of options with time to maturity close to one year it is difficult and questionable to recover the same information employing the model-free VRP measurements proposed in the econometric literature.

When only one variance risk premium is considered as regressor, its coefficient is always negative. This translates into the rule of thumb to buy the stock when the variance risk premium is low. Since low levels of variance risk premium usually correspond to high levels of VIX, and variance swaps in general, our findings are consistent with an old Wall Street's wisdom: "When the VIX is high, it's time to buy, when the VIX is low, it's time to go." [4]

In the case when we include in the regression the slope of the VRP term structure we observe that its coefficient becomes more significant than the one associated to the one-month maturity VRP – the absolute value of Newey-West and Hodrick t-statistics are always higher for $\Delta\text{VRP}(t)$ than for $\text{VRP}^G(t, 22)$ (from 24% to 309%). Since the coefficient is negative we conclude that if the VRP is decreasing in the long-term returns are expected to be higher. This fact can be interpreted as follows: The larger is the fear about long-term volatility risk perceived by the market, the larger is the profit from the investment demanded by investors. The empirical observation that the slope of the VRP term structure is a strong predictor of future market excess return, even much stronger than the one-month VRP, provides the argument for a modification of above market rule of thumb: When the VRP slope decreases future excess returns will be high. Then, defining $-\Delta\text{VRP}(t)$ as VRP term spread we might rephrase the Wall Street's adage "When the spread is high it's time to buy, when the spread is low it's time to go".

[4] The adage is taken from Bollerslev et al. (2014).

Quarterly return regressions

regressor:	constant	CAPE_t	TMSP_t	$\text{VRP}^M(t, 22)$	$\text{VRP}^G(t, 22)$	$\Delta\text{VRP}^G(t, 252)$	$\hat{R}^2$
Coefficient	-12.603	-	-	-	0.576	-7.287	0.090
H	-1.643	-	-	-	0.133	-1.378	0.035
NW	-1.455	-	-	-	0.222	-2.061	0.049
Coefficient	30.058	-0.858	-	-	-	-	0.036
H	1.574	-1.439	-	-	-	-	
NW	1.818	-1.753	-	-	-	-	
Coefficient	23.896	-0.926	-	-4.180	-	-	0.087
H	1.252	-1.545	-	-2.157	-	-	0.042
NW	1.387	-1.880	-	-1.892	-	-	0.028
Coefficient	26.807	-1.115	-	-	-4.996	-	0.107
H	1.418	-1.780	-	-	-2.182	-	0.051
NW	1.525	-2.122	-	-	-2.030	-	0.036
Coefficient	42.242	-2.057	-	-	3.727	-15.104	0.246
H	2.202	-2.996	-	-	0.839	-2.589	0.001
NW	2.641	-3.704	-	-	1.269	-3.451	0.000
Coefficient	89.172	-2.197	-10.642	-	-	-	0.121
H	2.607	-2.586	-1.833	-	-	-	0.035
NW	2.682	-2.727	-1.769	-	-	-	0.022
Coefficient	79.050	-2.144	-9.762	-3.549	-	-	0.158
H	2.275	-2.521	-1.664	-1.807	-	-	0.019
NW	2.519	-2.728	-1.704	-1.914	-	-	0.022
Coefficient	79.752	-2.269	-9.451	-	-4.308	-	0.174
H	2.327	-2.656	-1.608	-	-1.856	-	0.025
NW	2.573	-2.878	-1.684	-	-1.979	-	0.023
Coefficient	96.529	-3.247	-9.661	-	4.525	-15.266	0.318
H	2.747	-3.475	-1.641	-	1.005	-2.612	0.001
NW	3.339	-3.941	-2.029	-	1.471	-3.390	0.000
Coefficient	100.413	-3.403	-9.594	4.837	-	-15.376	0.326
H	2.800	-3.501	-1.637	1.574	-	-3.355	0.001
NW	3.714	-4.509	-2.076	1.606	-	-3.919	0.000

Table 5.2: Table shows the results from regressions of quarterly excess stock market returns on prior month variance risk premium and economic fundamentals. The values of CAPE_t are taken from Robert Schiller's webpage. Values of TMSP_t are computed using the ten-year Treasury yield and the three-month Treasury yield available from the webpage of Federal Reserve Board. $\text{VRP}^M(t, 22)$ corresponds to measurement obtained with equation (5.3.3) and $\text{VRP}^G(t, T)$ corresponds to measurement obtained with equation (5.1.6). Below the coefficients of regressions we report the Newey-West (NW) and Hodrick (H) t-statistics. In the last column, we report an adjusted R^2 and for multiple regressions Newey-West and Hodrick p-values of $\chi^2(n)$ $(n \geq 2)$ test that all coefficients are jointly equal to zero.

Annual return regressions

regressor:	constant	CAPE_t	TMSP_t	$\text{VRP}^M(t,22)$	$\text{VRP}^G(t,22)$	$\Delta\text{VRP}^G(t,252)$	$\hat{R}^2$
Coefficient	-0.688	-	-	-	1.863	-3.124	0.022
H	-0.121	-	-	-	0.891	-1.001	0.600
NW	-0.083	-	-	-	0.583	-1.097	0.497
Coefficient	38.768	-1.149	-	-	-	-	0.187
H	2.029	-1.765	-	-	-	-	
NW	2.527	-2.372	-	-	-	-	
Coefficient	37.304	-1.165	-	-0.993	-	-	0.197
H	1.991	-1.776	-	-0.848	-	-	0.194
NW	2.380	-2.498	-	-0.614	-	-	0.025
Coefficient	38.013	-1.209	-	-	-1.161	-	0.199
H	1.980	-1.861	-	-	-1.015	-	0.110
NW	2.383	-2.692	-	-	-0.901	-	0.004
Coefficient	48.348	-1.839	-	-	4.680	-10.113	0.378
H	2.305	-2.351	-	-	1.844	-2.498	0.047
NW	3.886	-5.629	-	-	2.299	-4.878	0.000
Coefficient	85.642	-2.211	-8.439	-	-	-	0.340
H	2.776	-2.580	-2.228	-	-	-	0.028
NW	3.773	-3.674	-2.118	-	-	-	0.001
Coefficient	84.345	-2.204	-8.326	-0.455	-	-	0.345
H	2.716	-2.575	-2.157	-0.374	-	-	0.061
NW	3.392	-3.589	-1.991	-0.327	-	-	0.000
Coefficient	84.424	-2.220	-8.285	-	-0.558	-	0.346
H	2.627	-2.613	-2.081	-	-0.434	-	0.014
NW	3.445	-3.786	-1.981	-	-0.510	-	0.000
Coefficient	95.693	-2.877	-8.426	-	5.375	-10.255	0.533
H	2.860	-2.984	-2.115	-	2.047	-2.531	0.007
NW	4.768	-5.690	-2.515	-	2.822	-4.502	0.000
Coefficient	97.109	-2.956	-8.225	4.555	-	-9.187	0.518
H	3.146	-3.297	-2.126	2.028	-	-2.521	0.003
NW	5.265	-6.349	-2.554	1.861	-	-3.112	0.000

Table 5.3: Table shows the results from regressions of annual excess stock market returns on prior month variance risk premium and economic fundamentals. The values of CAPE_t are taken from Robert Schiller's webpage. Values of TMSP_t are computed using the ten-year Treasury yield and the three-month Treasury yield available from the webpage of Federal Reserve Board. $\text{VRP}^M(t,22)$ corresponds to measurement obtained with equation (5.3.3) and $\text{VRP}^G(t,T)$ corresponds to measurement obtained with equation (5.1.6). Below the coefficients of regressions we report the Newey-West (NW) and Hodrick (H) t-statistics. In the last column, we report an adjusted R^2 and for multiple regressions Newey-West and Hodrick p-values of $\chi^2(n)$ $(n \geq 2)$ test that all coefficients are jointly equal to zero.

Conclusions

The objective of the thesis is fourfold. First, we propose and motivate a very general option pricing framework which includes a wide class of discrete time models featuring multiple components structure in both volatility and leverage and a flexible pricing kernel with multiple risk premia, in particular variance risk premium. Within this framework we characterise the recursive formulae for the analytical MGF under $\mathbb{P}$ and $\mathbb{Q}$ measures, the formal change of measure obtained using a general and flexible exponentially affine SDF, and the general characterisation of the analytical no-arbitrage condition. The usage of multi-dimensional Esscher transform is motivated by proving that it ensures Pareto optimal allocation.

Second, we introduce four new option pricing models: (i) a specific new class of realized volatility models, named LHARG, which extend the HARGL model of Corsi et al. (2013) by introducing various flexible types of leverage with heterogeneous structures and obtaining the full analytical tractability of the model, (ii) apply a new change of measure to a CGARCH model of Christoffersen et al. (2008) so that we take into account variance risk premium, (iii) extend the class of RV models by adding a jump component in volatility and its associated risk premium which provides a rapidly moving volatility factor and we label the model JLHARG, (iv) introduce a volatility model being a combination of RV and GARCH approach labeled GARCH-LHARG-RV. Moreover, we obtain an explicit one-to-one mapping between the parameters of the volatility dynamics under $\mathbb{P}$ and $\mathbb{Q}$ for models (i)-(iii) and we have closed-form option prices for all models.

Third, we empirically assess the importance of multi-components in volatility and variance risk premium in option pricing - proposed models suggest significant improvement compared to existing models in the literature. The best option pricing performance is achieved for GARCH-LHARG-RV, model being a combination of latent and realized volatility.

Finally, we propose and motivate an efficient methodology of estimating term structure of CGARCH-implied variance risk premium. We show that two-component GARCH model generates realistic hump-shape term structure of variance swap rate, contrary to the single volatility component model. Moreover we demonstrate the superiority of our VRP estimation procedure by comparing its ability to predict the stock market returns compared with the benchmark method available in the literature on variance risk premium. At the end of the thesis we provide an empiricial observation that the shape of term structure of variance risk premium (summarized by its slope) has a strong predictive power on future stock-index returns.

List of Figures

List of Tables

Bibliography

Ait-Sahalia, Y., Karaman, M., Mancini, L., 2015. The term structure of variance swaps and risk premia. Available at SSRN 2136820.

Alexander, C., 2001. Principal component analysis of volatility smiles and skews. Risk 14, S29–S32.

Alitab, D., Bormetti, G., Corsi, F., Majewski, A. A., 2015. A jump and smile ride: Continuous and jump variance risk premia in option pricing. Available at SSRN 2631155.

Andersen, T. G., Bollerslev, T., Diebold, F., Ebens, H., 2001a. The distribution of stock returns volatilities. Journal of Financial Economics 61, 43–76.

Andersen, T. G., Bollerslev, T., Diebold, F., Labys, P., 2001b. The distribution of realized exchange rate volatility. Journal of the American Statistical Association 96, 42–55.

Andersen, T. G., Bollerslev, T., Diebold, F., Labys, P., 2003. Modeling and forecasting realized volatility. Econometrica 71, 579–625.

Andersen, T. G., Bollerslev, T., Frederiksen, P., Ørregaard Nielsen, M., 2010. Continuous-time models, realized volatilities, and testable distributional implications for daily stock returns. Journal of Applied Econometrics 25 (2), 233–261.

Ané, T., Geman, H., 2000. Order flow, transaction clock, and normality of asset returns. The Journal of Finance 55 (5), 2259–2284.

Bachelier, L., 1900. Théorie de la spéculation. Gauthier-Villars.

Bakshi, G., Madan, D., Panayotov, G., 2010. Returns of claims on the upside and the viability of u-shaped pricing kernels. Journal of Financial Economics 97 (1), 130–154.

Barndorff-Nielsen, O. E., Shephard, N., 2001. Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. Journal of the Royal Statistical Society Series B (63), 167–241.

Barndorff-Nielsen, O. E., Shephard, N., 2002a. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 64, 253–280.

Barndorff-Nielsen, O. E., Shephard, N., 2002b. Estimating quadratic variation using realized variance. Journal of Applied Econometrics 17, 457–477.

Barndorff-Nielsen, O. E., Shephard, N., 2005. How accurate is the asymptotic approximation to the distribution of realized volatility? In: Andrews, D. W. F., Stock, J. H. (Eds.), Identification and Inference for Econometric Models. A Festschrift in Honour of T.J. Rothenberg. Cambridge University Press, pp. 306–331.

Barone-Adesi, G., Engle, R., Mancini, L., 2008. A GARCH option pricing with filtered historical simulation. Review of Financial Studies 21, 1223–1258.

Bates, D., 1996. Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche Mark options. Review of Financial Studies 9, 69–107.

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. The journal of political economy, 637–654.

Bollerslev, T., 1996. Generalize Autoregressive Conditional Heteroskedasticity. Journal of Econometrics 31, 307–327.

Bollerslev, T., Tauchen, G., Zhou, H., 2009. Expected stock returns and variance risk premia. Review of Financial Studies 22 (11), 4463.

