Scuola Normale Superiore

Classe di Scienze Matematiche e Naturali

Discrete Time Models for Financial Volatility and Jumps

Tesi di Perfezionamento in Matematica per la Finanza

Pisa 2017

CANDIDATO

Dario Alitab

RELATORI

Prof. Giacomo Bormetti

Prof. Fulvio Corsi

Acknowledgements

I would like to express my deepest gratitude to my advisors. I am grateful to Prof. Giacomo Bormetti for the exceptional guidance of my research path as his Ph.D. student. His open-mindedness, dedication and scientific rigour have been a source of inspiration and have given to me the opportunity to gain many skills and knowledge.
I warmly thank Prof. Fulvio Corsi for having shared his experience, ideas and all the productive discussions we had during these years. His contribution to our research work has been immensely precious.
To both of them, I want to say thanks for all the time they dedicated to me during the days spent in Pisa and, moreover, for all the extra-time they continued to devote until the submission of this thesis. This would not be possible without their invaluable support, advice and commitment.

I wish to sincerely thank Prof. Stefano Marmi, director of the Ph.D. programme. He gave me the opportunity to live this experience in a wonderful and remarkable place. As student, I appreciated the enthusiasm and the originality he put into his lessons and seminars. Furthermore, I thank him for funding to make my Ph.D studies productive.

I am thankful to Prof. Fabrizio Lillo for his beautiful courses and seminars. His research was an inspiration to me when I firstly got interested in quantitative finance. I also would like to thank him for the time, ideas and advices he gave me before starting the Ph.D. and after its conclusion.

I would also thank my dear colleague and co-author of our papers Adam Alek-

sander Majewski for his collaboration, helpful comments and fruitful discussions. I thank Giulia Livieri for her help, sharing ideas and important advices. Furthermore, I thank Damian Eduardo Taranto for all the interesting discussions about our research topics and financial mathematics.

Additionally, I want to thank the entire staff of the Scuola Normale Superiore for their hard work and dedication in every complementary aspect of our academic life as students.

Beyond being a significant period of study and cultural growth, during these years I had the fortune to meet a wonderful group of people. I would want to say thanks to each of them for all the amazing moments we spent together and for their sincere enduring friendship. We had a very good time.

Finally, my family. Thanks to my mother, thanks to my father and thanks to my brother for their support, their continuous encouragement and, above all, for their love. They were there in my times of trouble, no matter the distance, they were always near to me. This has meant a lot for me. I am so grateful for having had them by my side during this long extraordinary experience.

List of papers

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. (2016)).

Dario Alitab, Giacomo Bormetti, Fulvio Corsi and Adam A. Majewski, *Option Pricing with Realized Volatility and Persistent Jump Intensity*, (Alitab et al. (2017)).

Contents

Introduction

Volatility is one of the most recurrent keyword in finance. In its simplest form it is intended to be a measure of the degree of uncertainty governing the dynamics of the price of a stock or a commodity, an interest rate or a foreign exchange rate. We can think of it as the average deviation from the mean value of a set of data related to a studied quantity or what statisticians call *standard deviation*. For example by recording the daily returns of different stocks for a given period of time we could observe that some of them are more erratic than others, or in a different word, more "volatile". This concept has probably its origin from the very first work of financial mathematics, the thesis of the French mathematician, Louis Bachelier who postulated the existence of a "coefficient of instability" of the price of a financial security, giving a first example of Brownian motion model applied in a financial context (Bachelier (1900)).

Since then, the common view about the dynamics of a financial asset is given as a sum of a drift term indicating the "direction" of the temporal average behaviour and a diffusive term tuned by volatility accounting for the fluctuations coming from the arrival of new information. Nevertheless, this view has been proven to be incomplete since more rarely a security price can assume a value far beyond the local mean deviation. Sometimes it exhibits a "jump".

The concept of jump is introduced to describe those events which involve a change of extraordinary magnitude in the value of an asset over a very small time interval compared to the usual fluctuations happening in that time scale. It represents the reaction of market participants to the arrival of important new information whose effect is such to produce a discontinuity in the temporal evolution of the market value. This process is in fact different from the diffusive dynamics since it appears to be active only at random points remaining in an idle state most of

the times.

During the last thirty years, many empirical studies has evidenced that volatility is indeed a quite richer process, far from being considered a static parameter of the motion. It exhibits variability over time, a property know as heteroschedasticity, mean reversion, temporal clustering, long memory and self similarity. It also shows an asymmetric response to the sign of past realized returns, with negative ones having a higher impact on enhancing the level of future volatility with respect to the positive ones. This well-documented fact is known as leverage effect.

After Black and Scholes (1973)'s contribution, the typical modelling approach to describe the dynamics of asset prices has been done in continuous-time. This framework has the main advantage to provide a mathematically consistent and elegant method to obtain no-arbitrage prices for derivative contracts in many cases expressed in closed or semi-closed form. This setting has rapidly become mainstream. Nevertheless, Black-Scholes prices for options systematically differ from quoted ones. Some of the reasons are the unrealistic assumptions of time-constant volatility of the diffusive dynamics of the underlying return and the continuity of its temporal trajectories. To cope with this problem, continuous-time stochastic volatility models with jumps, known as jump-diffusion models, are introduced (among others Bates (1996b), Bakshi et al. (1997), Duffie et al. (2000)). Treating the volatility has a time-varying process and accounting for the presence of extreme movements in prices, these models are able to meet some of the empirical properties of volatility dynamics like clustering and the leverage effect and the significant negative skewness and the excess kurtosis of stock returns, resulting in a more realistic pricing and a better fit of the smile profiles of implied volatilities.

One of the issues of the jump-diffusion setting concerns the difficulties related to estimating continuous-time models with discretely sampled data. The typical solution is to directly model the dynamics under risk-neutral measure and calibrate it in order to best fit derivative's quoted prices. The alternative approach consists in building models in discrete-time. Originally, option pricing has not been one of the reason for the introduction of the discrete-time framework. Indeed, the mo-

tivation has been to propose models able to mimic the well-recognized fact that uncertainty of asset prices, as measured by volatility, changes through time and produce good forecasting of future volatilities, by using the historical information available. Seminal works by Engle (1982) and Bollerslev (1986), introducing the ARCH-GARCH processes, represent the first instances of this idea.

This thesis has the purpose to widen the family of discrete-time processes with different objectives. First, we propose a novel class of models that can reproduce the many stylized facts documented in empirical time series of asset prices, as jumps, volatility persistence and leverage effect. Secondly, despite the presence of different dynamical features, we build these models in order to be easily estimated using historical information. Third, we want this class to be available as a pricing framework capable to explain the well-known characteristics of the volatility surface implied by real option quotes, as the smile or the term structure. Finally, since we are also interested in modelling the connection between the physical process of an asset price and its risk-neutralization, we adopt for our models a flexible multi-dimensional pricing kernel including multiple risk premia for each source of randomness driving the dynamics (as discussed in Gagliardini et al. (2011) and Gourieroux and Monfort (2007)). This choice implies the further advantage that our framework allows to compute European option prices with semi-closed analytical pricing formulas.

The models that we are going to introduce belong to the class of affine processes. For such models we provide the methodology to compute the moment generating function of log-returns, an alternative way to specify the conditional probability distribution. Furthermore, we are able to determine the formal change of measure, coherently satisfying no-arbitrage condition, and derive the moment generating function under risk-neutral measure. Related to this, we prove that the risk-neutral process still belongs to the class of models proposed.

Our proposal consists of stochastic volatility models with jumps based on Realized Volatility (RV). This quantity is an easy-to-compute measure of the asset volatility typically constructed from the intraday price movements. Recent papers have opened a new field of research combining successfully RV literature with that on option pricing (see Stentoft (2008), Christoffersen et al. (2014), Corsi et al.

(2013) and Majewski et al. (2015)). The advantage in using realized volatility is that models which rely on it are easy to be estimated. In fact, since it is built directly from observed returns, this variable is in turn observable allowing the use of maximum likelihood estimation methods. Furthermore, a model based on realized volatility effectively incorporates the information contained in intraday data. This gives the model more flexibility to adapt rapidly to the market's movements and consequently improve option pricing. Nevertheless, most attention has been given to RV models accounting for the diffusive dynamics of asset prices. To the best of our knowledge, little work has been devoted to include a jump component in RV valuation model framework. A notable exception is the paper by Christoffersen et al. (2015) which develops an option valuation model that uses the observable realized volatility and a realized measure for jump variation to describe the dynamics in the diffusive volatility and in the jump intensity.
The main motivation of this work is to investigate the combined role of volatility dynamics and jump events within a novel class of realized volatility models. Including a jump component provides a rapidly moving factor which improves on the fitting properties of historical time series and allows to better explain the pronounced smile profiles of implied volatilities at short time-to-maturities. To cope with the presence of extreme events, we follows two different approaches. In a first class of models, we describe the jump dynamics as a compound Poisson process with constant jump intensity. This represents a parsimonious way to include a pure discontinuous component with the advantage of preserving simplicity of the estimation procedure without involving filtering procedures. Differently, our second model considers a jump component whose intensity is time-varying with autoregressive dynamics where the realized number of jumps detected each day plays the role of an idiosyncratic shock. Modelling a persistent intensity is adequate to reproduce the typical observed phenomenon of the clustering of extreme events. However, adding this new dynamic feature implies an estimation method which has to filter out the latent intensity variable. In both approaches, we model the diffusion component of the asset price as an autoregressive gamma process for the continuous realized volatility accounting for heterogeneous time-scale and leverage effect.

Concerning the intrinsic stochastic nature of volatility, a growing strand of literature is interested on the relevant topic of the variance risk premium. This is commonly intended as the compensation that a representative investor demands for bearing the risk related to unknown future variance of a given investment. Works by Bates (2000), Bakshi and Kapadia (2003) and Carr and Wu (2009) provide quantitative measures of the variance risk premium both in parametric and non-parametric approach. Successive studies have evidenced that the variance risk premium is able to predict stock market returns aggregated over different time horizons (see Bollerslev et al. (2009a) and Bakshi et al. (2010a)). Within our framework, we give explicit formulas to compute a dynamic model-implied measure of variance risk premium. The relevant aspect of this measure is that it can be easily split in order to separate the two contributions coming from the diffusive and the jump dynamics. A major advantage of adopting an analytical parametric measure is the possibility to compute risk premia over different maturity recovering their whole term structure.

The thesis is divided in four chapters. Chapter 1 starts with an overview of the available models proposed in the last decades to describe asset prices and volatility dynamics both in continuous and discrete-time settings. It summarizes the motivation behind the introduction of jump-diffusion processes in an arbitrage-free pricing framework and the advent of the econometric models belonging to the wide ARCH-GARCH class and more recently developed for realized volatility. In the following sections, we introduce the definition of realized volatility pointing out the practical issues encountered in building this measure from high-frequency time series and reporting references to different methods to deal with them. Finally, we concentrate on an important application of realized volatility and higher order realized measures: jump detection. We report some of the main results of a wide literature proposing a variety of statistical test to spot the discontinuities in the price path of a security based on asymptotic theory of a class of robust-to-jumps estimators built from high-frequency returns.

Chapter 2 deals with the topic of the risk-neutralization of discrete-time stochastic processes for asset pricing purposes. The central issue is modelling the stochastic discount factor or pricing kernel allowing to coherently connect the physical or

historical measure to the risk-neutral or martingale measure within an arbitrage-free framework. We give some examples of the typical stochastic discount factors encountered in financial literature focusing on the particular exponential-affine class. As we will see, this family has many advantages. For example, it allows to incorporate multiple risk-premia preserving full analytical tractability in many cases. In particular, the possibility to include different premia, adopting multidimensional pricing kernels, has been recently proven to be interestingly useful to explain well-known pricing puzzles (as discussed in Christoffersen et al. (2013)).
In Chapter 3 we present the novel class of models based on realized volatility accounting for heterogeneity, leverage and jumps. We summarize some of the well-established stylized facts recurring in empirical studies of financial time series and give motivation behind the proposal of heterogeneous autoregressive processes. We formally define the models and report the important results about analytical characterization of the probability distribution by computing the moment generating function. We also show how to determine the change of measure from the physical to the risk-neutral world by introducing an exponential-affine stochastic discount factor and demonstrate the existence of a mapping of model parameters from one measure to the other. Finally, we introduce a definition for the variance risk premium and give a set of recursive formulas to compute a model-implied estimate within our framework.
Chapter 4 provides the results of some empirical studies involving our models and using real data from financial datasets. We describe the procedures to construct our observable time series of returns and realized volatilities and separate the continuous and jump contributions of the dynamics. We provide parameters' estimation results and show how to calibrate the models for option pricing application. Then we assess the pricing performances by benchmarking our proposed class with standard and state-of-the-art models available in literature. We conclude with a final exploration about variance risk premium in order to verify if our model-implied measure of VRP is capable to predict stock market returns. We perform our forecasts through simple linear regressions of the S&P 500 excess returns on series of VRP computed both in a complete parametric way and in a semi-parametric one. As benchmark, we repeat the same regression us-

ing commonly recognized predictor variables, including the Cyclically Adjusted Price-to-Earnings, the Term Spread of interest rates and the market proxy of the variance risk premium determined from VIX index quotations. A comparative discussion of the results concludes.