Bollerslev, T., Todorov, V., 2011. Tails, fears, and risk premia. The Journal of Finance 66 (6), 2165–2211.

Bollerslev, T., Todorov, V., Xu, L., 2014. Tail risk premia and return predictability.

Bołt, W., Majewski, A. A., Szarek, T., 2012. An invariance principle for the law of the iterated logarithm for some markov chains. Studia mathematica 212 (1), 41–53.

Bormetti, G., Corsi, F., Majewski, A. A., 2015. Term structure of variance risk premium in multi-component garch models. Available at SSRN 2619278.

Brennan, M. J., 1979. The pricing of contingent claims in discrete time models. The journal of finance 34 (1), 53–68.

Broadie, M., Chernov, M., Johannes, M., 2007. Model specification and risk premia: Evidence from futures options. Journal of Finance 62, 1453–1490.

Bühlmann, H., Delbaen, F., Embrechts, P., Shiryaev, A. N., 1996. No-arbitrage, change of measure and conditional esscher transforms. CWI quarterly 9 (4), 291–317.

Bühlmann, H., Delbaen, F., Embrechts, P., Shiryaev, A. N., 1998. On esscher transforms in discrete finance models. Astin Bulletin 28 (02), 171–186.

Calvet, L., Fisher, A., 2002. Multifractality in asset returns: theory and evidence. Review of Economics and Statistics 84 (3), 381–406.

Camponovo, L., Scaillet, O., Trojani, F., 2014. Predictability hidden by anomalous observations. Swiss Finance Institute Research Paper (13-05).

Carr, P., 2005. Faqs in option pricing theory. J. Derivatives.

Carr, P., Wu, L., 2006. A tale of two indices. The Journal of Derivatives 13, 13–29.

Carr, P., Wu, L., 2009. Variance risk premiums. Review of Financial Studies 22, 1311–1341.

Christoffersen, P., Feunou, B., Jacobs, K., Meddahi, N., 2014. The economic value of realized volatility: Using high-frequency returns for option valuation. Journal of Financial and Quantitative Analysis 49 (03), 663–697.

Christoffersen, P., Heston, S., Jacobs, K., 2013. Capturing option anomalies with a variance-dependent pricing kernel. Review of Financial Studies, hht033.

Christoffersen, P., Jacobs, K., Mimouni, K., 2006. An empirical comparison of affine and non-affine models for equity index options. Manuscript, Springer.

Christoffersen, P., Jacobs, K., Ornthanalai, C., Wang, Y., 2008. Option valuation with long-run and short-run volatility components. Journal of Financial Economics 90 (3), 272–297.

Christoffersen, P., Redouane, E., Fenou, B., Jacobs, K., 2010. Options valuation with conditional heteroskedasticity and non-normality. Review of Financial Studies 23, 2139–2183.

Cochrane, J. H., 2005. Asset pricing. Vol. 1. Princeton university press Princeton, NJ.

Comte, F., Renault, E., 1998. Long memory in continuous time stochastic volatility models. Mathematical Finance 8, 291–323.

Cont, R., Da Fonseca, J., et al., 2002. Dynamics of implied volatility surfaces. Quantitative finance 2 (1), 45–60.

Cont, R., Tankov, P., 2004. Financial modelling with jump processes. Vol. 2. CRC press.

Corsi, F., 2009. A simple approximate long-memory model of realized-volatility. Journal of Financial Econometrics 7, 174–196.

Corsi, F., Fusari, N., La Vecchia, D., 2013. Realizing smiles: Options pricing with realized volatility. Journal of Financial Economics 107 (2), 284–304.

Corsi, F., Pirino, D., Renò, R., 2010. Threshold bipower variation and the impact of jumps on volatility forecasting. Journal of Econometrics 159 (2), 276 – 288.

Corsi, F., Renò, R., 2012. Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. Journal of Business & Economic Statistics 30 (3), 368–380.

Cox, J. C., Ingersoll Jr, J. E., Ross, S. A., 1985. A theory of the term structure of interest rates. Econometrica: Journal of the Econometric Society, 385–407.

Ding, Z., Granger, C. W., Engle, R. F., 1993. A long memory property of stock market returns and a new model. Journal of empirical finance 1 (1), 83–106.

Drechsler, I., Yaron, A., 2011. What's vol got to do with it. Review of Financial Studies 24 (1), 1–45.

Du, J., Kapadia, N., 2012. The tail in the volatility index. Tech. rep., Working paper.

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

Duan, J. C., Ritchken, P., Sun, Z., 2006. Approximating garch-jump models, jump-diffusion processes, and option pricing. Mathematical Finance 16, 21–52.

Duffie, D., 2010. Dynamic asset pricing theory. Princeton University Press.

Duffie, D., Filipović, D., Schachermayer, W., 2003. Affine processes and applications in finance. Annals of applied probability, 984–1053.

Duffie, D., Pan, J., Singleton, K., 2000. Transform analysis and asset pricing for affine jump-diffusions. Econometrica 68 (6), 1343–1376.

Dupire, B., 1994. Pricing with a smile. Risk 7 (1), 18–20.

Egloff, D., Leippold, M., Wu, L., 2010. The term structure of variance swap rates and optimal variance swap investments. The Journal of Financial and Quantitative Analysis 45 (5), 1279–1310.

Einstein, A., 1905. Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. Annalen der physik 322 (8), 549–560.

Engle, R., 1982. Autoregressive conditional heteroskedasticity with estimation of the variance of UK inflation. Econometrica 50, 987–1008.

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

Eraker, B., 2004. Do stock prices and volatility jump? Reconciling evidence from spot and option prices. Journal of Finance 59, 1367–1403.

Eraker, B., Johannes, M., Polson, N., 2003. The impact of jumps in volatility and returns. Journal of Finance 58, 1269–1300.

Esscher, F., 1932. On the probability function in the collective theory of risk. Skand. Aktuarie Tidskr. 15, 175–195.

Fang, F., Oosterlee, C. W., 2008. A novel pricing method for european options based on Fourier-Cosine series expansions. SIAM Journal on Scientific Computing 31, 826–848.

Filipović, D., 2013. Affine variance swap curve models. In: Seminar on Stochastic Analysis, Random Fields and Applications VII. Springer, pp. 381–393.

Filipovic, D., Gourier, E., Mancini, L., 2015. Quadratic variance swap models. forthcoming in Journal of Financial Economics.

Fisher, A., Calvet, L., Mandelbrot, B., 1997. Multifractality of dem rates. Tech. rep., Cowles Foundation Discussion Paper 1165.

Fouque, J.-P., Lorig, M. J., 2011. A fast mean-reverting correction to Heston's stochastic volatility model. SIAM Journal on Financial Mathematics 2 (1), 221–254.

Gagliardini, P., Gouriéroux, C., Renault, 2011. Efficient derivative pricing by extended method of moments. Econometrica Forthcoming.

Gatheral, J., 2011. The volatility surface: a practitioner's guide. Vol. 357. John Wiley & Sons.

Gatheral, J., Jaisson, T., Rosenbaum, M., 2014. Volatility is rough. Available at SSRN 2509457.

Geman, H., Madan, D. B., Yor, M., 2001. Time changes for lévy processes. Mathematical Finance 11 (1), 79–96.

Gerber, H. U., Shiu, E. S., 1993. Option pricing by Esscher transforms. HEC Ecole des hautes études commerciales.

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

Gourieroux, C., Jasiak, J., 2006. Autoregressive gamma process. Journal of Forecasting 25, 129–152.

Gourieroux, C., Monfort, A., 2007. Econometric specification of stochastic discount factor models. Journal of Econometrics 136 (2), 509–530.

Heston, L. S., 1993. Options with stochastic volatiltiy with applications to bond and currency options. The Review of Fiancial Studies 6, 327–343.

Heston, S., Nandi, S., 2000. A closed-form GARCH option valuation model. Review Financial Studies 13 (3), 585–625.

Hodrick, R. J., 1992. Dividend yields and expected stock returns: Alternative procedures for inference and measurement. Review of Financial studies 5 (3), 357–386.

Kardaras, C., 2010. Stochastic discount factors. Encyclopedia of Quantitative Finance.

Kou, S. G., 2002. A jump-diffusion model for option pricing. Management science 48 (8), 1086–1101.

LeBaron, B., 2001. Stochastic volatility as a simple generator of apparent financial power laws and long memory. Quantitative Finance 1 (6), 621–631.

Lehar, A., Scheicher, M., Schittenkopf, C., 2002. Garch vs. stochastic volatility: Option pricing and risk management. Journal of banking & finance 26 (2), 323–345.

Lowenstein, R., 2000. When genius failed: The rise and fall of Long-Term Capital Management. Random House LLC.

Maheu, J., McCurdy, T., 2004. News arrival, jump dynamics and volatility components for individual stock returns. Journal of Finance 59, 755–793.

Majewski, A. A., Bormetti, G., Corsi, F., 2015. Smile from the past: A general option pricing framework with multiple volatility and leverage components. Journal of Econometrics 187, 521–531.

Merton, R. C., 1976. Option pricing when underlying stock returns are discontinuosus. Journal of Financial Economics 3, 125–144.

Merton, R. C., 1980. On estimating the expected return on the market: An exploratory investigation. Journal of Financial Economics 8, 323–61.

Mueller, P., Vedolin, A., Yen, Y.-m., 2013. Bond variance risk premiums. Available at SSRN 1787478.

Müller, U., Dacorogna, M., Davé, R., Olsen, R., Pictet, O., von Weizsacker, J., 1997. Volatilities of different time resolutions - analyzing the dynamics of market components. Journal of Empirical Finance 4, 213–239.

Nelson, D. B., 1991. Conditional heteroskedasticity in asset returns: A new approach. Econometrica 59, 347–370.

Renault, E., 1997. Econometric models of option pricing errors. Econometric Society Monographs 28, 223–278.

Richter, A., Teichmann, J., 2014. Discrete time term structure theory and consistent recalibration models. http://arxiv.org/abs/1409.1830.

Rubinstein, M., 1976. The valuation of uncertain income streams and the pricing of options. The Bell Journal of Economics, 407–425.

Samuelson, P. A., 1965. Rational theory of warrant pricing. Industrial management review 6 (2), 13–32.

Schiller, R. J., 2000. Irrational exuberance. Princeton UP.

Wang, J., Eraker, B., 2015. A non-linear dynamic model of the variance risk premium. To appear in Journal of Econometrics.

Zhang, L., Aït-Sahalia, Y., Mykland, P. A., 2005. A tale of two time scales: Determining integrated volatility with noisy high frequency data. Journal of the American Statistical Association 100, 1394–1411.

Zumbach, G., Lynch, P., 2001. Heterogeneous volatility cascade in financial markets. Physica A: Statistical Mechanics and its Applications 298 (3), 521–529.

Appendix A

Proofs

A.1 MGF in general framework

We start from deriving the MGF of the log-returns $y_{t,T} = \log(S_T/S_t)$ under the risk-neutral measure $\mathbb{Q}$ conditional on the information available at time t. Applying the expression for the SDF given in (2.2.5), repeatedly and using the tower law of conditional expectation we obtain