Chapter 1

Stochastic volatility and jumps from continuous to discrete time

1.1 Continuous-time models

The main assumption behind the Black-Scholes model is that the dynamics of a stock price follows a geometric Brownian motion whose log-return is described by a Gaussian process with constant volatility parameter. Although theoretically important, this claim has been proven to be inadequate to explain the well-known "smile/skew" profile of implied volatilities curves obtained from quoted plain vanilla option prices as functions of their strikes. This fact has led to the development of two classes of models: local volatility and stochastic volatility.
The local volatility model, introduced by Dupire (1994) and Derman and Kani (1994), is a generalization of the Black-Scholes model that treats volatility as a deterministic function of the asset value. Indicating with S_t, the value of the asset at time t, this model implies the following stochastic differential equation (SDE)

$$\frac{dS_t}{S_t} = \mu_t dt + \sigma_L(S_t, t; S_0) dW_t$$

where μ_t is a drift term, W_t is a Wiener process and $\sigma_L(S_t, t; S_0)$ is the local volatility function, which in general can depend on the current asset value S_0. The advantage of this assumption is that it allows to adequately fit the entire

implied volatility surface. However, the main criticism against this approach is that it lacks economic interpretation by not giving a reasonable explanation of the volatility smile.

On the other hand, stochastic volatility models assume that the asset price S_t and its variance v_t are both stochastic processes satisfying the following SDE (see Gatheral (2006))

$$
\begin{aligned}
dS_t &= \mu_t S_t dt + m\sqrt{v_t} S_t dW_t^{(1)} \\
dv_t &= \alpha\left(S_t, v_t, t\right) dt + \eta\beta\left(S_t, v_t, t\right)\sqrt{v_t} dW_t^{(2)} \\
\langle dW_t^{(1)}, dW_t^{(2)} \rangle &= \rho dt
\end{aligned}
$$

where $\alpha : \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}$, $\beta : \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$, η is the volatility of volatility and ρ is the correlation between random asset price returns and changes in v_t. $dW_t^{(1)}$ and $dW_t^{(2)}$ are Wiener processes. These models are useful because they explain in a self-consistent way the volatility smile by assuming a realistic dynamics for the underlying. Remarkable instances of these large class of models are introduced by Hull and White (1987), Stein and Stein (1991), Heston (1993) and Hagan et al. (2002).

The reason behind modelling volatility as a stochastic process comes from empirical studies of stock price returns according to which the estimated volatility exhibits a random evolution through time. One of its effect is thickening the tails of returns distributions if compared to the normal distribution assumed in the Black-Scholes model. This aspect, implying a higher probability of tail events for the stock price dynamics, reconciles with the higher implied volatilities observed for out-of-the-money and in-the-money options with respect to at-the-money volatilities.

Another simplifying assumption introduced by Black and Scholes (1973) is that the price evolves in time according to continuous trajectories. Differently from this, it is empirically observed that daily and intraday asset returns time series exhibit extreme abrupt movements, or jumps, appearing as sharp discontinuities of the apparent continuous temporal evolution. These sudden unpredictable events contribute to fatten the tails of sampled returns distributions and are con-

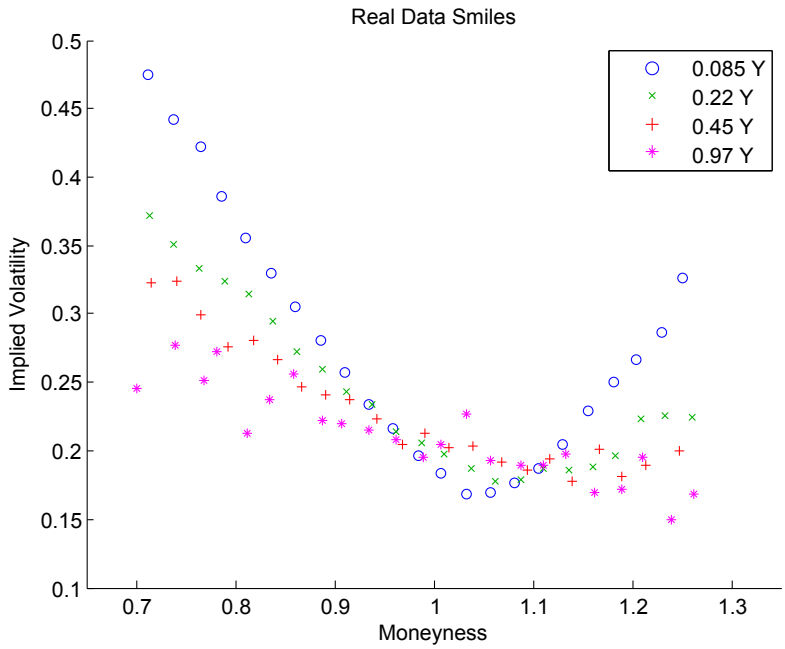

Figure 1.1: Implied volatility profile for options written on S&P 500 index with different maturities. The pronounced smile for short-maturity options is regarded as an evidence for the "jump fear" of the market agents.

sidered a relevant source of risk by market agents. It should be mentioned that heavy tails due to jump in price could be reproduced by a fine-tuned stochastic volatility model but its predictable diffusive dynamics would be at odd with the documented occurrence of discontinuities in sample paths prices.

It is documented that while stochastic volatility models are quite flexible to correctly reproduce the options smile, nevertheless they are often unable to fit the implied skew of short time-to-maturity options. As shown in Figure 1.1, smile profiles are affected by the maturity of the contracts. In particular the U-shape curves are steeper for very short maturity options and they become flatter moving to longer maturities. The common interpretation for this phenomenon is that market participants fear abrupt large movements in prices occurring on short term horizons that could change the value of the contracts from out-of-the-money to in-the-money implying big losses without the chance of "recovery". Coherently, this translates in higher prices for short maturity OTM options and consequently in higher implied volatilities and more pronounced smile profiles.

To cope with the possibility of rare extreme events, a new class of continuous-

time financial models combining a diffusive evolution of prices along with jumps at random times has been introduced. These are known as *jump-diffusion* models. According to them, the asset price is described as a Lévy process with a Gaussian component plus a jump part, which is a compound Poisson process with finitely many jumps in every time interval. The first example of such models is the Merton jump-diffusion model with Gaussian jumps (Merton (1976)) while a more refined specification is given by Kou (2002) introducing double exponential jumps. Another remarkable representative of this class is discussed by Bates (1996a) which combines compound Poisson jumps and stochastic volatility. These models have different advantages. They imply a closed form conditional characteristic function of the asset return which can be used to compute option prices with Fourier transform method. They are easy to be simulated with the possibility of being involved for efficient Monte Carlo pricing. They are able to reproduce the particularly steep skew of implied volatilities for short maturities. Nevertheless, they are difficult to be estimated since the underlying volatility is a latent process.

1.2 Discrete-time modelling approach

Although continuous-time models for stochastic volatility rapidly became a standard both in academia and industry, originally the fact that the asset price volatility is time-varying has been studied within the discrete-time framework of financial econometrics. As observed in the work by McNees (1979), the uncertainty related to the usual forecasts based on standard autoregressive models seems to vary widely over time. Furthermore, it has been also documented that large and small errors tend to cluster together in contiguous time periods as reported by Mandelbrot (1963) and Fama (1965). These evidences of heteroschedasticity of the volatility has motivated the development of the ARCH and GARCH processes by Engle (1982) and Bollerslev (1986), respectively.

The ease of tractability and the ability to reproduce the volatility clustering has determined the success of this class of models, that, since then, has been extended in order to account for other features of volatility dynamics. For example,

it is well-recognized that relevant information is carried by the direction of returns. Particularly, it is observed that stock market declines are followed by an increase of volatility while there is not evidence of a sizeable volatility decline when a rise occurs.. This asymmetric response of volatility to the sign of shocks in returns is known as leverage effect. The variety of GARCH models describing this phenomenon includes the EGARCH by Nelson (1991), the NGARCH by Engle and Ng (1993) and the TARCH by Glosten et al. (1993). Another relevant characteristic of volatility dynamics is its long memory or the high persistence of volatility autocorrelations at long time lags (usually up to several months). A typical approach is modelling volatility using a fractionally integrated process as in Baillie et al. (1996). Alternative models belongs to the multi-component family as the CGARCH by Engle and Lee (1999), recently extended by Christoffersen et al. (2008) and Bormetti et al. (2015b). These models introduce the novelty to consider the volatility process as a combination of different components each carrying information from different time-scales. In these sense they belong to the class of multi-factor models.

As concerns the possibility to incorporate jump components in price and volatility within this wide class of discrete-time models, we cite the alternative specifications of GARCH-Jump processes proposed in Maheu and McCurdy (2004), Duan et al. (2004, 2006) and Christoffersen et al. (2008).

A common characteristic of the cited GARCH-type models is that volatility is treated as a hidden process which has to be filtered from the observed series of returns. Besides, the fractionally integrated models are considered as convenient mathematical tools to achieve long memory but lack a clear economic interpretation as observed by Comte and Renault (1998). An alternative discrete-time approach to model time-varying conditional volatility has been proposed by Corsi (2009) which uses the realized volatility (RV) as an observable measure of the latent integrated volatility and introduces an heterogeneous autoregressive time dependence of present RV on past realizations aggregated on different time scales (daily, weekly and monthly). The advantages of using this specification named HAR-RV are many. First, it remains parsimonious in the number of parameters involved and easy to be estimated. Second, while formally not belonging to the

class of long-memory models, the HAR-RV is able to reproduce the volatility persistence observed in the empirical data as well as other stylized facts as fat-tailed return distributions. Third, by using heterogeneous past information from different time-scales, it reconciles with the so-called volatility cascade due to the heterogeneity of market agents' temporal horizons. This phenomenon has been studied for the first time by Muller et al. (1997).

Recently the HAR-RV specifications have been modified in order to build a new family of discrete-time stochastic volatility models named HARG-RV where realized volatility is described as an autoregressive gamma process (see Gourieroux and Jasiak (2006)) whose conditional distributions embed an heterogeneous dependence from past realizations aggregated on different time-horizons, as proposed by Corsi et al. (2013). The class also includes models able to reproduce the leverage effect as the HARGL-RV by Corsi et al. (2013) and the LHARG-RV by Majewski et al. (2015). This last version has the peculiarity to introduce a dependence on heterogeneous leverage terms in the volatility dynamics, following the indication coming by Corsi and Renò (2012). In this work, the authors show that extending the heterogeneous structure for the leverage effect is apt to mimic the empirically observed persistence of the impact of negative returns on future volatility and successfully improves forecasting. Starting with Stentoft (2008), realized volatility models have been successfully implemented for pricing purposes showing higher performances with respect to the GARCH class as documented in Corsi et al. (2013), Christoffersen et al. (2014) and Majewski et al. (2015).

The summarized class of realized volatility models has been developed to describe only the diffusive component of the dynamics of an asset and its volatility, thus neglecting the contribution of sudden jumps. Only recently, the possibility to use realized measure of volatility to discriminate the continuous and the discontinuous components of the time evolution of prices and explicitly jointly model them has been explored. The references of this new field are Christoffersen et al. (2015), Alitab et al. (2016) and Alitab et al. (2017).

Discrete-time models constitute a wide class of specifications and the advantages related to this approach are many. Firstly, financial time-series, transactions and quotes are only available in discrete time. So these model are naturally suited for

effectively incorporate this information. Secondly, they allow to accommodate different empirical properties of return and volatility process in a flexible way making more transparent the role of each parameter involved. Thirdly, by considering the class of affine process for financial application, it has been evidenced by Darolles et al. (2006) and Gourieroux et al. (2006) that, in discrete-time, this class is much larger than the equivalent continuous-time class proposed by Duffie et al. (2003), increasing considerably the type of admissible dynamics.

1.3 Observing the hidden process: Realized Volatility

A crucial feature of financial volatility is that it is not directly observable. For example, considering the daily returns of a stock or an index is not sufficient to deduce the daily volatility process since there is only one observation in a trading day. However, if intraday data, as high-frequency returns, are available, then it is possible to estimate the daily volatility. The idea originates from the seminal article of Merton (1980), according to which one can measure an asset's latent or hidden daily volatility using the sum of M intraday squared returns, when M tends to infinity. With the increasing availability of high-frequency information, this idea has made very significant progress with the definition of the realized volatility.

Formally, observing M intraday returns $y_{j,t}$ over a daily time interval $[t, t+T]$, the realized volatility is defined for the day t, as

$$\sqrt{\mathrm{RV}_t^{(M)}} = \sqrt{\sum_{j=1}^{M} y_{j,t}^2} \tag{1.1}$$

where $\mathrm{RV}_t^{(M)}$ is referred to as daily realized variance (RV)[1]. This quantity is simply the second non-centred sample moment of the return process over a day. The

[1]In the rest of the thesis the acronym RV will be used interchangeably to indicate the realized volatility or the realized variance. The exact meaning will be clear from the context.

M observations can be equally spaced (corresponding to temporal subintervals of length $\delta = T/M$) or unevenly spaced (for instance tick-by-tick with subintervals of variable length $\delta_j = t_{j+1} - t_j$).

The theoretical properties of realized volatility have been investigated in a large number of studies, among them Andersen and Bollerslev (1998), Andersen et al. (2001, 2003) and Barndorff-Nielsen and Shephard (2001, 2002a,b, 2005). The fundamental result is the proof that the realized variance represents a consistent estimator (when $\sup_j\{\delta_j\} \to 0$ as $M \to \infty$)[2] of the quadratic variation of a y_s continuous-time semimartingale, whose $y_{j,t}$ are the observed realizations at discrete times. We recall the enunciate of this result for continuous-time jump-diffusion process (for more general results involving semimartingales see Aït-Sahalia and Jacod (2014)).

Proposition 1. *Suppose y_s is a stochastic process obeying the following SDE*

$$dy_s = \alpha_s dt + \sigma_s dW_s + dJ_s$$

where α_s is predictable, σ_s is cadlag and $dJ_s = c_s dN_s$ where N_s is a Poisson process with adapted intensity process λ_s and c_s are i.i.d adapted random variables describing the size of the jumps. Given a time window $[t, t+T]$, the quadratic variation (QV) of such process is defined as

$$QV_{t,t+T} = \int_t^{t+T} \sigma_s^2 ds + \sum_{j=N_t}^{N_{t+T}} c_j^2. \tag{1.2}$$

Then, sampling M values $y_{i,t}$ in the time interval $[t, t+T]$, the realized variance converges in probability to the quadratic variation as $M \to \infty$.

Although theoretically justified under the assumption of a continuous stochastic model for the asset return process, the realized volatility approach suffers drawbacks in real applications. The main one of these issues is that the observed high-frequency trades or quotes are contaminated with market microstructure noise. This terminology refers to a variety of frictions related to the trading process as bid-ask bounces, discreteness of price changes, gradual response of prices

[2] In the rest of the chapter this condition will be synthetically indicated as $M \to \infty$.

to a block trade, inventory control effects. Intuitively, it can be argued that an asset's high-frequency price consists of the true efficient latent price plus a random observation error, due to microstructure noise. When the asset's squared return is calculated, the error is not averaged out when it is summed, systematically generating a bias in the variance estimate. This empirical fact indicates that the realized volatility estimator is not robust when the sampling interval is small. Furthermore, as observed by Zhang et al. (2005), it suggests that, from a probabilistic point of view, the observed log return process is not in fact a semimartingale.

To cope with this problem, several techniques has been proposed for estimating latent volatility in presence of microstructure noise. One example is based on the optimal choice of sampling frequency as in Aït-Sahalia et al. (2005) and Bandi and Russell (2011); other ones involve filters based on an AR(p) or MA(q) models of intraday data (Ebens (1999), Andersen et al. (2001) and Hansen et al. (2008)) or the so-called Realized Kernel (Barndorff-Nielsen et al. (2008, 2011b,a)) or the pre-averaging of high-frequency data (Jacod et al. (2009), Christensen et al. (2010) and Hautsch and Podolskij (2013)). Another remarkable methodology has been advanced by Zhang et al. (2005) and Aït-Sahalia et al. (2011), known as the "Two-Scale" method. It is based on the idea to combine two different sampling frequencies in order to take advantage of each of them. On one hand, in fact, high frequency data carry important information representing the closest available to the continuous time process but suffer the contamination due to market microstructure. On the other hand, sampling over longer time horizons reduces the impact of the microstructure bias but has the inconvenience of throwing away data which is not a desirable solution in statistics. The authors have solved this issue defining and combining two estimators, one making use of the full data by subsampling at a low time scale (say, 5 minutes) , the other one using every single observation (say, tick-by-tick), as in the definition of realized volatility. The result is an unbiased and consistent estimator for the integrated volatility despite the presence of market microstructure noise.

1.4 Jump detection tests

The advent of high-frequency data had a significant impact on the recent progress of the econometrics of financial volatility through the use of realized variances (volatilities). Along with these developments, it contributed to the introduction of nonparametric methods for the direct identification of jumps.
A jump detection test is correctly specified by two ingredients:

(i) a consistent measure of the diffusive or integrated volatility of the price process with the property to be robust to rare jump events;

(ii) the asymptotic distribution of the observed volatility in the limit of infinite return realizations under the assumption of absence of discontinuity of the price dynamics.

This assumption represents the so-called null hypothesis of the test which is rejected in accordance to an arbitrarily chosen statistical tolerance given by the $\alpha-$percentile of the asymptotic distribution (usually 95% or 99%). The first jump tests introduced in the literature have been developed by Barndorff-Nielsen and Shephard (2004, 2006). They compare the realized variance, which converges, as the time between observations approaches zero, to the total quadratic variation (equation (1.2)) and the bipower variation (BPV), which is a consistent estimate of the integrated variance as the number of observations tends to infinity and robust to discontinuities in the price path. From a practical point of view, we observe that all the robust measure recalled in this section could be affected by microstucture noise when they are built from high-frequency data as discussed for realized volatility. There exists a consensus that computing these measure using returns aggregated on a 5-minutes scale is a reasonable choice to reduce the noise still preserving the power of the test.

Proposition 2. *If y_s is a continuous-time jump-diffusion process and $y_{i,t}$ are M observations sampled discretely over the time window $[t, t+T]$ then the bipower variation, defined as*

$$\mathrm{BPV}_{t,t+T}^{(M)} = \mu_1^{-2} \sum_{i=2}^{M} |y_{i-1,t}| \, |y_{i,t}| \tag{1.3}$$

with $\mu_1 = \sqrt{2/\pi}$*, converges in probability to the integrated volatility* $\int_t^{t+T} \sigma_s^2 ds$*, as* $M \to \infty$*.*

The assessment of the statistical significance of the difference between RV and BPV is done by computing an estimate of the integrated quarticity. Barndorff-Nielsen and Shephard (2004, 2006) recommend the realized quadpower quarticity (QPV)

$$\mathrm{QPV}_{t,t+T}^{(M)} = \mu_1^{-4} \frac{1}{\delta} \sum_{i=4}^{T/\delta} |y_{i-3,t}| \, |y_{i-2,t}| \, |y_{i-1,t}| \, |y_{i,t}| \, .$$

The proposed test statistics is finally based on the following asymptotically standard normal test, or $z-$test,

$$z = \delta^{-1/2} \frac{\mathrm{RV}_{t,t+T}^{(M)} - \mathrm{BPV}_{t,t+T}^{(M)}}{\sqrt{\left(\frac{\pi^2}{4} + \pi - 5\right) \mathrm{QPV}_{t,t+T}^{(M)}}} \xrightarrow[p]{M \to \infty} \mathcal{N}(0,1)$$

where the null hypothesis that jumps did not occurred in a day is rejected at a 1% level of significance if the $z-$statistics is greater than 2.33. An alternative $z-$test which uses jump-robust realized tripower quarticity (TriPV) is suggested by Andersen et al. (2007a)

$$\mathrm{TriPV}_{t,t+T}^{(M)} = \mu_{\frac{4}{3}}^{-3} \frac{1}{\delta} \sum_{i=4}^{T/\delta} |y_{i-2,t}|^{\frac{4}{3}} \, |y_{i-1,t}|^{\frac{4}{3}} \, |y_{i,t}|^{\frac{4}{3}} \, ,$$

$$z = \delta^{-1/2} \frac{\mathrm{RV}_{t,t+T}^{(M)} - \mathrm{BPV}_{t,t+T}^{(M)}}{\sqrt{\left(\frac{\pi^2}{4} + \pi - 5\right) \mathrm{TPV}_{t,t+T}^{(M)}}} \xrightarrow[p]{M \to \infty} \mathcal{N}(0,1) \, ,$$

where $\mu_{\frac{4}{3}} = 2^{2/3} \Gamma(7/6) \, \Gamma(1/2)^{-1}$. Other competitor tests are based on the following robust to jumps estimators: threshold realized variance (Mancini (2009)), the median or the minimum realized variation (Andersen et al. (2012)) and the realized threshold multipower variation (Corsi et al. (2010)). The common characteristic among these tests is that the null hypothesis of absence of jumps can be statistically assessed within a certain time period such as a trading day. When rejected, the conclusion is that the asset return path has at least one discontinuity

over the chosen time interval. A different answer is given by the alternative tests developed by Andersen et al. (2007b, 2010) and Lee and Mykland (2008) which are able to reject the null hypothesis of continuity of the sample path at any observations, allowing for the identification of the time of a jump and consequently the total number of discontinuities during a day. The basic idea is to define a jump-robust measure for the instantaneous volatility for each observation using corresponding local movements of returns. Then, the ratio of a realized return at a given time to the estimated instantaneous volatility creates the test statistics for jumps.

It is worth mentioning that recent achievements question that jumps are only attributed to the discontinuous component of the price semimartingale process. For instance, Christensen et al. (2014) argue that jump events are often spurious detections resulting from the aggregation of returns at larger time scales. A second alternative hypothesis is given by Christensen et al. (2016) who show that historical time series does not rule out continuous-time models with no discontinuous component but where the drift coefficient may exhibit local bursts.

For the purpose of this thesis, we will use the methodology suggested by Corsi et al. (2010) along with the one described by Andersen et al. (2010). We will adopt the threshold multipower estimator (TMPV), defined as,

$$\mathrm{TMPV}_{t,t+T}^{(M),[\gamma_1,...,\gamma_N]} = \left(\prod_{k=1}^{N}\mu_{\gamma_k}\right)^{-1} \delta^{1-\frac{1}{2}(\gamma_1+...+\gamma_N)} \sum_{j=N}^{M}\prod_{k=1}^{N} |y_{j-k+1}|^{\gamma_k} \mathcal{I}_{\left\{|y_{j-k+1}|^2 \leq \vartheta_{j-k+1}\right\}}$$

where $\gamma_1, ..., \gamma_N > 0$ and $\vartheta_s : [t, t+T] \to \mathbb{R}^+$ is a strictly positive random threshold function. It is proved that this estimator admits a central limit theorem in the presence of jumps resulting a consistent measure of the multipower integrated volatility $\int_t^{t+T} \sigma_s^{\gamma_1+...+\gamma_N} ds$ (for the proof see Corsi et al. (2010)). The choice of the threshold estimates has the advantage of correcting the bias affecting the multipower variation for finite time intervals between observed returns which can be extremely large in the case two consecutive observations contain jumps. The

relative test is named Tz−test and has the following form

$$\mathrm{T}z = \delta^{-1/2} \frac{\left(\mathrm{RV}^{(M)}_{t,t+T} - \mathrm{TBPV}^{(M)}_{t,t+T}\right)\left(\mathrm{RV}^{(M)}_{t,t+T}\right)^{-1}}{\sqrt{\left(\frac{\pi^2}{4} + \pi - 5\right)\max\left\{1,\ \frac{\mathrm{TTriPV}^{(M)}_{t,t+T}}{\mathrm{TBPV}^{(M)}_{t,t+T}}\right\}}} \xrightarrow[p]{M\to\infty} \mathcal{N}(0,1)\,,$$

with $\mathrm{TBPV}^{(M)}_{t,t+T} = \mathrm{TMPV}^{(M),[1,1]}_{t,t+T}$ indicating the threshold bipower variation and $\mathrm{TTriPV}^{(M)}_{t,t+T} = \mathrm{TMPV}^{(M),\left[\frac{4}{3},\frac{4}{3},\frac{4}{3}\right]}_{t,t+T}$ standing for the threshold tripower quarticity. The Tz can be seen as a modification of the z−test advanced by Huang and Tauchen (2005) which involves in the numerator the relative difference between the realized measure of the variance and a realized estimate of its continuous component instead of the absolute difference.

The Tz−statistics allows to determine if at least one jump occurs during a trading day but it does not discern the total number of intraday jumps. In order to obtain this information, we propose to combine the test with the sequential identification procedure described by Andersen et al. (2010) that we summarize as follows: i) we perform the Tz−test for a trading day and determine if an extreme event has occurred or not; ii) if we detect a jump, we identify the largest intraday return (in absolute value), we substitute it with the mean of the other intraday returns and recompute our estimator of the realized variance of the trading day; iii) we repeat the test and if it does not reject the null we conclude that there is only one jump contained in the removed largest return; iv) if the test rejects the null it means that more than one jump happen during the day and the entire procedure is repeated removing the second largest return and so on. Finally, we get the series of days in which the price path shows at least one discontinuity and the corresponding number of intraday jumps, from which we are able to separate the continuous and the discontinuous contributions to the total return over the trading day.

Chapter 2

Risk-neutral measure for discrete-time processes

2.1 Stochastic discount factor modelling

For almost a decade discrete-time modelling approach, typical of financial econometrics, and no-arbitrage asset pricing remained separated fields mainly because the latter was essentially developed in the continuous-time framework of diffusion processes and jump-diffusion processes.

The central problem of pricing is to assign a rational valuation to a financial risky asset providing a stochastic payoff at future times. The word rational indicates that agents fix the price of their investment combining personal beliefs about the future states of the economy and their attitude towards risk under the assumption of absence of arbitrage opportunities. This economic request is mathematically formalized by introducing a pricing operator. This associates a price to every financial instrument embedding the compensation demanded by market agents for bearing investment uncertainty.

To fix the concept, we follow Bühlmann et al. (1998). We consider an economy composed by a risk-free asset with interest rate r and a risky security with price S_t and log-return

$$y_t = \log\left(\frac{S_t}{S_{t-1}}\right)$$

defined on a filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{1\leq t\leq T}, \mathbb{P}\right)$ satisfying standard conditions. We assume that the risky asset price evolves at discrete times till some horizon T, generating the space of the trajectories $(S_1, S_2, ..., S_T) \in \mathbb{R}^T$. We define a set of payoffs X belonging to the space $L_T^2 = L^2\left(\mathbb{R}^T, \mathbb{P}\right) = \{X : \mathbb{E}^{\mathbb{P}}\left[\sum_{i=1}^{T} X_i^2 < \infty\right]\}$. A pricing operator is any operator $\mathcal{O}_T : L_T^2 \to \mathbb{R}$ which associates a real value or price to any payoff X.