$$
\begin{aligned}
&\varphi^{\mathbb{Q}}(t,T,z)\\
&= \mathbb{E}^{\mathbb{Q}}\left[\mathrm{e}^{zy_{t,T}}|\mathcal{F}_t\right]\\
&= \mathbb{E}^{\mathbb{P}}\left[M_{t,t+1}\dots M_{T-1,T}\mathrm{e}^{zy_{t,T}}|\mathcal{F}_t\right]\\
&= \mathbb{E}^{\mathbb{P}}\left[M_{t,t+1}\dots M_{T-2,T-1}\mathrm{e}^{zy_{t,T-1}}\mathbb{E}^{\mathbb{P}}\left[M_{T-1,T}\mathrm{e}^{zy_T}|\mathcal{F}_{T-1}\right]|\mathcal{F}_t\right]\\
&= \mathbb{E}^{\mathbb{P}}\left[\begin{matrix} M_{t,t+1}\dots M_{T-2,T-1}\mathrm{e}^{zy_{t,T-1}-\mathcal{A}(-\nu_2,-\boldsymbol{\nu}_1,\mathbf{0})-\sum_{i=1}^{p}\boldsymbol{\mathcal{B}}_i(-\nu_2,-\boldsymbol{\nu}_1,\mathbf{0})\cdot\mathbf{f}_{T-i}} \\ \times\,\mathrm{e}^{-\sum_{j=1}^{q}\boldsymbol{\mathcal{C}}_j(-\nu_2,-\boldsymbol{\nu}_1,\mathbf{0})\cdot\boldsymbol{\ell}_{T-i}}\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{-\boldsymbol{\nu}_1\cdot\mathbf{f}_T+(z-\nu_2)y_T}|\mathcal{F}_{T-1}\right]\end{matrix}|\mathcal{F}_t\right]\\
&= \mathbb{E}^{\mathbb{P}}\left[\begin{matrix} M_{t,t+1}\dots M_{T-2,T-1}\mathrm{e}^{zy_{t,T-1}+\mathcal{A}(z-\nu_2,-\boldsymbol{\nu}_1,\mathbf{0})-\mathcal{A}(-\nu_2,-\boldsymbol{\nu}_1,\mathbf{0})} \\ \times\,\mathrm{e}^{\sum_{i=1}^{p}[\boldsymbol{\mathcal{B}}_i(z-\nu_2,-\boldsymbol{\nu}_1,\mathbf{0})-\boldsymbol{\mathcal{B}}_i(-\nu_2,-\boldsymbol{\nu}_1,\mathbf{0})]\cdot\mathbf{f}_{T-i}+\sum_{j=1}^{q}[\boldsymbol{\mathcal{C}}_j(z-\nu_2,-\boldsymbol{\nu}_1,\mathbf{0})-\boldsymbol{\mathcal{C}}_j(-\nu_2,-\boldsymbol{\nu}_1,\mathbf{0})]\cdot\boldsymbol{\ell}_{T-j}}\end{matrix}|\mathcal{F}_t\right]\\
&= \mathbb{E}^{\mathbb{P}}\left[M_{t,t+1}\dots M_{T-2,T-1}\mathrm{e}^{zy_{t,T-1}+\mathrm{a}^*_{T-1}+\sum_{i=1}^{p}\mathbf{b}^*_{T-1,i}\cdot\mathbf{f}_{T-i}+\sum_{j=1}^{q}\mathbf{c}^*_{T-1,j}\cdot\boldsymbol{\ell}_{T-j}}|\mathcal{F}_t\right]\\
&= \mathbb{E}^{\mathbb{P}}\left[\begin{matrix} M_{t,t+1}\dots M_{T-3,T-2}\mathrm{e}^{zy_{t,T-2}+\mathrm{a}^*_{T-1}} \\ \times\,\mathbb{E}^{\mathbb{P}}\left[M_{T-2,T-1}\mathrm{e}^{zy_{T-1}+\sum_{i=1}^{p}\mathbf{b}^*_{T-1,i}\cdot\mathbf{f}_{T-i}+\sum_{j=1}^{q}\mathbf{c}^*_{T-1,j}\cdot\boldsymbol{\ell}_{T-j}}|\mathcal{F}_{T-2}\right]\end{matrix}|\mathcal{F}_t\right]\\
&= \dots\\
&= \mathrm{e}^{\mathrm{a}^*_t+\sum_{i=1}^{p}\mathbf{b}^*_{t,i}\cdot\mathbf{f}_{t+1-i}+\sum_{j=1}^{q}\mathbf{c}^*_{t,j}\cdot\boldsymbol{\ell}_{t+1-j}}\,.
\end{aligned}
$$

Therefore MGF of the log-returns under $\mathbb{Q}$ is of the form

$$\varphi^{\mathbb{Q}}(t,T,z) = \mathrm{e}^{\mathrm{a}_t^* + \sum_{i=1}^{p} \mathbf{b}_{t,i}^* \cdot \mathbf{f}_{t+1-i} + \sum_{j=1}^{q} \mathbf{c}_{t,j}^* \cdot \boldsymbol{\ell}_{t+1-j}}\ , \tag{A.1.1}$$

where

$$\begin{aligned}
\mathrm{a}_s^* &= \mathrm{a}_{s+1}^* + \mathcal{A}(z-\nu_2, \mathbf{b}_{s+1,1}^* - \boldsymbol{\nu}_1, \mathbf{c}_{s+1,1}^*) - \mathcal{A}(-\nu_2, -\boldsymbol{\nu}_1, \mathbf{0}) \\
\mathbf{b}_{s,i}^* &= \begin{cases} \mathbf{b}_{s+1,i+1}^* + \boldsymbol{\mathcal{B}}_i(z-\nu_2, \mathbf{b}_{s+1,1}^* - \boldsymbol{\nu}_1, \mathbf{c}_{s+1,1}^*) - \boldsymbol{\mathcal{B}}_i(-\nu_2, -\boldsymbol{\nu}_1, \mathbf{0}) & \text{if } 1 \le i \le p-1 \\ \boldsymbol{\mathcal{B}}_i(z-\nu_2, \mathbf{b}_{s+1,1}^* - \boldsymbol{\nu}_1, \mathbf{c}_{s+1,1}^*) - \boldsymbol{\mathcal{B}}_i(-\nu_2, -\boldsymbol{\nu}_1, \mathbf{0}) & \text{if } i = p \end{cases} \\
\mathbf{c}_{s,j}^* &= \begin{cases} \mathbf{c}_{s+1,j+1}^* + \boldsymbol{\mathcal{C}}_j(z-\nu_2, \mathbf{b}_{s+1,1}^* - \boldsymbol{\nu}_1, \mathbf{c}_{s+1,1}^*) - \boldsymbol{\mathcal{C}}_j(-\nu_2, -\boldsymbol{\nu}_1, \mathbf{0}) & \text{if } 1 \le j \le q-1 \\ \boldsymbol{\mathcal{C}}_j(z-\nu_2, \mathbf{b}_{s+1,1}^* - \boldsymbol{\nu}_1, \mathbf{c}_{s+1,1}^*) - \boldsymbol{\mathcal{C}}_j(-\nu_2, -\boldsymbol{\nu}_1, \mathbf{0}) & \text{if } j = q \end{cases}
\end{aligned} \tag{A.1.2}$$

and $\mathrm{a}_T^* = 0$, $\mathbf{b}_{T,i}^* = \mathbf{c}_{T,j}^* = \mathbf{0} \in \mathbb{R}^k$ for $i = 1, \dots, p$ and $j = 1, \dots, q$.

Finally, the MGF under $\mathbb{P}$ readily follows by noticing that for $\nu_1 = \nu_2 = 0$ the SDF reduces to one, therefore $\varphi^{\mathbb{P}}(t,T,z) = \varphi^{\mathbb{Q}}(t,T,z)|_{(\nu_y, \boldsymbol{\nu}_{\mathbf{f}}) = \mathbf{0}}$.

A.2 LHARG-RV

A.2.1 Affine property of LHARG process

We have

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{zy_s + b\mathrm{RV}_s + c\ell_s} | \mathcal{F}_{s-1}\right] \\
&= \mathrm{e}^{zr}\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{(z\lambda+b)\mathrm{RV}_s}\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{z\sqrt{\mathrm{RV}_s}\epsilon_s + c(\epsilon_s - \gamma\sqrt{\mathrm{RV}_s})^2} | \mathrm{RV}_s\right] | \mathcal{F}_{s-1}\right] \\
&= \mathrm{e}^{zr}\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{\left(z\lambda + b - \frac{z^2}{4c} + \gamma z\right)\mathrm{RV}_s}\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{c(\epsilon_s - (\gamma - \frac{z}{2c})\sqrt{\mathrm{RV}_s})^2} | \mathrm{RV}_s\right] | \mathcal{F}_{s-1}\right] \\
&= \mathrm{e}^{zr - \frac{1}{2}\ln(1-2c)}\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{\left(z\lambda + b + \frac{\frac{1}{2}z^2 + \gamma^2 c - 2c\gamma z}{1-2c}\right)\mathrm{RV}_s} | \mathcal{F}_{s-1}\right].
\end{aligned} \tag{A.2.1}$$

In the last equality we have used the fact that if $Z \sim \mathcal{N}(0,1)$ then

$$\mathbb{E}\left[\exp\left(x(Z+y)^2\right)\right] = \exp\left(-\frac{1}{2}\ln(1-2x) + \frac{xy^2}{1-2x}\right). \tag{A.2.2}$$

Using eq.s (8)-(9) from Gourieroux and Jasiak (2006) we obtain

$$\begin{aligned}&\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{zy_s+b\mathrm{RV}_s+c\ell_s}|\mathcal{F}_{s-1}\right]\\&=\exp\left[zr-\frac{1}{2}\ln(1-2c)-\delta\mathcal{W}(x,\theta)+\mathcal{V}(x,\theta)\left(d+\sum_{i=1}^{p}\beta_i\mathrm{RV}_{s-i}+\sum_{j=1}^{q}\alpha_j\ell_{s-j}\right)\right],\end{aligned} \tag{A.2.3}$$

where

$$\mathcal{V}(x,\theta)=\frac{\theta x}{1-\theta x}\,,\qquad \mathcal{W}(x,\theta)=\ln\left(1-x\theta\right)\,,$$

and

$$x(z,b,c)=z\lambda+b+\frac{\frac{1}{2}z^2+\gamma^2c-2c\gamma z}{1-2c}\,.$$

From a direct inspection of the relation (2.1.6), we conclude that

$$\begin{aligned}\mathcal{A}(z,b,c)&=zr-\frac{1}{2}\ln(1-2c)-\delta\mathcal{W}(x,\theta)+d\mathcal{V}(x,\theta)\,,\\ \mathcal{B}_i(z,b,c)&=\mathcal{V}(x,\theta)\beta_i\,,\\ \mathcal{C}_j(z,b,c)&=\mathcal{V}(x,\theta)\alpha_j\,.\end{aligned} \tag{A.2.4}$$

Finally, plugging the above expressions for $\mathcal{A}$, $\mathcal{B}_i$ and $\mathcal{C}_j$ in eq. (A.1.2) we readily obtain the recurrence relations for MGF under physical measure.

A.2.2 Risk-neutral dynamics

First, we write MGF under risk-neutral measure. From Proposition 14 we have that under measure $\mathbb{Q}$ the MGF for LHARG has the form

$$\varphi^{\mathbb{Q}}(t,T,z)=\exp\left(\mathrm{a}^*_t+\sum_{i=1}^{p}\mathrm{b}^*_{t,i}\mathrm{RV}_{t+1-i}+\sum_{j=1}^{q}\mathrm{c}^*_{t,j}\ell_{t+1-j}\right)\,,$$

where

$$
\begin{aligned}
\mathrm{a}^*_s =&\mathrm{a}^*_{s+1} + zr - \frac{1}{2}\ln(1-2\mathrm{c}^*_{s+1,1}) - \delta\mathcal{W}(\mathrm{x}^*_{s+1},\theta) + \delta\mathcal{W}(\mathrm{y}^*_{s+1},\theta) \\
&+ d\mathcal{V}(\mathrm{x}^*_{s+1},\theta) - d\mathcal{V}(\mathrm{y}^*_{s+1},\theta) \\
\mathrm{b}^*_{s,i} =& \begin{cases} \mathrm{b}^*_{s+1,i+1} + \left(\mathcal{V}(\mathrm{x}^*_{s+1},\theta) - \mathcal{V}(\mathrm{y}^*_{s+1},\theta)\right)\beta_i & \text{for } 1 \le i \le p-1 \\ \left(\mathcal{V}(\mathrm{x}^*_{s+1},\theta) - \mathcal{V}(\mathrm{y}^*_{s+1},\theta)\right)\beta_i & \text{for } i = p \end{cases} \\
\mathrm{c}^*_{s,i} =& \begin{cases} \mathrm{c}^*_{s+1,i+1} + \left(\mathcal{V}(\mathrm{x}^*_{s+1},\theta) - \mathcal{V}(\mathrm{y}^*_{s+1},\theta)\right)\alpha_i & \text{for } 1 \le i \le q-1 \\ \left(\mathcal{V}(\mathrm{x}^*_{s+1},\theta) - \mathcal{V}(\mathrm{y}^*_{s+1},\theta)\right)\alpha_i & \text{for } i = q\,, \end{cases}
\end{aligned}
\tag{A.2.5}
$$

with

$$
\begin{aligned}
\mathrm{x}^*_{s+1} &= (z-\nu_2)\lambda + \mathrm{b}^*_{s+1,1} - \nu_1 + \frac{\frac{1}{2}(z-\nu_2)^2 + \gamma^2\mathrm{c}^*_{s+1,1} - 2\mathrm{c}^*_{s+1,1}\gamma(z-\nu_2)}{1-2\mathrm{c}^*_{s+1,1}}\,, \\
\mathrm{y}^*_{s+1} &= -\nu_2\lambda - \nu_1 + \frac{1}{2}\nu_2^2\,,
\end{aligned}
$$

and terminal conditions $\mathrm{a}^*_T = \mathrm{b}^*_{T,i} = \mathrm{c}^*_{T,j} = 0$ for $i = 1,\ldots,p$ and $j = 1,\ldots,q$.