Definition 1. *A payoff $X \in L_T^2$ is an arbitrage opportunity if $X_j \geq 0$ for every $j \in \{1, 2, .., T\}$ almost surely with non-zero probability of one component being positive and having price $\mathcal{O}_T(X) \leq 0$.*

The operator $\mathcal{O}_T$ is said to be positive if $\mathcal{O}_T(X) \geq 0$ for $X \geq 0$ almost surely with the inequality to be intended componentwise. The operator is said strictly positive if it is positive and additionally $\mathcal{O}_T(X) > 0$ if $\mathbb{P}(X_j > 0 \text{ for some } j \in \{1, 2, .., T\}) > 0$. The importance of the class of linear and strictly positive pricing operators is stated by the following

Theorem 1. *A market is free of arbitrage opportunities if and only if there exists a strictly positive linear pricing operator.*

The existence of the pricing operator implies the existence of a particular stochastic process as expressed by

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

By expanding the last expression in Theorem 2 as

$$\mathcal{O}_T(X) = \mathbb{E}\left[\tilde{M}X\right] = \mathbb{E}\left[\sum_{i=1}^{T} \tilde{M}_i X_i\right]$$

we observe that the process $\tilde{M} = \left(\tilde{M}_1, \tilde{M}_2, ..., \tilde{M}_T\right)$ has a simple economic interpretation of discounting the future value of a payoff X. Further, we introduce the process M such that $M_t = e^{rt}\tilde{M}_t$ componentwise and rewrite the expression of the price associated to each component $t > 0$ of the stream of payoffs X at

present time $t = 0$ as

$$X_{0,t} = \mathcal{O}_t\left(X_t\right) = \mathbb{E}^{\mathbb{P}}\left[\tilde{M}_t X_t\right] = \mathbb{E}^{\mathbb{P}}\left[e^{-rt} M_t X_t\right]. \tag{2.1}$$

With this choice we are able to separate the risk discounting from temporal discounting represented by the usual term e^{-rt} relative to the presence of a risk-free asset. Hence the process M represents the discounting term accounting for risks associated with the future states of the economy and it is called the Stochastic Discount Factor (SDF). Using equation (2.1) and observing that an investment of 1 unit of the risk-free asset B returns the deterministic payoff e^{rt} at time t, we can determine the mean of the SDF as

$$\mathbb{E}^{\mathbb{P}}\left[M_t\right] = \mathbb{E}^{\mathbb{P}}\left[e^{rt}\tilde{M}_t\right] = \mathcal{O}_t\left(B_t\right) = 1. \tag{2.2}$$

The expression (2.1) connecting the existence of a pricing operator to that one of a stochastic discount factor can be reformulated defining a new probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$. We denote by $\mathbb{P}_t$ the family of probability measures such that $\mathbb{P}_t = \mathbb{P}|\mathcal{F}_t$ and construct the equivalent probability measures $\mathbb{Q}_t$ satisfying relations $d\mathbb{Q}_t = M_t d\mathbb{P}_t$. We observe that the last relation determines the change of probability measure. Using notation $\mathbb{P} = \mathbb{P}_T$ and $\mathbb{Q} = \mathbb{Q}_T$, the $\mathbb{Q}$ probability that the value of the random variable S_t belongs to a set A is given by

$$\mathbb{Q}\left(S_t \in A\right) = \mathbb{E}^{\mathbb{P}}\left[\mathcal{I}_A\left(S_t\right) M_t\right] \tag{2.3}$$

where $\mathcal{I}_A$ is the indicator function of a set A. Introducing the change of measure from $\mathbb{P}$ to $\mathbb{Q}$, the price of a payoff X_t as expressed in (2.1) is rewritten according to the following

$$X_{0,t} = \mathbb{E}^{\mathbb{P}}\left[e^{-rt} M_t X_t\right] = \mathbb{E}^{\mathbb{Q}}\left[e^{-rt} X_t\right]. \tag{2.4}$$

We note that the last equality suggests that the quantity $e^{-rt}X_t$ is in fact a $\mathbb{Q}$-martingale. This result reconciles with the usual no-arbitrage pricing framework which represents the core of financial mathematics. The $\mathbb{Q}$ measure is well-known with the name of risk-neutral measure or equivalent martingale measure.

All the previous results can be summarized by the following fundamental theorem

Theorem 3 (The First Fundamental Theorem of Asset Pricing)**.** *The following conditions are equivalent:*

1. *There are no arbitrage opportunities.*
2. *A strictly positive linear pricing operator exists.*
3. *A stochastic discount factor exists.*
4. *A risk-neutral probability measure exists.*

Along with this theorem which states the conditions for the existence of strictly positive linear pricing operator and of stochastic discount factor, the Second Fundamental Theorem of Asset Pricing establishes conditions for their uniqueness. This theorem uses the concept of completeness of a market which indicates that agents can construct an investment strategy that replicates the payoffs of any claim available in the market.

Theorem 4 (The Second Fundamental Theorem of Asset Pricing)**.** *An arbitrage-free market is complete if and only if there exists a unique stochastic discount factor.*

Examples of complete markets are the multiperiod binomial model in discrete-time (Cox et al. (1979) and Rendleman and Bartter (1979)) and the Black-Scholes model in continuous-time (Black and Scholes (1973)). On the contrary, stochastic volatility models and short interest rate models are typical instances of incomplete markets.

From a modelling perspective, a stochastic discount factor represents a powerful tool to infer relevant information about the risk premia embedded in securities and derivatives prices by market agents in their trading activity. Related to this aspect, it has the important advantage to consistently reconcile the physical or historical information with the restrictions imposed by no arbitrage conditions. This issue is usually circumvented by modelling the price process directly in the risk-neutral measure, an approach which guarantees a flexible modelling framework but lacks in identification and separation of the different components contributing to price formation.

2.2 Examples of stochastic discount factors

This section provides some examples of stochastic discount factors in financial theory generally derived under either an equilibrium condition, or no-arbitrage assumptions in continuous-time specification (see Gourieroux and Monfort (2007) for more details). As we will see they belong to the class of exponential-affine functions of underlying state variables.

Example: *Consumption based CAPM (CCAPM).* The CCAPM model is based on the assumption that investors make their portfolio choices in order to maximise intertemporal expected utility of current and future consumption (see Lucas (1978), Breeden (1979), Mankiw and Shapiro (1986)). Within this framework, financial securities allow the consumer to manage her consumption over time, selling assets to finance consumption in "bad" periods of the economy and saving in "good" ones.
Let us assume the existence of a representative agent and indicate with U the utility function and with θ the intertemporal subjective discount rate. The equilibrium price p_t at time t of a financial asset with payoff X_{t+1} at time $t+1$ is given by the following

$$p_t = \mathbb{E}^{\mathbb{P}}\left[X_{t+1}\frac{q_t}{q_{t+1}}\theta\frac{U'(C_{t+1})}{U'(C_t)}\right]$$

where q_t is the cost of the consumption good, $U'(C_t) = \frac{dU}{dC}(C_t)$ and C_t is the quantity consumed at date t. From this equation we easily individuate the stochastic discount factor for the CCAPM

$$M_{t+1} = \frac{q_t}{q_{t+1}}\theta\frac{U'(C_{t+1})}{U'(C_t)}.$$

If we make the further assumption that agents have a power utility function, the SDF associated with the model becomes

$$M_{t+1} = \frac{q_t}{q_{t+1}}\theta\left(\frac{C_{t+1}}{C_t}\right)^{\gamma} = \exp\left[\log\theta - \log\frac{q_{t+1}}{q_t} + \gamma\log\frac{C_{t+1}}{C_t}\right].$$

We recognize that it is an exponential-affine function of two state variables, which are the logarithmic inflation rate and the logarithmic ratio of consumption.

Example: *Stochastic volatility model.* In its simplest one-factor form, a stochastic volatility model is described by a couple of one-dimensional Itô process satisfying the following stochastic differential equations

$$\begin{aligned} dS_t &= \mu_t S_t dt + \sigma_t S_t dW_t^S, \\ dY_t &= a(Y_t)dt + b(Y_t)dW_t^Y, \end{aligned}$$

where S_t is the asset price, Y_t is a driving factor, $\sigma_t = f(Y_t)$ is the stochastic volatility with function f smooth, positive and increasing and W_t^S, W_t^Y are typically correlated Brownian motions. Assuming a constant risk-free rate r, the application of the Girsanov's theorem determines the expression for the stochastic discount factor

$$\begin{aligned} M_{t,t+1} = &\exp(-r)\exp\left\{-(\mu-r)\int_t^{t+1}\frac{dW_s^S}{\sigma_s} - \frac{1}{2}(\mu-r)^2\int_t^{t+1}\frac{ds}{\sigma_s^2}\right\} \\ &\times \exp\left\{-\int_t^{t+1}\nu_s dW_s^Y - \frac{1}{2}\int_t^{t+1}\nu_s^2 ds\right\}, \end{aligned}$$

where ν_t is the risk premium related to stochastic volatility which arises from the incompleteness of the market and can be chosen as an arbitrary function of the time and past values of S and σ. We observe that the above expression has a discrete-time counterpart easily derived by approximating the integrals as

$$M_{t,t+1} = \exp(-r)\exp\left\{-(\mu-r)\frac{\epsilon_{t+q}^S}{\sigma_t} - \frac{1}{2}(\mu-r)^2\frac{1}{\sigma_t^2} - \nu_t\epsilon_{t+1}^Y - \frac{1}{2}\nu_t^2\right\}.$$

This is an exponential-affine function of standard normal innovations ϵ_t^S, ϵ_t^Y corresponding to return and volatility processes with generally path dependent coefficients.

2.3 Exponential-affine SDF

The basic problem in modelling asset prices is the specification of the stochastic discount factor which guarantees a rational no-arbitrage valuation for a contingent claim. As stated in Theorem 3, this issue is equivalent to find a process $M = (M_1, M_2, ..., M_T)$ such that, given an asset $X = (X_1, X_2, ..., X_T)$, the product $M_k X_k$ is a martingale under the physical (or historical) measure $\mathbb{P}$. This problem is usually circumvented by modelling the asset dynamics directly under the risk-neutral (or equivalent martingale) measure $\mathbb{Q}$ which absorbs the stochastic discount factor as expressed by equation (2.4). We suppose to have a market model as described in Section 2.1 and we assume that the log-return process y_t is driven by an $n-$dimensional vector of risk factors $\mathbf{f}_t$. Following Bühlmann et al. (1996, 1998), we adopt a multi-dimensional version of the Esscher Transform to build the SDF which achieves the martingale property described before. We choose the following SDF

$$M_t = \prod_{s=0}^{t-1} M_{s,s+1} \tag{2.5}$$

where

$$M_{s,s+1} = \frac{e^{-\boldsymbol{\nu}_{\mathbf{f}} \cdot \mathbf{f}_{s+1} - \nu_y y_{s+1}}}{\mathbb{E}^{\mathbb{P}}\left[e^{-\boldsymbol{\nu}_{\mathbf{f}} \cdot \mathbf{f}_{s+1} - \nu_y y_{s+1}} \middle| \mathcal{F}_s\right]} \tag{2.6}$$

which depends parametrically on $\boldsymbol{\nu}_{\mathbf{f}} \in \mathbb{R}^n$ and $\nu_y \in \mathbb{R}$. These are usually referred to as factor loadings or risk sensitivities or risk premia.

Expression (2.6) is inspired to the one-dimensional and unconditional version of parametric transform of probability measure originally introduced by Esscher (1932) in an actuarial context to approximate the distribution of aggregate claims. This change of measure has been pioneered in derivative pricing by Gerber and Shiu (1994) who extended the original idea by Esscher to a class of stochastic processes commonly used to model stock price dynamics including the Wiener process, the Poisson process, the gamma process and the inverse Gaussian process.

We observe that the selected SDF in formula (2.6) is exponential-affine. The

reasons behind this specification are manifold and has been widely documented in the asset pricing literature. First, it is derived naturally in equilibrium models like CCAPM (as reported in Section 2.2), consumption-based models with habit formation or Epstein-Zin preferences (see, among others, Campbell and Cochrane (1999), Bansal and Yaron (2004), Garcia et al. (2006), Eraker (2008)). Second, the SDFs implied by asset pricing models in continuous-time framework assume the exponential-affine form if they are time-discretized (as described in Section 2.2). Third, the exponential-affine SDF is particularly well adapted to compute the Moment Generating Function, the Fourier Transform or the Laplace Transform which are central tools in discrete-time asset pricing theory. Applications of this technique can be found in a general approach by Gourieroux and Monfort (2007) and Bertholon et al. (2008) or in the class of CAR (Compound-Autoregressive or discrete-time affine) models by Darolles et al. (2006). They are also involved for interest rate models by Gourieroux et al. (2002), exchange rate models by Gourieroux et al. (2010), credit risk models by Gourieroux et al. (2006) and, recently, for realized volatility models by Corsi et al. (2013), Majewski et al. (2015) and Christoffersen et al. (2015). Finally, a further reason for the convenience of choosing exponential-affine SDFs like (2.6) has been given by an economical argument in Bühlmann et al. (1998): Assuming that investors have exponential utility in the economy, the pricing operator defined by exponential-affine stochastic discount factor (or Esscher transform) is a solution to Pareto optimal allocation problem of wealth among agents in the economy.

2.4 Solving pricing puzzles with multidimensional SDF

Multidimensional stochastic discount factor of the form given in (2.6) has been recently involved to investigate some open questions in the literature on option valuation. Among them, Bates (1996b) pointed out the empirical issue to establish whether the distributions implicit in option prices are consistent with the time series properties of the underlying asset prices. Subsequent studies, addressing this problem, have evidenced difficulties to reconcile the empirical distributions of

spot returns with risk-neutral distributions embedded in derivative prices. This is troublesome because suggests that the existing stochastic discount factors are not adequate to explain option prices and that more general specifications are needed.