To derive the mapping of the parameters under which the risk-neutral MGF is formally equivalent to the physical MGF, we need to compare eq. (A.2.5) to eq. (3.1.19). In particular we have to find a set of starred parameters for which the recursions under $\mathbb{P}$ correspond to the expressions under $\mathbb{Q}$. More precisely, after defining

$$
\mathrm{x}^{**}_{s+1} = z\lambda^* + \mathrm{b}^*_{s+1,1} + \frac{\frac{1}{2}z^2 + (\gamma^*)^2\mathrm{c}^*_{s+1,1} - 2\mathrm{c}^*_{s+1,1}\gamma^* z}{1-2\mathrm{c}^*_{s+1,1}}\,,
$$

the following relations have to hold

$$
\delta\left(\mathcal{W}(\mathrm{x}^*_{s+1},\theta) - \mathcal{W}(\mathrm{y}^*,\theta)\right) = \delta^*\mathcal{W}(\mathrm{x}^{**}_{s+1},\theta^*)\,, \tag{A.2.6}
$$

$$
\beta_i\left(\mathcal{V}(\mathrm{x}^*_{s+1},\theta) - \mathcal{V}(\mathrm{y}^*,\theta)\right) = \beta^*_i\mathcal{V}(\mathrm{x}^{**}_{s+1},\theta^*)\,, \tag{A.2.7}
$$

$$
\alpha_j\left(\mathcal{V}(\mathrm{x}^*_{s+1},\theta) - \mathcal{V}(\mathrm{y}^*,\theta)\right) = \alpha^*_j\mathcal{V}(\mathrm{x}^{**}_{s+1},\theta^*)\,, \tag{A.2.8}
$$

$$
d\left(\mathcal{V}(\mathrm{x}^*_{s+1},\theta) - \mathcal{V}(\mathrm{y}^*,\theta)\right) = d^*\mathcal{V}(\mathrm{x}^{**}_{s+1},\theta^*)\,, \tag{A.2.9}
$$

with $y^* = -\lambda^2/2 - \nu_1 + \frac{1}{8}$. Eq. (A.2.6) can be rewritten as

$$\delta \log \left[1 - \frac{\theta}{1-\theta y^*} \left(x^*_{s+1} - y^* \right) \right] = \delta^* \log \left(1 - \theta^* x^{**}_{s+1} \right) ,$$

from which we obtain the sufficient conditions $\delta^* = \delta$, $\theta^* = \theta/(1-\theta y^*)$, and $x^*_{s+1} - y^* = x^{**}_{s+1}$. It is possible to verify by substitution that the latter relation is satisfied posing $\lambda^* = -1/2$ and $\gamma^* = \gamma + \lambda + 1/2$. The relation (A.2.7) is equivalent to

$$\frac{\beta_i}{1-\theta y^*} \frac{\theta}{1-\theta y^*} \frac{x^*_{s+1} - y^*}{\left[1 - \theta/(1-\theta y^*) \left(x^*_{s+1} - y^* \right) \right]} = \beta^*_i \frac{\theta^* x^{**}_{s+1}}{1 - \theta^* x^{**}_{s+1}} ,$$

which implies $\beta^*_i = \beta_i/(1-\theta y^*)$. Similar reasoning applies for eq.s (A.2.8) and (A.2.9).

A.3 k-CGARCH(p, q)

A.3.1 Affine property

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}} \left[e^{zr + z\lambda \mathcal{S}(\mathbf{h}_{s+1}) + z\sqrt{\mathcal{S}(\mathbf{h}_{s+1})}\epsilon_{s+1} + \mathbf{b}\mathbf{h}_{s+2} + \mathbf{c}\boldsymbol{\ell}_{s+1}} | \mathcal{F}_s \right] \\
&= e^{zr + z\lambda \mathcal{S}(\mathbf{h}_{s+1}) + \mathbf{bd} + \sum_{i=1}^{p} \mathbf{b}\mathbf{M}_i \mathbf{h}_{s+2-i} + \sum_{i=2}^{q} \mathbf{b}\mathbf{N}_i \boldsymbol{\ell}_{s+2-i}} \mathbb{E}^{\mathbb{P}} \left[e^{z\sqrt{\mathcal{S}(\mathbf{h}_{s+1})}\epsilon_{s+1} + (\mathbf{b}\mathbf{N}_1 + \mathbf{c})\boldsymbol{\ell}_{s+1}} | \mathcal{F}_s \right] \\
&= e^{zr + z\lambda \mathcal{S}(\mathbf{h}_{s+1}) + \mathbf{bd} + \sum_{i=1}^{p} \mathbf{b}\mathbf{M}_i \mathbf{h}_{s+2-i} + \sum_{i=2}^{q} \mathbf{b}\mathbf{N}_i \boldsymbol{\ell}_{s+2-i} + \left(\sum_{i=1}^{k}\sum_{j=1}^{k} (b_i n_{i,j} + c_j)\gamma_j^2 - \frac{\left(\sum_{i=1}^{k}\sum_{j=1}^{k} \left((b_i n_{i,j} + c_j)\gamma_j \right) - 0.5z \right)^2}{\left(\sum_{i=1}^{k}\sum_{j=1}^{k} (b_i n_{i,j} + c_j) \right)} \right) \mathcal{S}(\mathbf{h}_{s+1})} \\
&\times \mathbb{E}^{\mathbb{P}} \left[e^{\left(\sum_{i=1}^{k}\sum_{j=1}^{k} (b_i n_{i,j} + c_j) \right) \left(\epsilon_s - \frac{\sum_{i=1}^{k}\sum_{j=1}^{k} \left((b_i n_{i,j} + c_j)\gamma_j \right) - 0.5z}{\left(\sum_{i=1}^{k}\sum_{j=1}^{k} (b_i n_{i,j} + c_j) \right)} \sqrt{\mathcal{S}(\mathbf{h}_{s+1})} \right)^2} | \mathcal{F}_s \right]
\end{aligned} \tag{A.3.1}$$

Using property (A.2.2) we obtain

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[e^{zr+z\lambda\mathcal{S}(\mathbf{h}_{s+1})+z\sqrt{\mathcal{S}(\mathbf{h}_{s+1})}\epsilon_{s+1}+\mathbf{b}\mathbf{h}_{s+2}+\mathbf{c}\boldsymbol{\ell}_{s+1}}|\mathcal{F}_s\right] \\
&= \exp\left(\begin{array}{l}
zr + z\lambda\mathcal{S}\left(\mathbf{h}_{s+1}\right) + \mathbf{bd} + \sum_{i=1}^{p}\mathbf{bM}_i\mathbf{h}_{s+2-i} + \sum_{i=2}^{q}\mathbf{bN}_i\boldsymbol{\ell}_{s+2-i} \\
+ \left(\sum_{i=1}^{k}\sum_{j=1}^{k}(b_i n_{i,j}+c_j)\gamma_j^2 - \frac{\left(\sum_{i=1}^{k}\sum_{j=1}^{k}\left((b_i n_{i,j}+c_j)\gamma_j\right)-0.5z\right)^2}{\left(\sum_{i=1}^{k}\sum_{j=1}^{k}(b_i n_{i,j}+c_j)\right)}\right)\mathcal{S}\left(\mathbf{h}_{s+1}\right) \\
-\frac{1}{2}\ln\left(1-2\left(\sum_{i=1}^{k}\sum_{j=1}^{k}(b_i n_{i,j}+c_j)\right)\right) \\
+\frac{\left(\sum_{i=1}^{k}\sum_{j=1}^{k}\left((b_i n_{i,j}+c_j)\gamma_j\right)-0.5z\right)^2}{\left(1-2\left(\sum_{i=1}^{k}\sum_{j=1}^{k}(b_i n_{i,j}+c_j)\right)\right)\left(\sum_{i=1}^{k}\sum_{j=1}^{k}(b_i n_{i,j}+c_j)\right)}\mathcal{S}\left(\mathbf{h}_{s+1}\right)
\end{array}\right)
\end{aligned}
\tag{A.3.2}
$$

Taking into account that $\mathcal{S}\left(\mathbf{h}_s\right) = h_s^{(1)} + \ldots + h_s^{(k)}$ we finally obtain

$$
\begin{aligned}
&\boldsymbol{\mathcal{A}}(z,\mathbf{b},\mathbf{c}) = zr + \mathbf{bd} - \frac{1}{2}\ln\left(1-2\left(\sum_{i=1}^{k}\sum_{j=1}^{k}(b_i n_{i,j}+c_j)\right)\right) \\
&\boldsymbol{\mathcal{B}}_1(z,\mathbf{b},\mathbf{c}) = \mathbf{bM}_1 + \left(z\lambda + \sum_{i=1}^{k}\sum_{j=1}^{k}(b_i n_{i,j}+c_j)\gamma_j^2 + 2\frac{\left(\sum_{i=1}^{k}\sum_{j=1}^{k}\left((b_i n_{i,j}+c_j)\gamma_j\right)-0.5z\right)^2}{1-2\left(\sum_{i=1}^{k}\sum_{j=1}^{k}(b_i n_{i,j}+c_j)\right)}\right)\mathbf{1} \\
&\boldsymbol{\mathcal{B}}_i(z,\mathbf{b},\mathbf{c}) = \mathbf{bM}_i \qquad \text{for } i \in \{2,\ldots,p\} \\
&\boldsymbol{\mathcal{C}}_1(z,\mathbf{b},\mathbf{c}) = 0 \\
&\boldsymbol{\mathcal{C}}_j(z,\mathbf{b},\mathbf{c}) = \mathbf{bN}_j \qquad \text{for } j \in \{2,\ldots,q\}
\end{aligned}
\tag{A.3.3}
$$

where $\mathbf{1} = (1,\ldots,1)' \in \mathbb{R}^k$, $\mathbf{b},\mathbf{c} \in \mathbb{R}^k$, $n_{i,j}$ are elements of matrix $\mathbf{N}_1$ and $\cdot$ stands for the scalar product in $\mathbb{R}^k$.

A.3.2 Moment Generating Function

Under the physical measure $\mathbb{P}$ the MGF of the log-returns $y_{t,T} = \log(S_T/S_t)$ conditional on the information available at time t is of the form

$$\varphi^{\mathbb{P}}(t,T,z) = \mathrm{e}^{\mathrm{a}_t + \sum_{i=1}^{p} \mathbf{b}_{t,i}\cdot \mathbf{h}_{t+2-i} + \sum_{j=2}^{q} \mathbf{c}_{t,j}\cdot \boldsymbol{\ell}_{t+1-j}}\,, \tag{A.3.4}$$

where

$$\begin{aligned}
\mathrm{a}_s &= \mathrm{a}_{s+1} + zr + \mathbf{b}_{s+1,1}\mathbf{d} - \frac{1}{2}\ln\left(1 - 2\left(\sum_{i=1}^{k}\sum_{j=1}^{k}(\mathrm{b}_{s+1}^{((i))}n_{i,j} + \mathrm{c}_{s+1}^{((j))})\right)\right)\\
\mathbf{b}_{s,i} &= \begin{cases} \mathbf{b}_{s+1,i+1} + \mathbf{b}_{s+1,1}M_i + X & \text{if } i = 1\\ \mathbf{b}_{s+1,i+1} + \mathbf{b}_{s+1,1}M_i & \text{if } 2 \le i \le p-1\\ \boldsymbol{\mathcal{B}}_i(z, \mathbf{b}_{s+1,1}, \mathbf{c}_{s+1,1}) & \text{if } i = p \end{cases}\\
\mathbf{c}_{s,j} &= \begin{cases} \mathbf{c}_{s+1,j+1} + \mathbf{b}_{s+1,1}N_j & \text{if } 2 \le j \le q-1\\ \mathbf{b}_{s+1,1}N_j & \text{if } j = q \end{cases}
\end{aligned} \tag{A.3.5}$$

where

$$X = \left(z\lambda + \sum_{i=1}^{k}\sum_{j=1}^{k}(\mathrm{b}_{s+1}^{((i))}n_{i,j} + \mathrm{c}_{s+1}^{((j))})\gamma_j^2 + 2\frac{\left(\sum_{i=1}^{k}\sum_{j=1}^{k}\left((\mathrm{b}_{s+1}^{((i))}n_{i,j} + \mathrm{c}_{s+1}^{((j))})\gamma_j\right) - 0.5z\right)^2}{1 - 2\left(\sum_{i=1}^{k}\sum_{j=1}^{k}(\mathrm{b}_{s+1}^{((i))}n_{i,j} + \mathrm{c}_{s+1}^{((j))})\right)}\right)\mathbf{1} \tag{A.3.6}$$

and $\mathrm{a}_T = 0$, $\mathbf{b}_{T,i} = \mathbf{c}_{T,j} = \mathbf{0} \in \mathbb{R}^k$ for $i = 1,\ldots,p$ and $j = 1,\ldots,q$, $n_{i,j}$ are elements of matrix $\mathbf{N}_1$.