These studies have found many other stylized facts and empirical puzzles related to option pricing. For instance, a well-established fact is that implied volatilities tend to exceed realized volatilities which can be explained by the presence of a price for variance risk (according to the diffused convention in literature this is usually described by a negative variance premium as in Bakshi and Kapadia (2003)). This "expectations puzzle" claims that implied variance can not be considered an unbiased estimate of future variance. Another issue consists in the explanation of the "overreaction puzzle" following the findings by Stein (1989) e Poteshman (2001) indicating that long-term implied variance overreacts to change in short-term variance. In addition, several papers [1] have recognized that available models seem to lack a complete explanation of the cross-section of option prices, specifically the prices of out-of-the-money options. In particular, empirical stochastic discount factors implied in index options are not monotonically decreasing in underlying returns as standard financial theory assumes (as in Rubinstein (1976) and Brennan (1979)) but appear to be U-shaped functions of returns (see Figure 2.1). This poses the so-called "pricing kernel puzzle" (Jackwerth (2000)). Among the numerous existing studies on this subject, Christoffersen et al. (2013) have recently advanced a unified explanation for these puzzles by introducing a stochastic discount factor that is more general than the standard pricing kernels that are monotonic in asset returns. As proposed by the authors, the key to solve the puzzles comes with incorporating a variance premium by using a multidimensional stochastic discount factor that is monotonic both in returns and variance. Specifically, they generalize the standard Duan (1995) and Heston and Nandi (2000) GARCH models including the following stochastic

[1] Veronesi (2001), Chabi-Yo et al. (2008), Bates (2008), Bakshi et al. (2010b), Brown and Jackwerth (2012)

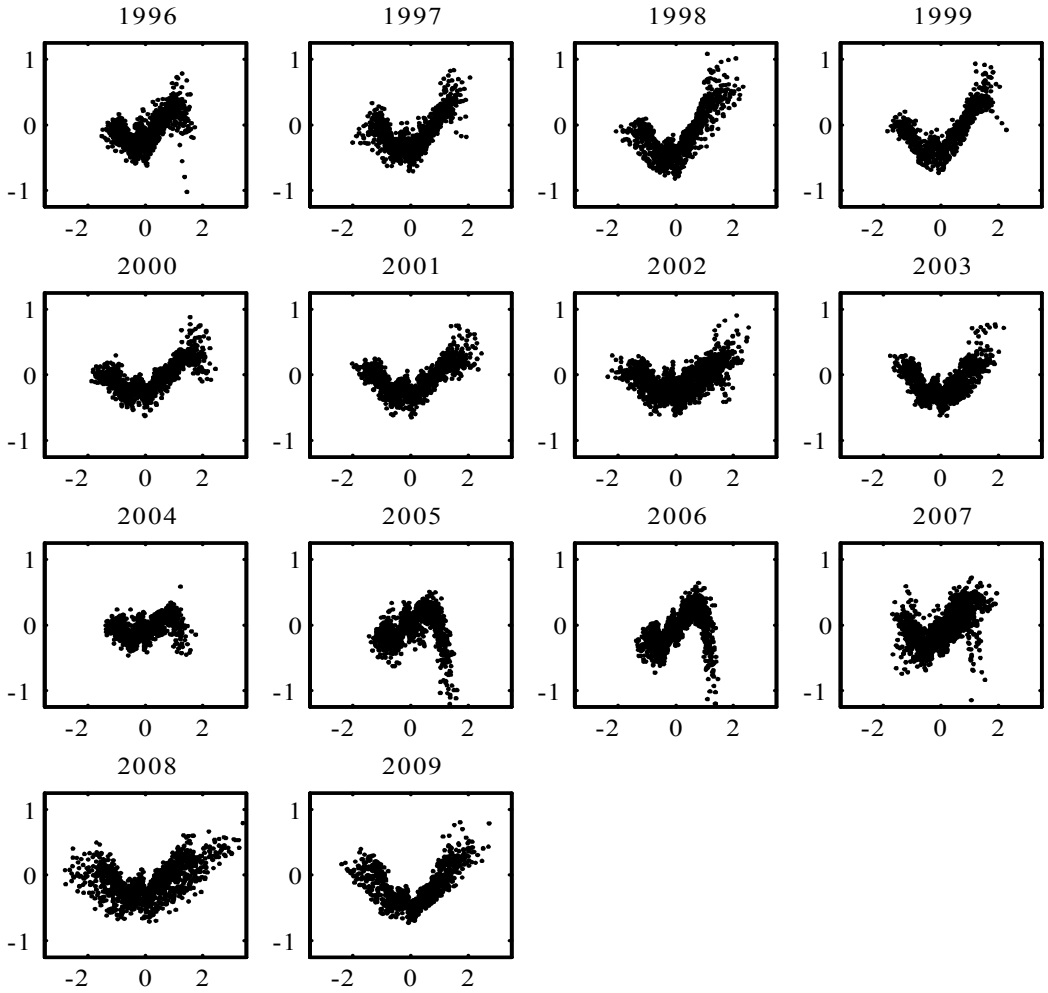

Figure 2.1: Empirical SDFs obtained as the log ratios of risk-neutral one-month returns densities and physical one-month returns histograms. The return sample is derived for the S&P500 total return index from 1996 to 2009. From Christoffersen et al. (2013).

discount factor

$$M_t = M_0 \left(\frac{S_t}{S_0}\right)^{\phi} \exp\left(\delta t + \eta \sum_{s=1}^{t} h_s + \xi \left(h_{t+1} - h_1\right)\right) \tag{2.7}$$

where $S(t)$ and $h(t)$ respectively denote the asset price process and its variance process. Parameters δ and η describe the time-preference, while ϕ and ξ tune the aversion to equity and variance risk, respectively. We observe that this specification belongs to the exponential-affine class.

With this choice, it is proven the following proposition establishing the relation between variance process and model parameters in the passage from the physical

to the risk-neutral measure.

Proposition 3. *Assuming that the daily log-return process y_t follows the Heston and Nandi (2000) GARCH process under physical measure,*

$$y_t = r + \left(\mu - \frac{1}{2}\right) h_t + \sqrt{h_t} z_t,$$
$$h_t = \omega + \beta h_{t-1} + \alpha \left(z_{t-1} - \gamma \sqrt{h_{t-1}}\right)^2,$$

where r is the daily risk-free rate and z_t is a standard normal noise, then the risk-neutral log-return process corresponding to the physical process and the stochastic discount factor (2.7) is still a GARCH process

$$y_t = r - \frac{1}{2} h_t^* + \sqrt{h_t^*} z_t^*,$$
$$h_t^* = \omega^* + \beta h_{t-1}^* + \alpha^* \left(z_{t-1}^* - \gamma^* \sqrt{h_{t-1}^*}\right)^2,$$

where z_t^* is a standard normal noise and the following relations hold

$$h_t^* = \frac{h_t}{1 - 2\alpha\xi},$$
$$\omega^* = \frac{\omega}{1 - 2\alpha\xi},$$
$$\alpha^* = \frac{\alpha}{(1 - 2\alpha\xi)^2},$$
$$\gamma^* = \gamma - \phi.$$

We observe that the physical parameters are mapped into the risk-neutral by some functions depending on the values of the pricing kernel parameters ϕ and ξ. In addition, the risk-neutral variance process is also modified by the presence of the scaling factor $(1 - 2\alpha\xi)$. This is a crucial result due to the particular specification of the stochastic discount factor in (2.7) in clear contrast to the Heston and Nandi (2000) model for which the physical and the risk-neutral variance process are assumed to be the same. In this framework, the kernel parameter ξ plays an important role because it influences the level, persistence and volatility of the variance. In particular, when $\xi > 0$, the pricing kernel puts more weight on the tails of innovations and the risk-neutral variance h_t^* exceeds the physical vari-

ance h_t, in accordance to what usually observed from real data. In the following corollary, we recall some results related to the GARCH model combined with the pricing kernel in (2.7) (for details see Christoffersen et al. (2013)).

Corollary 1. *If the equity premium is positive ($\mu > 0$), the variance premium is negative ($\xi > 0$) and variance is negatively correlated with returns (leverage effect with $\gamma > 0$) then:*

1. *The risk-neutral variance h_t^* exceeds the physical variance h_t;*
2. *The risk-neutral expected future variance exceeds the physical expected future variance;*
3. *The risk-neutral variance process is more persistent than physical process;*
4. *The risk-neutral variance of variance exceeds the physical variance of variance.*

This summarizes how a pricing kernel including a variance premium can explain the questions posed by the cited puzzles concerning the relationship between the implied option variance and the variance filtered from observed time-series. Furthermore, this model is able to capture the observed U-shaped profile of the ratio of the risk-neutral distribution and the physical one. In fact, although the stochastic discount factor in (2.7) is a monotonic function of the log-return and variance, its projection onto the log-return alone implies a U-shape, as formalized by the following corollary (for the proof see Appendix C in Christoffersen et al. (2013)).

Corollary 2. *The logarithm of the stochastic discount factor in (2.7) is a quadratic function of the log-return,*

$$\ln\left(\frac{M_t}{M_{t-1}}\right) = \frac{\xi\alpha}{h_t}\left(y_t - r\right)^2 - \mu\left(y_t - r\right) + \left(\eta + \xi\left(\beta - 1\right) + \xi\alpha\left(\mu - \frac{1}{2} + \gamma\right)^2\right) h_t + \delta + \xi\omega + \phi r.$$

So, the logarithm of the pricing kernel has a parabolic shape with concavity dependent on the variance premium parameter ξ. If $\xi > 0$ (implying a negative

variance risk premium), we observe that the pricing kernel has a U-shape. This result helps to understand the phenomenon of the relative thickness of the tails of the physical and risk-neutral distribution and states the importance to account for more general specification of the pricing kernel including variance risk premium.

Chapter 3

Heterogeneous autoregressive gamma models with leverage and jumps

In this chapter, we introduce two class of models where the log-return dynamics is completely determined by specifying the RV dynamics chosen among the family of the HAR-RV processes. These processes, introduced by Corsi (2009), successfully describe the impact that past realized variances aggregated on different time scales (daily, weekly and monthly) have on the current level of realized variance. Later, Corsi et al. (2013) have studied the application of the HARG-RV model to option pricing in discrete time, also introducing the HARGL extension which accounts for the leverage through a daily binary component. More recently, Majewski et al. (2015) have widened the HARG-RV class and included a heterogeneous parabolic structure for leverage, defining the LHARG-RV model.

We extend the LHARG-RV model to account for the possibility of extreme movements in the evolution of the log-return. The newly proposed models are labelled as JLHARG-RV or Heterogeneous AutoRegressive Gamma model for Realized Volatility with Leverage and Jumps and LHARG-ARJ-RV or Heterogeneous AutoRegressive Gamma model for Realized Volatility with Leverage and AutoRegressive Jumps.

For both, we assume that the diffusive or continuous component of the dynam-

ics is described by an autoregressive gamma process (see Gourieroux and Jasiak (2006)) of RV whose conditional mean is assumed to be a linear function of the past realized variances and leverage terms aggregated over different time scales (daily, weekly, and monthly).
The main difference between the two classes is the modelling of the jump component. As concerns the JLHARG, the discontinuous component is introduced by modelling the jump realized variation as a compound Poisson process with constant intensity and jump size sampled from a gamma distribution. This choice allows us to describe the pure discontinuous dynamics in a parsimonious way and preserve the direct estimation of the model on observed data without filtering.
Differently, LHARG-ARJ model assumes that jumps in log-returns are driven by a compound non homogeneous Poisson process with sizes sampled from a normal distribution and intensity linearly dependent on its first-order lagged value. Introducing a dynamics for the jump intensity allows to account for the documented clustering of jump events implying a more realistic model. On the other hand, the estimation procedure has to include a filtering technique since jump intensities are not observable.

3.1 Long Memory, leverage effect and jumps in financial time series

In this section, we briefly recall some of the well-documented stylized facts emerging from the statistical analysis of financial time series (see Cont (2001) for a comprehensive overview). For our purposes we will concentrate on those facts that have been of much interest in the development of modern volatility models and their applications to asset pricing: the long memory of volatility process, the leverage effect and the presence of jumps in asset price dynamics.
Long memory in volatility refers to the observed property that the effects of volatility shocks persist over long period of time. Formally, the volatility process σ_t, with autocorrelation function ρ_k at lag k, is said to have long memory

property if

$$\lim_{n\to\infty} \sum_{k=-n}^{n} \rho_k = \infty.$$

Nevertheless, this definition is not useful for practical purposes since nor the volatility process is directly observed nor we can deal with infinite time series. Alternatively, a typical test used to identify the high persistence of volatility consists in verifying that the autocorrelation function of a volatility measure has a power law decay. For example, using realized volatility RV_t, it can be tested if the following holds as k increases

$$\mathrm{Corr}\left(\mathrm{RV}_t, \mathrm{RV}_{t+k}\right) \sim \frac{C}{k^{\beta}} \tag{3.1}$$

with $0 < \beta < 1$. Different studies suggest that this behaviour is well reproduced by a power law with $\beta \in [0.2, 0.4]$ when absolute or squared returns series are involved (see Cont et al. (1997), Liu et al. (1997) and Cont (1998)). In Figure 3.1 we report the autocorrelation function of a realized volatility series built from HF returns of S&P500 Index Futures. We observe the typical characteristic of long-memory with the autocorrelation function assuming values higher than the 95% confidence level of no correlation up to 100 days.

Another consolidated fact is the so-called leverage effect, first observed by Black (1976). It represents the existing negative relationship between stock returns and volatility which implies the tendency of volatility to increase when the price drops. It originates from an asymmetric response of volatility to the sign of past shocks in price with negative ones having a higher impact than positive shocks of the same magnitude. In Figure 3.2 we report the correlation function (in absolute value) between negative/positive returns of S&P500 Index Futures and future levels of volatility, evidencing this aspect. The presence of a similar effect is particularly relevant for option prices. On one hand, it determines that implied volatilities of at-the-money contracts tend to increase after price drops as the realized volatilities do. On the other hand, it implies the observation of asymmetric volatility smiles. This phenomenon is also known as skew of the volatility surface

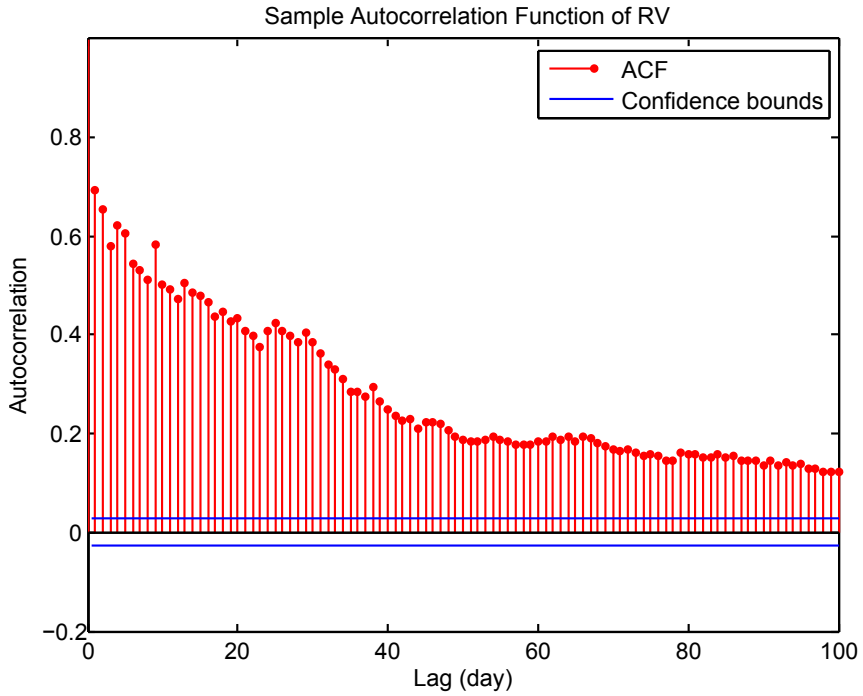

Figure 3.1: Autocorrelation function of RV time series for S&P500 Index Futures returns from 3 July 1990 to 28 June 2011.

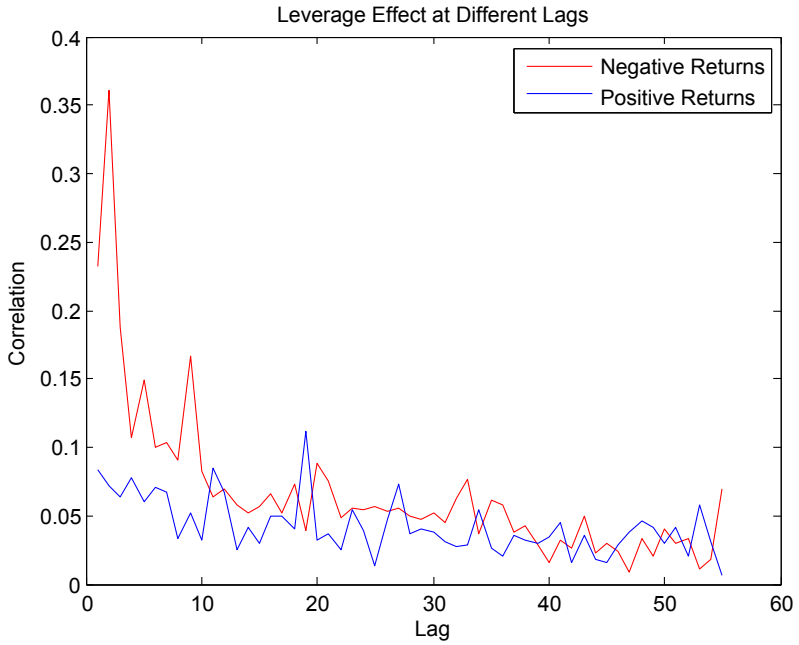

Figure 3.2: Leverage effect for S&P500 Index Futures returns. In this figure, the correlation functions of for negative/positive returns versus lagged volatilities are computed in absolute value.

or more commonly as the "smirk". This skew reflects the fact that a negative volatility-return correlation induces a negative skew in the distribution of returns

themselves.

Finally, we want to address to the commonly observed sudden sharp changes in the evolution of asset prices known as jumps. For years, the typical way of describing the dynamics of returns of financial securities has been the Brownian motion, basically because of its simplicity, intuitiveness and tractability. Nevertheless, a graphical comparison between a real returns series and Gaussian random variable with same mean and standard deviation poses questions on the goodness of the assumption (see Figure 3.3). In fact, by counting the observed returns above the

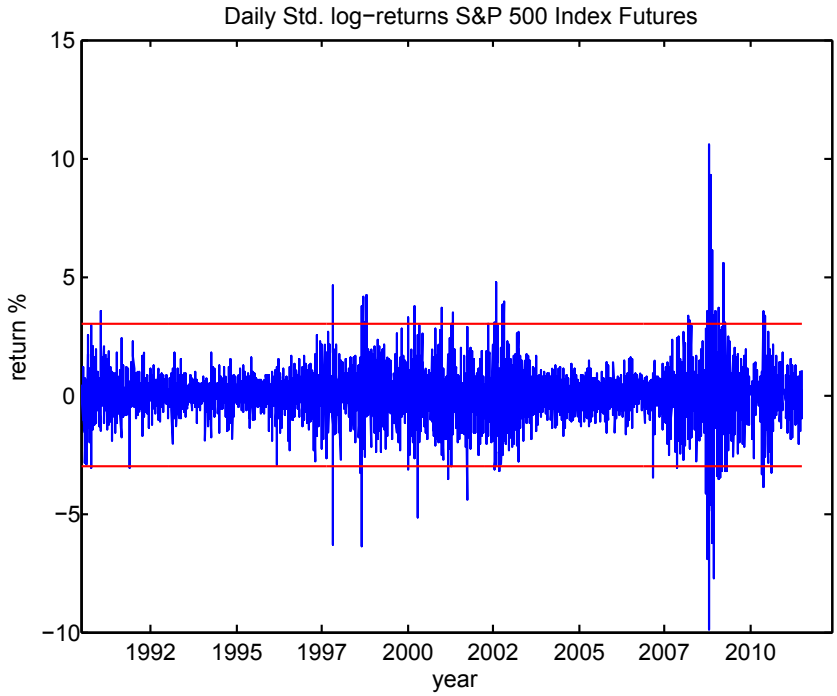

Figure 3.3: Series of daily returns for S&P500 Index Futures returns from July 1990 to June 2011. Red lines identify the three standard deviations confidence interval for a Gaussian random variable with the same mean and variance.

range of three sample standard deviations, we note that the theoretical standard normal probability of observing such events (0.27%) is at odd with the realized frequency of $80/5237 \approx 1.53\%$ occurred in the sample. Typically, this is the basic argument criticizing the assumption of the standard Brownian motion as a plausible model for returns dynamics, although not excluding the possibility of a more complex diffusive behaviour *based* on Brownian motion. In fact, as observed by Bibby and Sørensen (2003), an appropriate choice of a nonlinear time-dependent diffusion coefficient could generate diffusion processes with arbitrary heavy tails.

This is the case, for instance, of stochastic volatility models.

A more stringent argument against diffusive models is suggested by observing returns at finer temporal scale. For example, it is not uncommon to find the behaviour reported in Figure 3.4 for a series of intraday 5-min returns which shows a sudden drop/peak corresponding to a jump in the price. This behaviour can be

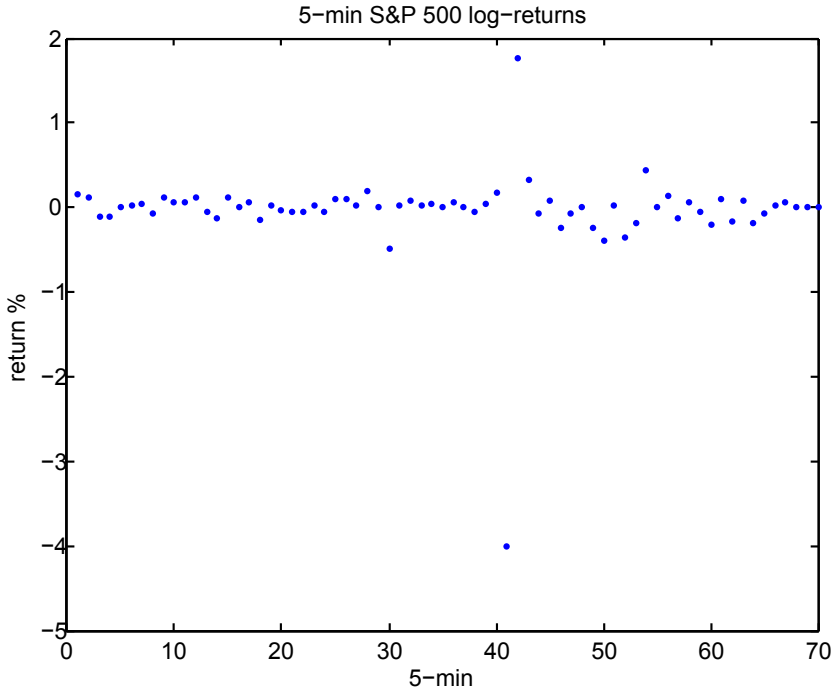

Figure 3.4: Series of 5-minute intraday returns for S&P500 Index Futures on 8 January 1991.

difficultly ascribed to a diffusive dynamics since independently by its complexity it implies the continuity of its trajectory when the temporal scale tends to the continuous limit. Indeed, these abrupt changes suggests the presence of a purely discontinuous component of the dynamics of asset prices which adds to the continuous diffusion. Furthermore several studies indicate that an unexpected jump event triggers subsequent jumps, a phenomenon known as jump clustering (see Christoffersen et al. (2012); Chen and Poon (2013); Aït-Sahalia et al. (2015); Bormetti et al. (2015a)).

The origin of jumps is typically attributed to arrivals of significant information such as macroeconomic news, surprises in corporate earnings and swings in monetary policy or government spending. Jumps represent a major source of risk for market agents in particular for those operating with short-term horizons. Nowa-

days, the request for models capable to account for them and the risk-premium asked for compensation of sudden extreme events is central in the financial debate.

3.2 Motivating idea of heterogeneous autoregressive processes

As previously said, long memory is an important property that recent volatility models are designed to capture. ARFIMA processes are the typical example of this class. Using realized volatility, ARFIMA specification provides a test for long memory allowing to compute explicitly the variance of integrated process

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

where β is the decay exponent of the autocorrelation function in (3.1). The test indicates the presence of long memory with β parameter if the empirically computed $V(\Delta)$ behaves like $\Delta^{2-\beta}$. However, it has been shown the existence of process satisfying the above long memory test which, contrarily, have not an autocorrelation function with power law decay (as reported in LeBaron (2001), Corsi (2009), Gatheral et al. (2014)). These ambiguities have kept open the question whether financial volatility process is a real long memory process or it just resembles one. Furthermore, doubts has raised about the use of sophisticated long memory models that are difficult to be economically interpreted and estimated, encouraging to build simpler additive process that could mimic the demanded property.

A relevant instance of a parsimonious process not formally belonging to the class of long memory model but able to reproduce the volatility persistence observed in real data is the Heterogeneous Autoregressive process (HAR) introduced by Corsi (2009). The motivating idea is inspired by the so-called Heterogeneous Market Hypothesis formulated by Muller et al. (1997), which is proposed as a possible explanation of the empirically observed phenomenon that volatility defined over a coarse time grid significantly predicts volatility defined over a finer time grid. Market heterogeneity refers in widely sense to the differences between agents'

perceptions of the market, risk profiles, endowments, degree of information, prior beliefs and other characteristics. It is argued that many of these differences among market participants translate into a sensitivity to different time horizons. In fact, a financial market consists of participants having different investment frequencies. These form a wide spectrum going from high intraday frequencies (such as for dealers and high frequency traders) to low trading frequencies (such as for institutional investors, pension funds or insurance companies). The aggregation of agents' perceptions and reactions to events on financial market with different trading frequencies cause different volatility components that contribute to the formation of the total market volatility.

According to this view, the asymmetric relation between volatilities over longer time intervals and shorter time intervals can be economically explained by observing that short-term traders are interested in the level of long-term volatility because they form expectations about future size of trends and risk. On the other hand, the short-term volatility does not influence the trading strategies of long-term agents with low trading frequencies.

HAR-RV model is set in order to reflect this volatility cascade by introducing a hierarchical dependence of present realized volatility from three volatility components corresponding to past time horizons of one day (short-term), one week (medium-term) and one month (long-term). These components are constructed as the mean of the daily realized variances collected over the past day, week and month, respectively. The model is formally expressed as

$$\mathrm{RV}_t = d + \beta_d \mathrm{RV}_{t-1}^{(d)} + \beta_w \mathrm{RV}_{t-1}^{(w)} + \beta_d \mathrm{RV}_{t-1}^{(m)} + w_t$$

where d is a real constant, w_t is an error term and

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

The remarkable result about HAR specification is the ability to reproduce the volatility persistence, along with other stylized facts, while theoretically having the characteristics of a short-memory process, which asymptotically should not

exhibit hyperbolic decay of the autocorrelation function. Besides, in spite of the simplicity of its structure and estimation (based on simple linear regression), the HAR model shows a good forecasting performance comparable to the much more complicated and troublesome to estimate long-memory ARFIMA (see also Corsi et al. (2012) for more results and HAR extensions).
Recently, Corsi et al. (2013) have proposed a stochastic volatility model for realized volatility in which the source of randomness is described by a non-central gamma distribution whose non-centrality term is driven by an HAR process. This model has been named as Heterogeneous AutoRegressive Gamma or HARG and represents a starting point of the modelling settings we are going to introduce.