In the case of a process with one time lag in volatility and leverage (k-CGARCH$(1,1)$) the MGF of the log-returns $y_{t,T} = \log(S_T/S_t)$ conditional on the information available at time t is of the form

$$\varphi^{\mathbb{P}}(t,T,z) = \mathrm{e}^{\mathrm{a}_t + \mathbf{b}_t\cdot \mathbf{h}_{t+1}}\,, \tag{A.3.7}$$

where

$$\begin{aligned} \mathrm{a}_s &= \mathrm{a}_{s+1} + zr + \mathbf{b_{s+1}d} - \frac{1}{2}\ln\left(1 - 2\left(\sum_{i=1}^{k}\sum_{j=1}^{k}\mathrm{b}_{s+1}^{((i))}n_{i,j}\right)\right), \\ \mathbf{b}_s &= \mathbf{b}_{s+1} + \mathbf{b}_{s+1}M + X\,, \end{aligned} \tag{A.3.8}$$

with

$$X = \left(z\lambda + \sum_{i=1}^{k}\sum_{j=1}^{k}\mathrm{b}_{s+1}^{((i))}n_{i,j}\gamma_j^2 + 2\frac{\left(\sum_{i=1}^{k}\sum_{j=1}^{k}\mathrm{b}_{s+1}^{((i))}n_{i,j}\gamma_j - 0.5z\right)^2}{1 - 2\sum_{i=1}^{k}\sum_{j=1}^{k}\mathrm{b}_{s+1}^{((i))}n_{i,j}}\right)\mathbf{1}\,, \tag{A.3.9}$$

and $\mathrm{a}_T = 0$, $\mathbf{b}_T = \mathbf{0} \in \mathbb{R}^k$.

A.3.3 No arbitrage condition

We derive the no-arbitrage condition for a general SDF

$$M_{s,s+1} = \frac{\mathrm{e}^{-\boldsymbol{\nu}_{\mathbf{h}}\cdot\mathbf{h}_{s+2} - \nu_y y_{s+1}}}{\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{-\boldsymbol{\nu}_{\mathbf{h}}\cdot\mathbf{h}_{s+2} - \nu_y y_{s+1}}\middle|\mathcal{F}_s\right]}, \tag{A.3.10}$$

where $\boldsymbol{\nu}_{\mathbf{h}} \in \mathbb{R}^k$. To derive the no-arbitrage condition we plug functions (A.3.3) to conditions (2.2.3). Except from the following condition

$$\boldsymbol{\mathcal{B}}_1(1 - \nu_y, -\boldsymbol{\nu}_{\mathbf{h}}, \mathbf{0}) = \boldsymbol{\mathcal{B}}_1(-\nu_y, -\boldsymbol{\nu}_{\mathbf{h}}, \mathbf{0}), \tag{A.3.11}$$

all conditions are trivial. After doing some computations condition (A.3.11) translates into no-arbitrage condition

$$\nu_y = \lambda + \frac{1}{2} + 2\sum_{i=1}^{k}\nu_h^{(i)}\sum_{j=1}^{k}n_{i,j}(\gamma_j + \lambda), \tag{A.3.12}$$

where $\boldsymbol{\nu}_{\mathbf{h}} = (\nu_h^{(1)}, \ldots, \nu_h^{(k)})$. When we apply SDF (3.2.7) we need to substitute vector $\boldsymbol{\nu}_{\mathbf{h}}$ by a scalar ν_h.

A.3.4 Risk-neutral dynamics

Firstly we compute

$$\varphi^{\mathbb{Q}}(t,t+1,z) = \mathbb{E}^{\mathbb{Q}}\left[e^{zy_{t+1}}|\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{P}}\left[M_{t,t+1}e^{zy_{t+1}}|\mathcal{F}_t\right], \tag{A.3.13}$$

where $M_{t,t+1}$ is defined in (A.3.10). After some computations, using no-arbitrage condition (A.3.12) and the form of functions $\mathcal{A}$, $\mathcal{B}_i$, $\mathcal{C}_j$ for $i = 1,\dots,p$, $j = 1,\dots,q$, we obtain

$$\varphi^{\mathbb{Q}}(t,t+1,z) = \exp\left(zr - \frac{\mathcal{S}\left(\mathbf{h}_{t+1}\right)}{2+4\sum_{i=1}^{k}\nu_h^{(i)}\sum_{j=1}^{k}n_{i,j}}z + \frac{\mathcal{S}\left(\mathbf{h}_{t+1}\right)}{2+4\sum_{i=1}^{k}\nu_h^{(i)}\sum_{j=1}^{k}n_{i,j}}z^2\right). \tag{A.3.14}$$

From (A.3.14) we deduce that

$$y_t = r - \frac{1}{2}\mathcal{S}\left(\mathbf{h}_t^*\right) + \epsilon_t^*\sqrt{\mathcal{S}\left(\mathbf{h}_t^*\right)}, \tag{A.3.15}$$

where

$$\mathbf{h}_t^* = \frac{\mathbf{h}_t}{1+2\sum_{i=1}^{k}\nu_h^{(i)}\sum_{j=1}^{k}n_{i,j}}, \tag{A.3.16}$$

and ϵ_t^* has distribution $\mathcal{N}(0,1)$ under measure $\mathbb{Q}$. Comparing (A.3.15) with (3.2.2) we obtain the following relation between ϵ_t^* and ϵ_t:

$$\epsilon_t^* = \sqrt{1+2\sum_{i=1}^{k}\nu_h^{(i)}\sum_{j=1}^{k}n_{i,j}}\left(\epsilon_t + \left(\frac{1}{2(1+2\sum_{i=1}^{k}\nu_h^{(i)}\sum_{j=1}^{k}n_{i,j})} + \lambda\right)\sqrt{\mathcal{S}\left(\mathbf{h}_t\right)}\right). \tag{A.3.17}$$

We conclude by applying no-arbitrage condition (3.2.8), the relations (A.3.16) and (A.3.17) in dynamics of $\mathbf{h}_t$ described by equation (3.2.3) to obtain the following mappings of parameters:

$$\begin{aligned}
\lambda^* &= -1/2, \\
\mathbf{d}^* &= \frac{\mathbf{d}}{1 + 2\sum_{i=1}^{k} \nu_h^{(i)} \sum_{j=1}^{k} n_{i,j}}, \\
\mathbf{M}_i^* &= \mathbf{M}_i \text{ for } i = 1, \ldots, p \\
\mathbf{N}_j^* &= \frac{\mathbf{N}_j}{\left(1 + 2\sum_{i=1}^{k} \nu_h^{(i)} \sum_{j=1}^{k} n_{i,j}\right)^2} \text{ for } i = 1, \ldots, q \\
\gamma_l^* &= \gamma_l + \nu_y + 2\sum_{i=1}^{k} \nu_h^{(i)} \sum_{j=1}^{k} n_{i,j}(\gamma_l - \gamma_j) \text{ for } 1 \leq l \leq k.
\end{aligned} \tag{A.3.18}$$

The dynamics of the process under risk-neutral measure is described by equations (A.3.15)-(A.3.16). From there we can see that the dynamics of each factor of volatility $\mathrm{h}_t^{(i*)}$ is equal the dynamics of $\mathrm{h}_t^{(i)}$ divided by some constant, the same for every factor $i \in 1, \ldots, k$. Considering $k+1$-dimensional Esscher transform does not improve the flexibility of our model which would suffer from identification problem and it is enough to apply 2-dimensional Esscher transform substituting vector $\boldsymbol{\nu}_{\mathbf{h}}$ by a scalar ν_h.

Obtaining efficient $k+1$-dimensional change of measure would be possible by increasing the number of sources of randomness in the dynamics under measure $\mathbb{P}$. Let us consider a class of GARCH models with multi time scales associated with independent innovations, which we define as MTS-GARCH(p, q), where p and q stands for the order of regression:

$$y_{t+1} = r + \lambda \sum_{i=1}^{k} \mathrm{h}_t^{(i)} + \sum_{i=1}^{k} \sqrt{\mathrm{h}_t^{(i)}}\ \epsilon_{t+1}^{(i)}, \tag{A.3.19}$$

where r is the risk-free rate, λ is the market price of risk, and $\epsilon_t^{(1)}, \ldots, \epsilon_t^{(k)}$ are i.i.d. $\mathcal{N}(0, 1)$. We model $\mathbf{h}_{t+1}$ as

$$\mathbf{h}_{t+1} = \mathbf{d} + \sum_{i=1}^{p} \mathbf{M}_i \mathbf{h}_{t+1-i} + \sum_{j=1}^{q} \mathbf{N}_j \boldsymbol{\ell}_{t+1-j}, \tag{A.3.20}$$

where $\mathbf{M}_i, \mathbf{N}_j \in \mathbb{R}^{k\times k}$ for $i = 1, \ldots, p$ and $j = 1, \ldots, q$, $\mathbf{d} \in \mathbb{R}^k$, and vectors $\boldsymbol{\ell}_{t-j}$ are of the form

$$\boldsymbol{\ell}_{t+1-j} = \begin{bmatrix} \left(\epsilon^{(1)}_{t+1-j} - \gamma_1 \sqrt{\mathrm{h}^{(1)}_{t+1-j}}\right)^2 \\ \vdots \\ \left(\epsilon^{(k)}_{t+1-j} - \gamma_k \sqrt{\mathrm{h}^{(k)}_{t+1-j}}\right)^2 \end{bmatrix}. \tag{A.3.21}$$

Since innovations in (A.3.19) are independent we have that $y \sim \mathcal{N}\left(r + \mathcal{S}\left(\mathbf{h}_t\right), \sqrt{\mathcal{S}\left(\mathbf{h}_t\right)}\right)$ and equation (A.3.19) is equivalent to (3.2.2). The dynamics of the 2TS-GARCH, analog of CJOW, is described by following equations:.

$$\begin{aligned} y_{t+1} &= r + \lambda\left(h^{(1)}_{t+1} + h^{(2)}_{t+1}\right) + \sqrt{h^{(1)}_{t+1}}\epsilon^{(1)}_{t+1} + \sqrt{h^{(2)}_{t+1}}\epsilon^{(2)}_{t+1}, \\ h^{(1)}_{t+1} &= \beta_1 h^{(1)}_t + \alpha_1\left(\left(\epsilon^{(1)}_t\right)^2 - 1 - 2\gamma_1\epsilon^{(1)}_t\sqrt{h^{(1)}_t + h^{(2)}_t}\right), \\ h^{(2)}_{t+1} &= \omega + \beta_2 h^{(2)}_t + \alpha_2\left(\left(\epsilon^{(2)}_t\right)^2 - 1 - 2\gamma_2\epsilon^{(2)}_t\sqrt{h^{(1)}_t + h^{(2)}_t}\right). \end{aligned} \tag{A.3.22}$$

Multiple independent innovation enables to apply efficiently $k+1$ dimensional Esscher transform, so that we can associate risk with each component of volatility. Moreover, class of MTS-GARCH(p, q) processes belongs to generalised affine processes satisfying the generalised version of Assumption 6 in Section 2.1, and as a consequence we are able to use a generalisation of the framework presented in Chapter 2.

A.3.5 'U shape' of pricing kernel

Since log-returns have Gaussian distribution under measure $\mathbb{P}$, the physical probability density function of log-return y_t conditioned on variance $\mathcal{S}\left(\mathbf{h}_t\right)$ can be written as follows

$$f(y_t) = \frac{1}{\sqrt{2\pi\mathcal{S}\left(\mathbf{h}_t\right)}} \exp\left(-\frac{\left(y_t - r - \lambda\mathcal{S}\left(\mathbf{h}_t\right)\right)^2}{2\mathcal{S}\left(\mathbf{h}_t\right)}\right). \tag{A.3.23}$$

On the other hand, log-returns have Gaussian distribution also under measure $\mathbb{Q}$, and due to the fact that $\mathcal{S}\left(\mathbf{h^*}_t\right) = \xi^2\mathcal{S}\left(\mathbf{h}_t\right)$, the risk-neutral probability density function of log-return y_t

conditioned on variance $\mathcal{S}(\mathbf{h}_t)$ can be written as follows

$$f^*(y_t) = \frac{1}{\sqrt{2\pi\xi^2\mathcal{S}(\mathbf{h}_t)}} \exp\left(-\frac{\left(y_t - r + \frac{\xi^2}{2}\mathcal{S}(\mathbf{h}_t)\right)^2}{2\xi^2\mathcal{S}(\mathbf{h}_t)}\right). \tag{A.3.24}$$

Dividing (A.3.24) by (A.3.23) and taking the logarithm we obtain the formula (3.2.14).

A.4 GARCH-LHARG-RV

A.4.1 Moment Generating Function under physical measure

First, we rewrite equation (3.3.4) as

$$\Theta(\mathbf{RV}_t, \mathbf{L}_t) = d + \hat{\beta}_d \mathrm{RV}_t^{(d)} + \hat{\beta}_w \mathrm{RV}_t^{(w)} + \hat{\beta}_m \mathrm{RV}_t^{(m)} + \alpha_d \ell_t^{(d)} + \alpha_w \ell_t^{(w)} + \alpha_m \ell_t^{(m)}, \tag{A.4.1}$$

where

$$\begin{aligned}
\ell_t^{(d)} &= \left(\epsilon_t - \gamma\sqrt{\mathrm{RV}_t + h_t}\right)^2, \\
\ell_t^{(w)} &= \frac{1}{4}\sum_{i=1}^{4}\left(\epsilon_{t-i} - \gamma\sqrt{\mathrm{RV}_{t-i} + h_{t-i}}\right)^2, \\
\ell_t^{(m)} &= \frac{1}{17}\sum_{i=5}^{21}\left(\epsilon_{t-i} - \gamma\sqrt{\mathrm{RV}_{t-i} + h_{t-i}}\right)^2.
\end{aligned}$$

and $d = -(\alpha_d + \alpha_w + \alpha_m)$, $\hat{\beta}_l = \beta_l - \alpha_l\gamma^2$ for $l = d, w, m$. Then we rewrite (A.4.1) as

$$\Theta(\mathbf{RV^c}_t, \mathbf{L}_t) = d + \sum_{i=1}^{22}\beta_i \mathrm{RV}^c_{t+1-i} + \sum_{j=1}^{22}\alpha_j\left(\epsilon_{t+1-j} - \gamma\sqrt{\mathrm{RV}_{t+1-j}}\right)^2 \tag{A.4.2}$$

with

$$\beta_i = \begin{cases} \hat{\beta}_d & \text{for } i = 1 \\ \hat{\beta}_w/4 & \text{for } 2 \le i \le 5 \\ \hat{\beta}_m/17 & \text{for } 6 \le i \le 22 \end{cases} \qquad \alpha_i = \begin{cases} \alpha_d & \text{for } i = 1 \\ \alpha_w/4 & \text{for } 2 \le i \le 5 \\ \alpha_m/17 & \text{for } 6 \le i \le 22 \end{cases}. \tag{A.4.3}$$

Now we can show that GARCH-LHARG process satisfies an affine property.

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{zy_{s+1}+b_1h_{s+2}+b_2RV_{s+1}+c\ell_{s+1}}|\mathcal{F}_s\right]\\
&=\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{zr+z\lambda(h_{s+1}+RV_{s+1})+z\sqrt{h_{s+1}+RV_{s+1}}\epsilon_{s+1}+b_1h_{s+2}+b_2RV_{s+1}+c\ell_{s+1}}|\mathcal{F}_s\right]\\
&=\mathrm{e}^{zr+z\lambda h_{s+1}+b_1\omega+b_1\beta_hh_{s+1}}\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{z\lambda RV_{s+1}+z\sqrt{h_{s+1}+RV_{s+1}}\epsilon_{s+1}+b_1\alpha_h\left(\epsilon_{s+1}-\gamma_h\sqrt{h_{s+1}+RV_{s+1}}\right)^2+b_2RV_{s+1}+c\ell_{s+1}}|\mathcal{F}_s\right]\\
&=\mathrm{e}^{zr+z\lambda h_{s+1}+b_1\omega+b_1\beta_hh_{s+1}}\\
&\times\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{\left(b_1\alpha_h\gamma_h^2+c\gamma^2-\frac{(b_1\alpha_h\gamma_h+c\gamma-0.5z)^2}{b_1\alpha_h+c}\right)(h_{s+1}+RV_{s+1})+z\lambda RV_{s+1}+b_2RV_{s+1}+(b_1\alpha_h+c)\left(\epsilon_{s+1}-\frac{b_1\alpha_h\gamma_h+c\gamma-0.5z}{b_1\alpha_h+c}\sqrt{h_{s+1}+RV_{s+1}}\right)^2}|\mathcal{F}_s\right]\\
&=\mathrm{e}^{zr+z\lambda h_{s+1}+b_1\omega+b_1\beta_hh_{s+1}-\frac{1}{2}\ln(1-2(b_1\alpha_h+c))+\left(b_1\alpha_h\gamma_h^2+c\gamma^2+2\frac{(b_1\alpha_h\gamma_h+c\gamma-0.5z)^2}{1-2(b_1\alpha_h+c)}\right)h_{s+1}}\\
&\times\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{\left(b_1\alpha_h\gamma_h^2+c\gamma^2+2\frac{(b_1\alpha_h\gamma_h+c\gamma-0.5z)^2}{1-2(b_1\alpha_h+c)}+z\lambda+b_2\right)RV_{s+1}}|\mathcal{F}_s\right]
\end{aligned}
\tag{A.4.4}
$$

Using eq.s (8)-(9) from Gourieroux and Jasiak (2006) we obtain

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{zy_{s+1}+b_1h_{s+2}+b_2RV_{s+1}+c\ell_{s+1}}|\mathcal{F}_s\right]\\
&=\exp\left[\begin{aligned}&zr+b_1\omega-\frac{1}{2}\ln\left(1-2(b_1\alpha_h+c)\right)+x_1h_{s+1}\\&-\delta\mathcal{W}(x_2,\theta)+\mathcal{V}(x_2,\theta)\left(d+\sum_{i=1}^{p}\beta_i\mathrm{RV}_{s-i}+\sum_{j=1}^{q}\alpha_j\ell_{s-j}\right)\end{aligned}\right],
\end{aligned}
\tag{A.4.5}
$$

where

$$
\mathcal{V}(x,\theta)=\frac{\theta x}{1-\theta x}\,,\qquad \mathcal{W}(x,\theta)=\ln\left(1-x\theta\right)\,,
$$

and

$$
\begin{aligned}
x^h(z,\mathbf{b},c)&=z\lambda+b_1\beta_h+2\frac{(b_1\alpha_h\gamma_h+c\gamma-0.5z)^2}{1-2(b_1\alpha_h+c)}+b_1\alpha_h\gamma_h^2+c\gamma^2\\
x^r(z,\mathbf{b},c)&=z\lambda+b_2+2\frac{(b_1\alpha_h\gamma_h+c\gamma-0.5z)^2}{1-2(b_1\alpha_h+c)}+b_1\alpha_h\gamma_h^2+c\gamma^2
\end{aligned}
$$

From direct inspection of the relation (2.1.6), we conclude that

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{zy_{s+1}+b_1h_{s+2}+b_2RV_{s+1}+c\ell_{s+1}}|\mathcal{F}_s\right] \\
&=\exp\left(\mathcal{A}(z,\mathbf{b},c)+\mathcal{B}_1^{(1)}(z,\mathbf{b},c)h_{s+1}+\sum_{i=1}^{22}\mathcal{B}_i^{(2)}(z,\mathbf{b},c)RV_{s+1-i}+\sum_{i=1}^{22}\mathcal{C}_j(z,\mathbf{b},c)\ell_{s+1-j}\right)
\end{aligned}
\tag{A.4.6}
$$

with

$$
\begin{aligned}
\mathcal{A}(z,\mathbf{b},c) &= zr+b_1\omega-\frac{1}{2}\ln\left(1-2(b_1\alpha+c)\right)-\delta\mathcal{W}(x^r,\theta)+d\mathcal{V}(x^r,\theta)\,, \\
\mathcal{B}_1^{(1)}(z,\mathbf{b},c) &= \mathrm{x}^h(z,\mathbf{b},c)\,, \\
\mathcal{B}_i^{(2)}(z,\mathbf{b},c) &= \mathcal{V}(\mathrm{x}^r(z,\mathbf{b},c),\theta)\beta_i\,, \\
\mathcal{C}_j(z,\mathbf{b},c) &= \mathcal{V}(\mathrm{x}^r(z,\mathbf{b},c),\theta)\alpha_j\,.
\end{aligned}
\tag{A.4.7}
$$

Finally, plugging the above expressions for $\mathcal{A}$, $\mathcal{B}_i$ and $\mathcal{C}_j$ with $\nu_1=0$ in eq. (A.1.2) we readily obtain the MGF under the physical measure:

$$
\varphi^{\mathbb{P}}(t,T,z)=\exp\left(\mathrm{a}_t+\mathrm{b}_t^h h_{t+1}+\sum_{i=1}^{22}\mathrm{b}_{t,i}^r\mathrm{RV}_{t+1-i}+\sum_{j=1}^{22}\mathrm{c}_{t,j}\ell_{t+1-j}\right)\,,
$$

where a_t, b_t^h, $\mathrm{b}_{t,i}^r$, $\mathrm{c}_{t,j}$ are given by recursive relations and $\ell_t=\left(\epsilon_t-\gamma\sqrt{\mathrm{RV}_t+h_t}\right)^2$.

$$
\begin{aligned}
\mathrm{a}_s =& \mathrm{a}_{s+1}+zr+\mathrm{b}_{s+1}^h\omega-\frac{1}{2}\ln\left(1-2(\mathrm{b}_{s+1}^h\alpha_h+\mathrm{c}_{s+1})\right) \\
&-\delta\mathcal{W}(\mathrm{x}_{s+1}^r,\theta)+d\mathcal{V}(\mathrm{x}_{s+1},\theta) \\
\mathrm{b}_s^h =& \mathrm{x}_{s+1}^{h*} \\
\mathrm{b}_{s,i}^r =& \begin{cases}\mathrm{b}_{s+1,i+1}^r+\mathcal{V}(\mathrm{x}_{s+1}^r,\theta)\beta_i & \text{for } 1\le i\le 21\\ \mathcal{V}(\mathrm{x}_{s+1}^r,\theta)\beta_i & \text{for } i=22\end{cases} \\
\mathrm{c}_{s,i} =& \begin{cases}\mathrm{c}_{s+1,i+1}+\mathcal{V}(\mathrm{x}_{s+1}^{r*},\theta)\alpha_i & \text{for } 1\le i\le 21\\ \mathcal{V}(\mathrm{x}_{s+1}^r,\theta)\alpha_i & \text{for } i=22\,,\end{cases}
\end{aligned}
\tag{A.4.8}
$$

with

$$
\begin{aligned}
\mathrm{x}^h_{s+1} &= z\lambda + (\beta_h + \alpha_h\gamma_h^2)\mathrm{b}^h_{s+1} + 2\frac{(\mathrm{b}^h_{s+1}\alpha\gamma_h + \mathrm{c}_{s+1}\gamma - 0.5z)^2}{1-2(\mathrm{b}^h_{s+1}\alpha + \mathrm{c}_{s+1})} + \mathrm{c}_{s+1}\gamma^2,\\
\mathrm{x}^r_{s+1} &= z\lambda + \mathrm{b}^r_{s+1} + \alpha_h\gamma_h^2\mathrm{b}^h_{s+1} + 2\frac{(\mathrm{b}^h_{s+1}\alpha_h\gamma_h + \mathrm{c}_{s+1}\gamma - 0.5z)^2}{1-2(\mathrm{b}^h_{s+1}\alpha + \mathrm{c}_{s+1})} + c_{s+1}\gamma^2,
\end{aligned}
$$

A.4.2 Moment Generating Function under risk-neutral measure

Under the risk-neutral measure $\mathbb{Q}$ the MGF for LHARG has the form

$$
\varphi^{\mathbb{Q}}(t,T,z) = \exp\left(\mathrm{a}^*_t + \mathrm{b}^{h*}_t h_{t+1} + \sum_{i=1}^{p}\mathrm{b}^{r*}_{t,i}\mathrm{RV}_{t+1-i} + \sum_{j=1}^{q}\mathrm{c}^*_{t,j}\ell_{t+1-j}\right),
$$

where

$$
\begin{aligned}
\mathrm{a}^*_s =& \mathrm{a}^*_{s+1} + zr + \mathrm{b}^{h*}_{s+1}\omega - \frac{1}{2}\ln\left(1 - 2((\mathrm{b}^{h*}_{s+1} - \nu_h)\alpha_h + \mathrm{c}^*_{s+1})\right) + \frac{1}{2}\ln\left(1 + 2\nu_h\alpha_h\right)\\
& - \delta\mathcal{W}(\mathrm{x}^{r*}_{s+1},\theta) + \delta\mathcal{W}(\mathrm{y}^{r*}_{s+1},\theta) + d\mathcal{V}(\mathrm{x}^{r*}_{s+1},\theta) - d\mathcal{V}(\mathrm{y}^{r*}_{s+1},\theta)\\
\mathrm{b}^{h*}_s =& \mathrm{x}^{h*}_{s+1} - \mathrm{y}^{h*}_{s+1}\\
\mathrm{b}^{r*}_{s,i} =& \begin{cases} \mathrm{b}^{r*}_{s+1,i+1} + \left(\mathcal{V}(\mathrm{x}^{r*}_{s+1},\theta) - \mathcal{V}(\mathrm{y}^{r*}_{s+1},\theta)\right)\beta_i & \text{for } 1 \le i \le p-1\\ \left(\mathcal{V}(\mathrm{x}^{r*}_{s+1},\theta) - \mathcal{V}(\mathrm{y}^{r*}_{s+1},\theta)\right)\beta_i & \text{for } i = p \end{cases}\\
\mathrm{c}^*_{s,i} =& \begin{cases} \mathrm{c}^*_{s+1,i+1} + \left(\mathcal{V}(\mathrm{x}^{r*}_{s+1},\theta) - \mathcal{V}(\mathrm{y}^{r*}_{s+1},\theta)\right)\alpha_i & \text{for } 1 \le i \le q-1\\ \left(\mathcal{V}(\mathrm{x}^{r*}_{s+1},\theta) - \mathcal{V}(\mathrm{y}^{r*}_{s+1},\theta)\right)\alpha_i & \text{for } i = q\,, \end{cases}
\end{aligned}
\tag{A.4.9}
$$