3.3 JLHARG-RV models

3.3.1 Physical dynamics

We consider a risky asset with the following log-return dynamics

$$y_t = r + \left(\lambda - \frac{1}{2}\right) \mathrm{RV}_t + \sqrt{\mathrm{RV}_t}\epsilon_t, \tag{3.2}$$

where r is the risk-free rate, λ is the market price of risk, ϵ_t are *i.i.d.* standard normal innovations, and RV_t is realized variance at day t. Here the realized variance is intended as the total realized variance given by the sum of two separate components: a continuous component CRV_t and a jump component JRV_t. This approach is motivated by the empirical analyses of Andersen et al. (2001), 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 $y_t/\sqrt{\mathrm{RV}_t}$ standardized returns.[1]

[1] "*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*

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 3.5, we observe the same feature for the S&P500 Futures in our sampling period. The two-sample Kolmogorov-Smirnov test between

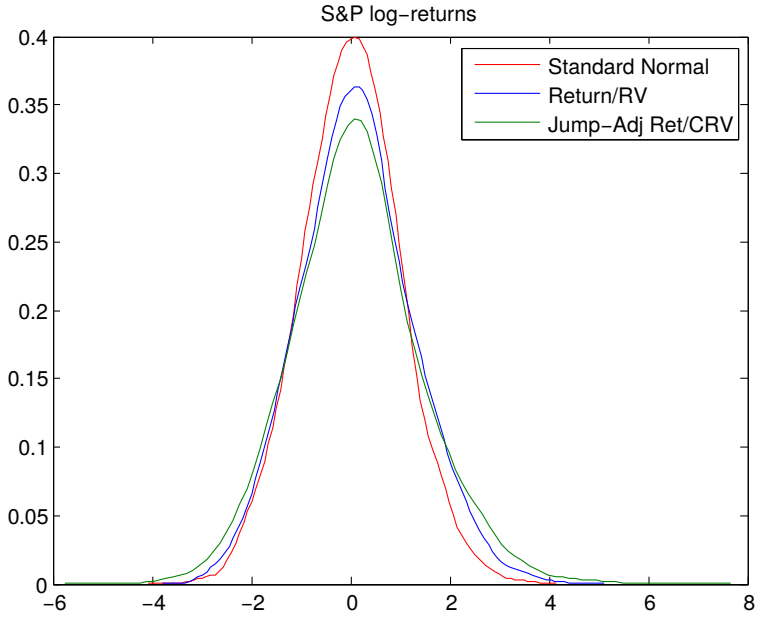

Figure 3.5: Standardized log-return distribution. Comparison of the S&P500 Futures log-return distribution under different rescaling measures: Standard normal distribution (red line), jump-adjusted standardized log-return by CRV (green line) and standardized log-return by total RV (blue line).

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. Consequently, we employ the log-return dynamics

$$y_t = r + \left(\lambda - \frac{1}{2}\right)(\mathrm{CRV}_t + \mathrm{JRV}_t) + \sqrt{\mathrm{CRV}_t + \mathrm{JRV}_t}\epsilon_t \tag{3.3}$$

as a reasonable approximation of the process observed at daily frequency.

Given the information at time t, $\mathcal{F}_t$, a new realization of the RV components

standardized returns, so the impact is muted." - Andersen et al. (2010)

is obtained by sampling at time $t+1$ from two conditionally independent distributions. The continuous part of RV depends on past realizations of CRV and of a leverage component ℓ_t, which corresponds to a quadratic function of the total realized variance

$$\ell_t = \left(\epsilon_t - \gamma\sqrt{\mathrm{CRV}_t + \mathrm{JRV}_t}\right)^2 .$$

Then, introducing the notation $\mathbf{CRV}_t = (\mathrm{CRV}_{t-21}, \ldots, \mathrm{CRV}_t)$ and $\mathbf{L}_t = (\ell_{t-21}, \ldots, \ell_t)$, the continuous component of RV is drawn from a noncentred gamma distribution

$$\mathrm{CRV}_{t+1}|\mathcal{F}_t \sim \bar{\gamma}(\delta, \Theta(\mathbf{CRV}_t, \mathbf{L}_t), \theta)\,, \tag{3.4}$$

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

$$\Theta(\mathbf{CRV}_t, \mathbf{L}_t) = d + \beta_d \mathrm{CRV}_t^{(d)} + \beta_w \mathrm{CRV}_t^{(w)} + \beta_m \mathrm{CRV}_t^{(m)} + \alpha_d \ell_t^{(d)} + \alpha_w \ell_t^{(w)} + \alpha_m \ell_t^{(m)}\,, \tag{3.5}$$

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

$$\begin{aligned}
&\mathrm{CRV}_t^{(d)} = \mathrm{CRV}_t, && \ell_t^{(d)} = \ell_t\,,\\
&\mathrm{CRV}_t^{(w)} = \tfrac{1}{4}\textstyle\sum_{i=1}^{4} \mathrm{CRV}_{t-i}, && \ell_t^{(w)} = \tfrac{1}{4}\textstyle\sum_{i=1}^{4} \ell_{t-i}\,,\\
&\mathrm{CRV}_t^{(m)} = \tfrac{1}{17}\textstyle\sum_{i=5}^{21} \mathrm{CRV}_{t-i}, && \ell_t^{(m)} = \tfrac{1}{17}\textstyle\sum_{i=5}^{21} \ell_{t-i}
\end{aligned}$$

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.

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 $\tilde{\theta}$

$$\mathrm{JRV}_{t+1}|\mathcal{F}_t \sim \sum_{i=1}^{n_{t+1}} Y_i \quad \text{with} \quad n_{t+1} \sim \mathcal{P}(\tilde{\Theta}) \quad \text{and} \quad Y_i \; i.i.d. \sim \gamma(\tilde{\delta}, \tilde{\theta}). \tag{3.6}$$

Equations (3.3)-(3.6) completely characterise the log-return dynamics as an Au-

toregressive Gamma model in Realized Volatility with Heterogeneous Leverage and Jumps, and we acronym it JLHARG-RV model. The crucial advantage of the JLHARG model is that it satisfies the affine property. The importance of affine processes in finance - due to their analytical tractability - has been acknowledged in many studies (see Duffie et al. (2000); Darolles et al. (2006); Majewski et al. (2015) among others). We prove the following

Proposition 4. *Under* $\mathbb{P}$*, the MGF of the log-return* $y_{t,T} = \sum_{k=t+1}^{T} y_k$ *for JL-HARG model has the following form*

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

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

Proof: See Appendix A.1.

3.3.2 Risk-neutral dynamics

According to the contents of Chapter 2, the main ingredient to have a discrete-time model suited for asset pricing is the choice of a stochastic discount factor allowing the passage from the physical measure to the martingale or risk-neutral measure. Following this purpose, we risk-neutralize the process employing an SDF within the family of exponential-affine factors, whose high flexibility allows to incorporate multiple factor-dependent risk premia. This approach has been extensively used in literature (for example in Gagliardini et al. (2011), Corsi et al. (2013), Christoffersen et al. (2013) and Majewski et al. (2015)). We propose an SDF of the following form

$$M_{s,s+1} = \frac{e^{-\nu_c \mathrm{CRV}_{s+1} - \nu_j \mathrm{JRV}_{s+1} - \nu_y y_{s+1}}}{\mathbb{E}^{\mathbb{P}}\left[e^{-\nu_c \mathrm{CRV}_{s+1} - \nu_j \mathrm{JRV}_{s+1} - \nu_y y_{s+1}}|\mathcal{F}_s\right]}, \tag{3.7}$$

which represents the Esscher transform from the physical log-return density to the risk neutral one. An important characteristic of the SDF (3.7) is to clearly identify the sources of risk and explicitly compensate them with separated risk premia.

Specifically, this form allows to have both the continuous (ν_c) and discontinuous (ν_j) variance risk premia, in addition to the standard equity premium (ν_y). The equity premium has to satisfy the following no-arbitrage condition.

Proposition 5. *The JLHARG model defined by equations (3.3) – (3.6) with SDF given by (3.7) satisfies the no-arbitrage condition if and only if*

$$\nu_y = \lambda\,.$$

Proof: Appendix A.2.

Another significant consequence of adopting the SDF (3.7) is the fact that it allows to preserve analytical tractability of the model under the martingale measure. Moreover, we are able to provide a one-to-one mapping of the parameters under $\mathbb{P}$ to those under the $\mathbb{Q}$ measure, ensuring that the risk-neutral log-return dynamics is still governed by a JLHARG process according to the following proposition.

Proposition 6. *Under risk-neutral measure $\mathbb{Q}$ the realized variance follows a JLHARG process with parameters*

$$\begin{aligned}
&\beta_d^* = \frac{\beta_d}{1-\theta \mathrm{y}^{c*}},\ \beta_w^* = \frac{\beta_w}{1-\theta \mathrm{y}^{c*}},\ \beta_m^* = \frac{\beta_m}{1-\theta \mathrm{y}^{c*}},\\
&\alpha_d^* = \frac{\alpha_d}{1-\theta \mathrm{y}^{c*}},\ \alpha_w^* = \frac{\alpha_w}{1-\theta \mathrm{y}^{c*}},\ \alpha_m^* = \frac{\alpha_m}{1-\theta \mathrm{y}^{c*}},\\
&\theta^* = \frac{\theta}{1-\theta \mathrm{y}^{c*}},\ \delta^* = \delta,\ \gamma^* = \gamma + \lambda,\\
&d^* = \frac{d}{1-\theta \mathrm{y}^{c*}}\\
&\tilde{\Theta}^* = \frac{\tilde{\Theta}}{\left(1-\tilde{\theta} \mathrm{y}^{j*}\right)^{\tilde{\delta}}},\ \tilde{\delta}^* = \tilde{\delta},\ \tilde{\theta}^* = \frac{\tilde{\theta}}{1-\tilde{\theta} \mathrm{y}^{j*}},
\end{aligned} \tag{3.8}$$

where $\mathrm{y}^{c*} = -\lambda^2/2 - \nu_c$ *and* $\mathrm{y}^{j*} = -\lambda^2/2 - \nu_j$.

Proof: Appendix A.3.

Knowing the dynamics of the process under $\mathbb{Q}$, the moment generating function under the risk-neutral measure is a straightforward consequence of Proposition 4.

Corollary 3. *Under $\mathbb{Q}$ the MGF of the JLHARG model is formally the same as in Proposition 4 with equity risk premium $\lambda^* = 0$, and d^*, δ^*, θ^*, $\tilde{\Theta}^*$, $\tilde{\delta}^*$, $\tilde{\theta}^*$ γ^*, α_l^*, β_l^* for $l = d, w, m$ as in (3.8).*

We point out that the risk premia in the vector (ν_c, ν_j) are the only parameters that need to be calibrated on option data. Then, all the parameters governing the dynamics of the process under $\mathbb{Q}$ can be explicitly computed from the values estimated under $\mathbb{P}$ starting from the values of (ν_c, ν_j) through the relations given by (3.8).

3.3.3 Nested models

The JLHARG-RV family nests a variety of RV models as special cases. The first instance is the JHARG model which preserves the heterogeneous autoregressive structure for RV but lacks the leverage term. This model can be seen as a natural extension of the HARG model, by Corsi et al. (2013), accounting for a discontinuous component. The second model is the JLHARG model with Parabolic Leverage (P-JLHARG) that we obtain setting $d = 0$ in (3.5). The third one is a JLHARG with zero-mean leverage term (ZM-JLHARG) inspired by the Component GARCH model of Christoffersen et al. (2008). At variance with the latter paper, we follow the approach of Majewski et al. (2015) including time heterogeneity in leverage through the following relations

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

The linear $\Theta(\mathbf{CRV}_t, \mathbf{L}_t)$ in this case reads

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

which can be reduced to the form (3.5) setting $d = -(\alpha_d + \alpha_w + \alpha_m)$, $\beta_l = \beta_l - \alpha_l \gamma^2$ for $l = d, w, m$. The larger flexibility of the leverage term $\bar{\ell}_t$ allows the model to better describe the skewness and kurtosis of the empirical data.

3.3.4 Model estimation and statistical properties

Belonging to the class of the RV models, a relevant advantage of JLHARG-RV models is the ease of the parameters estimation under $\mathbb{P}$ measure. This derives from the direct observability of the realized volatility which avoids the use of filtering procedure of the latent volatility process. So the general scheme of an estimation procedure (starting with the empirical computation of RV series) for this class of models can be summarized by the following passages:

(i) to build a realized volatility series from observed returns series choosing one of the many available methods in literature (see Section 1.3);

(ii) to identify the discontinuous component using a jump detection test (see Section 1.4) which detects the spikes in RV time series and separates it from the continuous component.