with

$$
\begin{aligned}
\mathrm{x}^{h*}_{s+1} &= (z-\nu_y)\lambda + (\alpha_h\gamma_h^2 + \beta_h)(\mathrm{b}^{h*}_{s+1} - \nu_h) + 2\frac{((\mathrm{b}^{h*}_{s+1} - \nu_h)\alpha_h\gamma^h + c^*_{s+1}\gamma - 0.5(z-\nu_y))^2}{1-2((\mathrm{b}^{h*}_{s+1} - \nu_h)\alpha_h + c^*_{s+1})} + c^*_{s+1}\gamma^2,\\
\mathrm{y}^{h*}_{s+1} &= -\nu_y\lambda - \nu_h(\beta_h + \alpha_h\gamma_h^2) + 2\frac{(-\nu_h\alpha_h\gamma_h + 0.5\nu_y)^2}{1+2\nu_h\alpha_h},\\
\mathrm{x}^{r*}_{s+1} &= (z-\nu_y)\lambda + (\mathrm{b}^{r*}_{s+1} - \nu_r) + \alpha_h\gamma_h^2(\mathrm{b}^{h*}_{s+1} - \nu_h) + 2\frac{((\mathrm{b}^{h*}_{s+1} - \nu_h)\alpha_h\gamma^h + c^*_{s+1}\gamma - 0.5(z-\nu_y))^2}{1-2((\mathrm{b}^{h*}_{s+1} - \nu_h)\alpha_h + c^*_{s+1})} + c^*_{s+1}\gamma^2,\\
\mathrm{y}^{r*}_{s+1} &= -\nu_y\lambda - \nu_h\alpha_h\gamma_h^2 - \nu_r + 2\frac{(-\nu_h\alpha_h\gamma_h + 0.5\nu_y)^2}{1+2\nu_h\alpha_h},
\end{aligned}
$$

and terminal conditions $\mathrm{a}^*_T = \mathrm{b}^{h*}_T = \mathrm{b}^{r*}_{T,i} = \mathrm{c}^*_{T,j} = 0$ for $i = 1,\ldots,p$ and $j = 1,\ldots,q$.

A.5 JLHARG-RV

A.5.1 MGF computations under $\mathbb{P}$ measure

Similarly to GARCH-LHARG process we begin by fiting the leverage and hetergoneous structure into our general framework. The expression (3.4.4) is rewritten as

$$\Theta(\mathbf{RV^c}_t, \mathbf{L}_t) = d + \sum_{i=1}^{22} \beta_i \mathrm{RV}^c_{t+1-i} + \sum_{j=1}^{22} \alpha_j \left(\epsilon_{t+1-j} - \gamma\sqrt{\mathrm{RV}_{t+1-j}}\right)^2 \tag{A.5.1}$$

with

$$\beta_i = \begin{cases} \beta_d - \alpha_d\gamma^2 & \text{for } i = 1 \\ (\beta_w - \alpha_w\gamma^2)/4 & \text{for } 2 \le i \le 5 \\ (\beta_m - \alpha_m\gamma^2)/17 & \text{for } 6 \le i \le 22 \end{cases} \qquad \alpha_i = \begin{cases} \alpha_d & \text{for } i = 1 \\ \alpha_w/4 & \text{for } 2 \le i \le 5 \\ \alpha_m/17 & \text{for } 6 \le i \le 22 \end{cases}, \tag{A.5.2}$$

where $d = -(\alpha_d + \alpha_w + \alpha_m)$.

We begin by showing that JLHARG process satisfies an affine property, namely the following relation holds true

$$\mathbb{E}\left[e^{zy_{s+1}+\mathbf{b}\cdot\mathbf{rv}_{s+1}+\mathbf{c}\cdot\ell_{s+1}}|\mathcal{F}_s\right] = e^{\mathcal{A}(z,\mathbf{b},\mathbf{c})+\sum_{i=1}^{p}\mathcal{B}_i(z,\mathbf{b},\mathbf{c})\cdot\mathbf{rv}_{s+1-i}+\sum_{j=1}^{q}\mathcal{C}_i(z,\mathbf{b},\mathbf{c})\cdot\ell_{s+1-j}}, \tag{A.5.3}$$

for some functions $\mathcal{A} : \mathbb{R}\times\mathbb{R}^k\times\mathbb{R}^k \to \mathbb{R}$, $\mathcal{B}_i : \mathbb{R}\times\mathbb{R}^k\times\mathbb{R}^k \to \mathbb{R}^k$, $\mathcal{C}_j : \mathbb{R}\times\mathbb{R}^k\times\mathbb{R}^k \to \mathbb{R}^k$, where $\mathrm{RV}_t = (\mathrm{RV}^c_t, \mathrm{RV}^j_t)$ and $\mathbf{b}, \mathbf{c} \in \mathbb{R}^k$ and $\cdot$ is the scalar product in $\mathbb{R}^k$. We derive the form of the functions $\mathcal{A}$, $\mathcal{B}_i$, $\mathcal{C}_j$ which allow to compute the MGF for JLHARG.

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[e^{zy_t+\mathbf{b}\cdot\mathbf{rv}_t+c\ell_t}|\mathcal{F}_{t-1}\right]\\
&=\mathbb{E}^{\mathbb{P}}\left[e^{z(r+\lambda\mathrm{RV}_t+\sqrt{\mathrm{RV}_t}\epsilon_t)+\mathbf{b}\cdot\mathbf{rv}_t+c\ell_t}|\mathcal{F}_{t-1}\right]\\
&=\mathbb{E}^{\mathbb{P}}\left[e^{z(r+\lambda\mathrm{RV}_t)+\mathbf{b}\cdot\mathbf{rv}_t}\mathbb{E}^{\mathbb{P}}\left[e^{z\sqrt{\mathrm{RV}_t}\epsilon_t+c(\epsilon_t-\gamma\sqrt{\mathrm{RV}_t})^2}|\mathrm{RV}_t\right]|\mathcal{F}_{t-1}\right]\\
&=\mathbb{E}^{\mathbb{P}}\left[e^{z(r+\lambda\mathrm{RV}_t)+b_1\mathrm{RV}_t^c+b_2\mathrm{RV}_t^j-\frac{1}{2}\ln(1-2c)+\left(\frac{\frac{z^2}{2}+\gamma^2c-2c\gamma z}{1-2c}\right)\mathrm{RV}_t}|\mathcal{F}_{t-1}\right]\\
&=\mathbb{E}^{\mathbb{P}}\left[e^{zr-\frac{1}{2}\ln(1-2c)+\left(z\lambda+b_1+\frac{\frac{z^2}{2}+\gamma^2c-2c\gamma z}{1-2c}\right)\mathrm{RV}_t^c+\left(z\lambda+b_2+\frac{\frac{z^2}{2}+\gamma^2c-2c\gamma z}{1-2c}\right)\mathrm{RV}_t^j}|\mathcal{F}_{t-1}\right]\\
&=e^{zr-\frac{1}{2}\ln(1-2c)}\mathbb{E}^{\mathbb{P}}\left[e^{\left(z\lambda+b_1+\frac{\frac{z^2}{2}+\gamma^2c-2c\gamma z}{1-2c}\right)\mathrm{RV}_t^c}|\mathcal{F}_{t-1}\right]\mathbb{E}^{\mathbb{P}}\left[e^{\left(z\lambda+b_2+\frac{\frac{z^2}{2}+\gamma^2c-2c\gamma z}{1-2c}\right)\mathrm{RV}_t^j}|\mathcal{F}_{t-1}\right].
\end{aligned}
\tag{A.5.4}
$$

In the third passage we use property (A.2.2). From Gourieroux and Jasiak (2006) for a non-centred gamma distributed random variable, we obtain

$$
\mathbb{E}^{\mathbb{P}}\left[e^{x\mathrm{RV}_t^c}|\mathcal{F}_{t-1}\right]=\exp\left(\delta\mathcal{W}\left(x,\theta\right)+\mathcal{V}\left(x,\theta\right)\left(d+\sum_{i=1}^{p}\beta_i\mathrm{RV}_{s-i}^c+\sum_{j=1}^{q}\alpha_j\ell_{s-j}\right)\right),
\tag{A.5.5}
$$

where

$$
\mathcal{V}(x,\theta)=\frac{\theta x}{1-\theta x},\quad \mathcal{W}(x,\theta)=\ln(1-x\theta),
\tag{A.5.6}
$$

and

$$
x(z,b,c)=z\lambda+b_1+\frac{\frac{1}{2}z^2+\gamma^2c-2c\gamma z}{1-2c}.
\tag{A.5.7}
$$

For the computation of the last expectation in the last line of expression (A.5.4), we use the property that if Z_t is a compound Poisson process with rate ω and sizes D_i i.i.d. then

$$
\mathbb{E}\left[e^{xZ_t}|\mathcal{F}_{t-1}\right]=\exp\left(\omega\left(M_D(x)-1\right)\right),
\tag{A.5.8}
$$

where $M_D(x)$ is the moment-generating function of the random variable D of the jump size. Since the sizes of the jumps in realized volatility are distributed according to a gamma distribution, we have that

$$M_D(x) = \frac{1}{\left(1 - x\tilde{\theta}\right)^{\tilde{\delta}}}. \tag{A.5.9}$$

According to expressions (A.5.8) and (A.5.9), we obtain

$$\mathbb{E}^{\mathbb{P}}\left[e^{\left(z\lambda + b_2 + \frac{\frac{z^2}{2} + \gamma^2 c - 2c\gamma z}{1-2c}\right)\mathrm{RV}_t^j}|\mathcal{F}_{t-1}\right] = \exp\left(\tilde{\Theta}\mathcal{J}\left(x, \tilde{\theta}, \tilde{\delta}\right)\right), \tag{A.5.10}$$

where

$$\mathcal{J}(x, \tilde{\theta}, \tilde{\delta}) = \frac{1 - (1 - \tilde{\theta}x)^{\tilde{\delta}}}{(1 - \tilde{\theta}x)^{\tilde{\delta}}}. \tag{A.5.11}$$

Gathering all the previous results, we finally have

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[e^{zy_t + \mathbf{b}\cdot\mathbf{RV}_t + c\ell_t}|\mathcal{F}_{t-1}\right] = \\
&\exp\left[zr - \frac{1}{2}\ln(1-2c) + \mathcal{V}(x_1, \theta)\left(d + \sum_{i=1}^{p}\beta_i \mathrm{RV}_{t-i}^c + \sum_{j=1}^{q}\alpha_j \ell_{t-j}\right)\right. \\
&\left. -\delta\mathcal{W}(x_1, \theta) + \tilde{\Theta}\mathcal{J}(x_2, \tilde{\theta}, \tilde{\delta})\right]
\end{aligned} \tag{A.5.12}$$

where we distinguish the two functions $x_1 = x(z, b_1, c)$ and $x_2 = x(z, b_2, c)$ and the expression for x is given by (A.5.7). The direct comparison of the last expression with A.5.3 allow to derive explicitly the form of the functions of the exponential affine form:

$$\begin{aligned}
\mathcal{A}(z, \mathbf{b}, c) &= zr - \frac{1}{2}\ln(1-2c) - \delta\mathcal{W}(x_1, \theta) + d\mathcal{V}(x_1, \theta) + \tilde{\Theta}\mathcal{J}(x_2, \tilde{\theta}, \tilde{\delta}) \\
\mathcal{B}_i(z, b_1, c) &= \mathcal{V}(x_1, \theta)\beta_i \\
\mathcal{C}_j(z, b_1, c) &= \mathcal{V}(x_1, \theta)\alpha_j.
\end{aligned} \tag{A.5.13}$$

Finally, plugging the above expressions for $\mathcal{A}$, $\mathcal{B}_i$ and $\mathcal{C}_j$ in eq. (A.1.2) with SDF's parametrs equal zero we readily obtain the MGF under the physical measure:

$$\varphi^{\mathbb{P}}(t, T, z) = \mathbb{E}^{\mathbb{P}}\left[e^{zy_{t,T}}|\mathcal{F}_t\right] = \exp\left(\mathrm{a}_t + \sum_{i=1}^{p}\mathrm{b}_{t,i}\mathrm{RV}_{t+1-i}^c + \sum_{i=1}^{q}\mathrm{c}_{t,i}\ell_{t+1-i}\right) \tag{A.5.14}$$

where

$$
\begin{aligned}
\mathrm{a}_s &= \mathrm{a}_{s+1} + zr - \frac{1}{2}\log(1 - 2\mathrm{c}_{s+1,1}) + d\mathcal{V}(\mathrm{x}^c_{s+1}, \theta) - \delta\mathcal{W}(\mathrm{x}^c_{s+1}, \theta) + \tilde{\Theta}\mathcal{J}(\mathrm{x}^j_{s+1}, \tilde{\theta}) \\
\mathrm{b}_{s,i} &= \begin{cases} \mathrm{b}_{s+1,i} + \mathcal{V}(\mathrm{x}^c_{s+1}, \theta)\beta_i & for\ 1 \leq i \leq p-1 \\ \mathcal{V}(\mathrm{x}^c_{s+1}, \theta)\beta_i & for\ i = p \end{cases} \\
\mathrm{c}_{s,i} &= \begin{cases} \mathrm{c}_{s+1,i} + \mathcal{V}(\mathrm{x}^c_{s+1}, \theta)\alpha_i & for\ 1 \leq i \leq q-1 \\ \mathcal{V}(\mathrm{x}^c_{s+1}, \theta)\alpha_i & for\ i = q \end{cases}
\end{aligned}
\tag{A.5.15}
$$

where

$$
\mathrm{x}^c_{s+1} = z\lambda + \mathrm{b}_{s+1,1} + \frac{\frac{1}{2}z^2 + \gamma^2\mathrm{c}_{s+1,1} - 2\mathrm{c}_{s+1,1}\gamma z}{1 - 2\mathrm{c}_{s+1,1}} \tag{A.5.16}
$$

$$
\mathrm{x}^j_{s+1} = z\lambda + \frac{\frac{1}{2}z^2 + \gamma^2\mathrm{c}_{s+1,1} - 2\mathrm{c}_{s+1,1}\gamma z}{1 - 2\mathrm{c}_{s+1,1}} \tag{A.5.17}
$$

The functions $\mathcal{V}$, $\mathcal{W}$ and $\mathcal{J}$ are defined as

$$
\mathcal{V}(x, \theta) = \frac{\theta x}{1 - \theta x}, \quad \mathcal{W}(x, \theta) = \ln(1 - x\theta), \quad \mathcal{J}(x, \tilde{\theta}, \tilde{\delta}) = \frac{1 - (1 - \tilde{\theta}x)^{\tilde{\delta}}}{(1 - \tilde{\theta}x)^{\tilde{\delta}}} \tag{A.5.18}
$$

and the terminal condition are $\mathrm{a}_T = \mathrm{b}_{T,i} = \mathrm{c}_{T,j} = 0$ for $i = 1, 2, ..., p$ and $j = 1, 2, ..., q$.

A.5.2 Risk-neutral dynamics

Firstly we observe that risk-neutral MGF can be expressed with a recursive set of expressions, involving a combination of the functions $\mathcal{A}$, $\mathcal{B}_i$, $\mathcal{C}_j$. From Theorem 9, the MGF for JLHARG model under measure $\mathbb{Q}$ has the following form

$$
\varphi^{\mathbb{Q}}_{\nu_r,\nu_j,\nu_y}(t, T, z) = \mathbb{E}^{\mathbb{Q}}\left[e^{zy_{t,T}} | \mathcal{F}_t\right] = \exp\left(\mathrm{a}^*_t + \sum_{i=1}^{p} \mathrm{b}^*_{t,i}\mathrm{RV}^c_{t+1-i} + \sum_{i=1}^{q} \mathrm{c}^*_{t,i}\ell_{t+1-i}\right),
$$

where

$$\begin{aligned}
\mathrm{a}_s^* &= \mathrm{a}_{s+1}^* + zr - \frac{1}{2}\log(1-2\mathrm{c}_{s+1,1}^*) + d\mathcal{V}(\mathrm{x}_{s+1}^{c\,*},\theta) - d\mathcal{V}(\mathrm{y}_{s+1}^{c\,*},\theta) \\
&- \delta\mathcal{W}(\mathrm{x}_{s+1}^{c\,*},\theta) + \delta\mathcal{W}(\mathrm{y}_{s+1}^{c\,*},\theta) + \tilde{\Theta}\mathcal{J}(\mathrm{x}_{s+1}^{j\,*},\tilde{\theta}) - \tilde{\Theta}\mathcal{J}(\mathrm{y}_{s+1}^{j\,*},\tilde{\theta}) \\
\mathrm{b}_{s,i}^* &= \begin{cases} \mathrm{b}_{s+1,i}^* + \left(\mathcal{V}(\mathrm{x}_{s+1}^{c\,*},\theta) - \mathcal{V}(\mathrm{y}_{s+1}^{c\,*},\theta)\right)\beta_i & for\ 1 \le i \le p-1 \\ \left(\mathcal{V}(\mathrm{x}_{s+1}^{c\,*},\theta) - \mathcal{V}(\mathrm{y}_{s+1}^{c\,*},\theta)\right)\beta_i & for\ i = p \end{cases} \\
\mathrm{c}_{s,j}^* &= \begin{cases} \mathrm{c}_{s+1,j}^* + \left(\mathcal{V}(\mathrm{x}_{s+1}^{c\,*},\theta) - \mathcal{V}(\mathrm{y}_{s+1}^{c\,*},\theta)\right)\alpha_j & for\ 1 \le j \le q-1 \\ \left(\mathcal{V}(\mathrm{x}_{s+1}^{c\,*},\theta) - \mathcal{V}(\mathrm{y}_{s+1}^{c\,*},\theta)\right)\alpha_j & for\ j = q \end{cases}
\end{aligned} \tag{A.5.19}$$

where

$$\begin{aligned}
x_{s+1}^{c\,*} &= (z-\nu_y)\lambda + \mathrm{b}_{s+1,1}^* - \nu_c + \frac{\frac{1}{2}(z-\nu_y)^2 + \gamma^2\mathrm{c}_{s+1,1}^* - 2\mathrm{c}_{s+1,1}^*\gamma(z-\nu_y)}{1-2c_{s+1,1}^*} \\
x_{s+1}^{j\,*} &= (z-\nu_y)\lambda - \nu_j + \frac{\frac{1}{2}(z-\nu_y)^2 + \gamma^2\mathrm{c}_{s+1,1}^* - 2\mathrm{c}_{s+1,1}^*\gamma(z-\nu_y)}{1-2c_{s+1,1}^*} \\
y_{s+1}^{l\,*} &= -\nu_y\lambda - \nu_l + \frac{1}{2}\nu_y^2,
\end{aligned}$$

with $l = r, j$ and the terminal conditions are $\mathrm{a}_T^* = \mathrm{b}_{T,i}^* = \mathrm{c}_{T,j}^* = 0$ for $i = 1, 2, ..., p$ and $j = 1, 2, ..., q$.

The first passage consists in comparing expression (A.5.19) with (A.5.15). We have to find a set of new parameters for which the recursive expressions for $\mathrm{a}_t^*, \mathrm{b}_t^*, \mathrm{c}_t^*$ under $\mathbb{Q}$ correspond to the expressions under $\mathbb{P}$. We start defining

$$\begin{aligned}
x_{s+1,i}^{c\,**} &= z\lambda^* + b_{s+1,1}^* + \frac{\frac{1}{2}z^2 + (\gamma^*)^2 c_{s+1,1}^* - 2c_{s+1,1}^*\gamma^* z}{1-2c_{s+1,1}^*}, \\
x_{s+1,i}^{j\,**} &= z\lambda^* + \frac{\frac{1}{2}z^2 + (\gamma^*)^2 c_{s+1,1}^* - 2c_{s+1,1}^*\gamma^* z}{1-2c_{s+1,1}^*}.
\end{aligned}$$

Then, the following relations have to hold

$$\delta\left(\mathcal{W}\left(\mathrm{x}_{s+1}^{c\,*},\theta\right)-\mathcal{W}\left(\mathrm{y}^{c\,*},\theta\right)\right)=\delta^{*}\mathcal{W}\left(\mathrm{x}_{s+1}^{c\,**},\theta^{*}\right) \tag{A.5.20}$$

$$\beta_i\left(\mathcal{V}\left(\mathrm{x}_{s+1}^{c\,*},\theta\right)-\mathcal{V}\left(\mathrm{y}^{c\,*},\theta\right)\right)=\beta_i^{*}\mathcal{V}\left(\mathrm{x}_{s+1}^{c\,**},\theta^{*}\right) \tag{A.5.21}$$

$$\alpha_j\left(\mathcal{V}\left(\mathrm{x}_{s+1}^{c\,*},\theta\right)-\mathcal{V}\left(\mathrm{y}^{c\,*},\theta\right)\right)=\alpha_j^{*}\mathcal{V}\left(\mathrm{x}_{s+1}^{c\,**},\theta^{*}\right) \tag{A.5.22}$$

$$\tilde{\Theta}\left(\mathcal{J}\left(\mathrm{x}_{s+1}^{j\,*},\tilde{\theta}\right)-\mathcal{J}\left(\mathrm{y}^{j\,*},\tilde{\theta}\right)\right)=\tilde{\Theta}^{*}\mathcal{J}\left(\mathrm{x}_{s+1}^{j\,**},\tilde{\theta}^{*}\right) \tag{A.5.23}$$

with $\mathrm{y}^{c\,*}=-\lambda^2/2-\nu_r+\frac{1}{8}$ and $\mathrm{y}^{j\,*}=-\lambda^2/2-\nu_j+\frac{1}{8}$.
Equation (A.5.20) can be explicitly written as

$$\delta\log\left[1-\frac{\theta}{1-\theta\mathrm{y}^{c\,*}}\left(\mathrm{x}_{s+1}^{c\,*}-\mathrm{y}^{c\,*}\right)\right]=\delta^{*}\log\left(1-\theta^{*}\mathrm{x}_{s+1}^{c\,**}\right),$$

which implies the following three sufficient conditions

$$\begin{aligned}\delta^{*}&=\delta\\ \theta^{*}&=\frac{\theta}{1-\theta\mathrm{y}^{c\,*}}\\ \mathrm{x}_{s+1}^{c\,**}&=\mathrm{x}_{s+1}^{c\,*}-\mathrm{y}^{c\,*}.\end{aligned} \tag{A.5.24}$$

It can be easily verified that the last condition (A.5.24) is satisfied by substituting

$$\begin{aligned}\lambda^{*}&=-\frac{1}{2},\\ \gamma^{*}&=\gamma+\lambda+\frac{1}{2}.\end{aligned}$$

The equation (A.5.21) can be equivalently expressed in the form

$$\frac{\beta_i}{1-\theta\mathrm{y}^{c\,*}}\frac{\theta}{1-\theta\mathrm{y}^{c\,*}}\frac{\mathrm{x}_{s+1}^{c\,*}-\mathrm{y}^{c\,*}}{1-\theta/(1-\theta\mathrm{y}^{c\,*})\left(\mathrm{x}_{s+1}^{c\,*}-\mathrm{y}^{c\,*}\right)}=\beta_i^{*}\frac{\theta^{*}\mathrm{x}_{s+1}^{c\,**}}{1-\theta^{*}\mathrm{x}_{s+1}^{c\,**}}$$

which gives another sufficient condition for the mapping

$$\beta_i^{*}=\frac{\beta_i}{1-\theta\mathrm{y}^{c\,*}}.$$

An analogous consideration about the third condition (A.5.22) allows to obtain the condition on α_i^*,

$$\alpha_i^* = \frac{\alpha_i}{1-\theta \mathrm{y}^{c\,*}} \, .$$

Relation (A.5.2) gives us the expressions for β_d^*, β_w^*, β_m^*, α_d^*, α_w^* and α_m^*. Finally, equation (A.5.23) provides the last sufficient condition

$$\frac{\tilde{\Theta}}{\left(1-\tilde{\theta}\mathrm{y}^{j\,*}\right)^{\tilde{\delta}}} \frac{1-\left(\left(1-\tilde{\theta}\mathrm{x}_{s+1}^{j\,*}\right) / \left(1-\tilde{\theta}\mathrm{y}^{j\,*}\right)\right)^{\tilde{\delta}}}{\left(\left(1-\tilde{\theta}\mathrm{x}_{s+1}^{j\,*}\right) / \left(1-\tilde{\theta}\mathrm{y}^{j\,*}\right)\right)^{\tilde{\delta}}} = \tilde{\Theta}^* \frac{1-(1-\tilde{\theta}^*\mathrm{x}_{s+1}^{j\,**})^{\tilde{\delta}^*}}{(1-\tilde{\theta}^*\mathrm{x}_{s+1}^{j\,**})^{\tilde{\delta}^*}},$$

which is satisfied if

$$\begin{aligned}
\tilde{\delta}^* &= \tilde{\delta}\,, \\
\tilde{\Theta}^* &= \frac{\tilde{\Theta}}{\left(1-\tilde{\theta}\mathrm{y}^{j\,*}\right)^{\tilde{\delta}}}\,, \\
\tilde{\theta}^* &= \frac{\tilde{\theta}}{1-\tilde{\theta}\mathrm{y}^{j\,*}}\,, \\
\mathrm{x}_{s+1}^{j\,**} &= \mathrm{x}_{s+1}^{j\,*} - \mathrm{y}^{j\,*}\,.
\end{aligned} \tag{A.5.25}$$

As it can be seen the last condition (A.5.25) is redundant when compared to the condition (A.5.24).