After having determined the two time series for the RV components, the estimation of the parameters of the JLHARG-RV processes is determined using the Maximum Likelihood Estimator. According to the model specified in equation (3.4) and (3.6), 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

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

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

Both log-likelihoods have a term involving an infinite series. To overcome this issue a truncation of the infinite sum can be operated to an arbitrary high term. In order to reduce the dimensionality of the space of parameters, two of them can be fixed by variance targeting with the aim to obtain the exact match of the observed sample mean of the realized variance. For example, by choosing as target parameters δ and $\tilde{\delta}$, they can be computed using the following expressions for the unconditional mean of CRV_t and JRV_t implied by the model, respectively

$$\mathbb{E}\left[\mathrm{CRV}_t\right] = \frac{\theta\delta}{1-\theta\left(\sum_{i=d,w,m}\beta_i\right)} \quad \text{for JHARG and ZM-JLHARG} \tag{3.9}$$

$$\mathbb{E}\left[\mathrm{CRV}_t\right] = \frac{\theta\left(\delta+\sum_{i=d,w,m}\alpha_i\left(1+\gamma^2\mathbb{E}\left[\mathrm{JRV}_t\right]\right)\right)}{1-\theta\left(\sum_{i=d,w,m}\beta_i+\gamma^2\sum_{i=d,w,m}\alpha_i\right)} \quad \text{for P-JLHARG,} \tag{3.10}$$

$$\mathbb{E}\left[\mathrm{JRV}_t\right] = \tilde{\delta}\tilde{\Theta}\tilde{\theta}. \tag{3.11}$$

Finally, the market price of risk λ in equation (3.3) is determined by regressing the centered and normalized log-return on the realized volatility. This regression is performed by rewriting the equation (3.3) as

$$\frac{y_{t+1}-r}{\sqrt{\mathrm{CRV}_{t+1}+\mathrm{JRV}_{t+1}}} = \left(\lambda-\frac{1}{2}\right)\sqrt{\mathrm{CRV}_{t+1}+\mathrm{JRV}_{t+1}}+\epsilon_{t+1}\,.$$

In view of an option pricing application, an important role for reproducing the shape of the implied volatility surface is played by the skewness and kurtosis term structures generated by the underlying dynamics. Adding a heterogeneous leverage considerably improves the skewness and the excess kurtosis of the log-return probability distribution. Within our framework, we not only preserve the heterogeneity of the leverage, but we also add a discontinuous component which captures extreme price movements. With this choice, our JLHARG class of models is able to reproduce a stronger leverage effect. In Figure 3.6 we show the skewness and the excess kurtosis from a simulation of the P-JLHARG model with realistic values of parameters at different aggregation time – from one 1 day to 250 days – under both $\mathbb{P}$ and $\mathbb{Q}$ measures. The model is able to reproduce significant negative values of skewness and positive excess kurtosis under the physical measure. When moving to the $\mathbb{Q}$ measure, the effect, induced by the presence of the variance risk premia ν_c and ν_j, is strengthened.

3.4 LHARG-ARJ-RV model

3.4.1 Physical dynamics

The model that we are going to present in this section has been introduced to account for dynamics of extreme events with time-varying jump intensity. In order to include the empirical properties observed in returns and volatility time series of jumps clustering (see Lee and Mykland (2008)), the dynamics of jumps' intensity is based on a GARCH-type approach with autoregressive structure and an innovation contribution expressed in terms of past realized number of jumps. Here we describe the dynamics of log-returns of a risky asset on daily scale by the following,

$$y_t = r + \left(\lambda_c - \frac{1}{2}\right) \mathrm{CRV}_t + (\lambda_j - \eta)\left(\Lambda^2 + \delta^2\right) n_t + \sqrt{\mathrm{CRV}_t}\epsilon_t + \sum_{i=1}^{n_t} X_{t,i}, \quad (3.12)$$

where r is a risk-free rate in the economy, λ_c and λ_j are continuous and jump, respectively, components of equity risk premium, CRV_t is the continuous component of realized variance representing the variation due to diffusion with ϵ_t i.i.d.

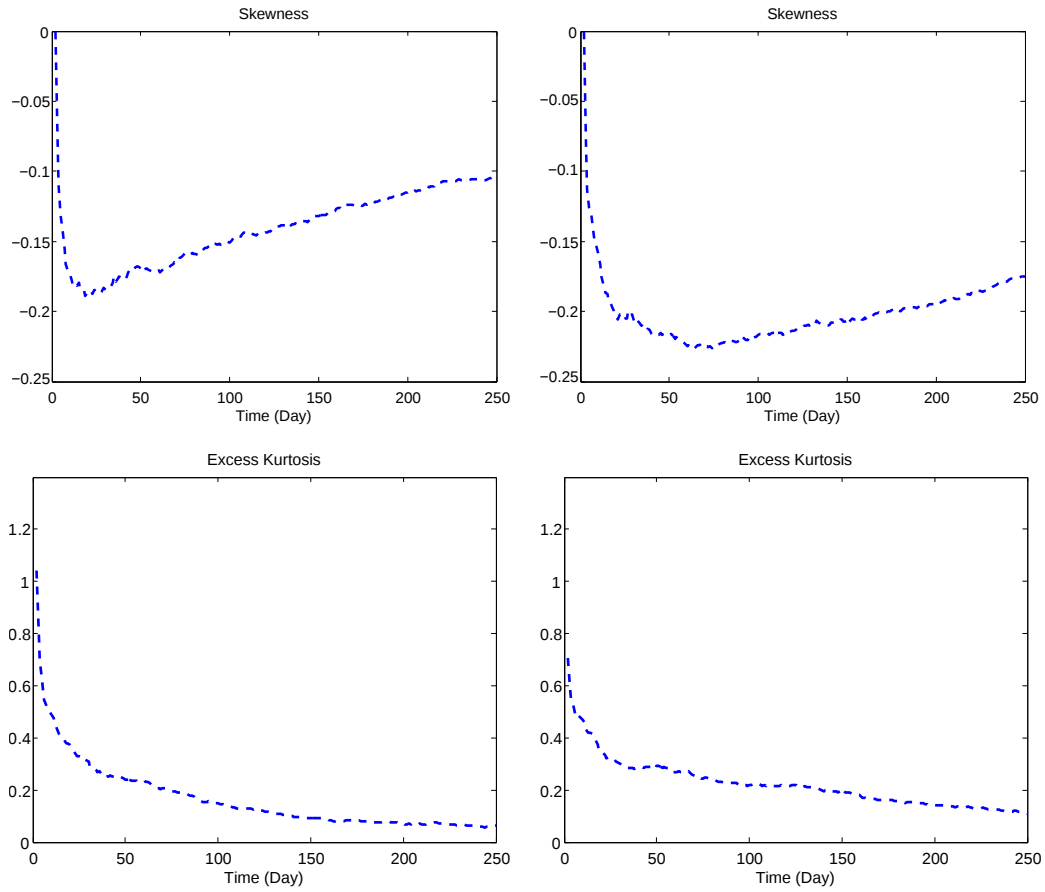

Figure 3.6: Skewness and excess kurtosis of the JLHARG process under physical measure $\mathbb{P}$ (left column) and risk-neutral measure $\mathbb{Q}$ (right column).

standard normal random variables, n_t is the total number of jumps occurring at day t and $X_{t,i}$ are i.i.d. random variables capturing the sizes and directions of the jumps. We assume that $X_{t,i}$ are sampled from a standard normal distribution $\mathcal{N}(\Lambda, \delta^2)$ where $\Lambda \in \mathbb{R}$ and $\delta \in \mathbb{R}^+$ are parameters specifying the mean and standard deviation of the jump size, respectively. The continuous component of realized variance, CRV_t is modelled conditioned on information at day t as a non-centred gamma random variable distributed according to

$$\mathrm{CRV}_t|\mathcal{F}_t \sim \bar{\gamma}(\kappa, \Theta(\mathbf{CRV}_{t-1}, \mathbf{L}_{t-1}), \theta) \tag{3.13}$$

with

$$\Theta(\mathbf{CRV}_t, \mathbf{L}_t) = \beta_d \mathrm{CRV}_t^{(d)} + \beta_w \mathrm{CRV}_t^{(w)} + \beta_m \mathrm{CRV}_t^{(m)} + \alpha_d \ell_t^{(d)} + \alpha_w \ell_t^{(w)} + \alpha_m \ell_t^{(m)}. \tag{3.14}$$

In the previous equation, the quantities

$$\begin{array}{ll} \mathrm{CRV}_t^{(d)} = \mathrm{CRV}_t, & \ell_t^{(d)} = \left(\epsilon_t - \gamma\sqrt{\mathrm{CRV}_t}\right)^2 - 1 - \gamma^2 \mathrm{CRV}_t, \\ \mathrm{CRV}_t^{(w)} = \frac{1}{4}\sum_{i=1}^{4} \mathrm{CRV}_{t-i}, & \ell_t^{(w)} = \frac{1}{4}\sum_{i=1}^{4}\left[\left(\epsilon_{t-i} - \gamma\sqrt{\mathrm{CRV}_{t-i}}\right)^2 - 1 - \gamma^2 \mathrm{CRV}_{t-i}\right], \\ \mathrm{CRV}_t^{(m)} = \frac{1}{17}\sum_{i=5}^{21} \mathrm{CRV}_{t-i}, & \ell_t^{(m)} = \frac{1}{17}\sum_{i=5}^{21}\left[\left(\epsilon_{t-i} - \gamma\sqrt{\mathrm{CRV}_{t-i}}\right)^2 - 1 - \gamma^2 \mathrm{CRV}_{t-i}\right], \end{array} \tag{3.15}$$

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 jump component of variance, JRV_{t+1}, defined as follows

$$\mathrm{JRV}_t = \sum_{i=1}^{n_t} X_{t,i}^2 \tag{3.16}$$

is naturally linked with the sum of squared daily jump variation which we model

as a compound Poisson variable with i.i.d. normally distributed jumps

$$n_t|\omega_t \sim \text{Poisson}\,(\omega_t) \quad \text{and} \quad X_{t,i} \sim \mathcal{N}(\Lambda, \delta^2). \tag{3.17}$$

The number of jumps depends on time-varying intensity ω_t whose dynamics is described by autoregressive process

$$\omega_{t+1} = \bar{\omega} + \xi\omega_t + \zeta n_t. \tag{3.18}$$

where $\bar{\omega}, \xi, \zeta \in \mathbb{R}^+$ are parameters of the model. As follows from equation (3.18) jump intensity on day $t+1$ depends on the intensity day before (ω_t) and on the number of intra-day jumps that occurred the day before (n_t). This structure allows the possibility that an extreme event can increase locally the intensity of the jump process in the following days reproducing the empirical clustering of jump events.

We observe that the parameter η in (3.12) has to be interpreted as a jump compensator ensuring that the conditionally expected total return is given by

$$\mathbb{E}\left[\exp\left(y_{t+1}\right)|\text{CRV}_{t+1}, n_{t+1}\right] = \exp(r + \lambda_c \text{CRV}_{t+1} + \lambda_j\left(\Lambda^2 + \delta^2\right) n_{t+1}). \tag{3.19}$$

This condition imposes that η is a particular function of other jump parameters expressed as

$$\eta = \frac{\Lambda + \frac{1}{2}\delta^2}{\Lambda^2 + \delta^2},$$

which guarantees that parameters λ_c and λ_j are interpreted as of two equity premia compensating for diffusive volatility and jump exposure, respectively. The complete model described by equations (3.12)-(3.18), with LHARG dynamics of continuous component of RV and Auto-Regressive in Jump is labelled LHARG-ARJ. This model still has the characteristic to satisfy the affine property which guarantees the availability of moment generating function in an exponential-affine form.

Proposition 7. *The MGF of log-return* $y_{t,T} = \sum_{i=0}^{T-1} y_{t+i}$ *in LHARG-ARJ model*

conditioned on the information available at time t is of the form

$$\mathbb{E}^{\mathbb{P}}\left[e^{zy_{t,T}}|\mathcal{F}_t\right] = e^{\mathrm{a}_t+\sum_{i=1}^{22}\mathrm{b}_{t,i}\mathrm{CRV}_{t+1-i}+\sum_{j=1}^{22}\mathrm{c}_{t,j}\ell_{t+1-j}+\mathrm{d}_{t+1}(\xi\omega_t+\zeta n_t)}. \tag{3.20}$$

with recursive relation

$$\begin{aligned}
\mathrm{a}_s &= \mathrm{a}_{s+1} + zr - \frac{1}{2}\ln(1-2c_{s+1,1}) - \kappa\mathcal{W}\left(x_{s+1},\theta\right) + d\mathcal{V}\left(x_{s+1},\theta\right) + \mathrm{d}_{s+1}\bar{\omega},\\
\mathrm{b}_{s,i} &= \begin{cases} \mathrm{b}_{s+1,i} + \beta_i\mathcal{V}\left(x_{s+1},\theta\right) & \text{for } 1 \le i < 22 \\ \beta_i\mathcal{V}\left(x_{s+1},\theta\right) & \text{for } i = 22 \end{cases}\\
\mathrm{c}_{s,j} &= \begin{cases} \mathrm{c}_{s+1,j} + \alpha_j\mathcal{V}\left(x_{s+1},\theta\right) & \text{for } 1 \le j < 22 \\ \alpha_j\mathcal{V}\left(x_{s+1},\theta\right) & \text{for } j = 22 \end{cases}\\
\mathrm{d}_s &= e^{v+\zeta\mathrm{d}_{s+1}} - 1 + \xi\mathrm{d}_{s+1}.
\end{aligned} \tag{3.21}$$

where

$$x_{s+1} = z\lambda_c + b_{s+1,1} + \frac{\frac{z^2}{2} + \gamma^2 c_{s+1,1} - 2c_{s+1,1}\gamma z}{1-2c_{s+1,1}}, \tag{3.22}$$

$$v = \frac{z^2\delta^2}{2} + z\Lambda + z\left(\lambda_j - \eta\right)\left(\Lambda^2+\delta^2\right), \tag{3.23}$$

and

$$\mathcal{V}(x,\theta) = \frac{\theta x}{1-\theta x}, \quad \mathcal{W}(x,\theta) = \ln(1-x\theta), \tag{3.24}$$

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

Proof: See Appendix B.1.

3.4.2 Risk-neutral dynamics

The risk-neutralization of the LHARG-ARJ model is obtained in a similar fashion as JLHARG models. In this case, we propose to apply a four-dimensional Esscher transform of the type

$$M_{t-1,t} = \frac{\mathrm{e}^{-\nu_c \mathrm{CRV}_t - \nu_j \mathrm{JRV}_t - \mu_c \sqrt{\mathrm{CRV}_t}\epsilon_t - \mu_j \sum_{i=1}^{n_t} X_{t,i}}}{\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{-\nu_c \mathrm{CRV}_t - \nu_j \mathrm{JRV}_t - \mu_c \sqrt{\mathrm{CRV}_t}\epsilon_t - \mu_j \sum_{i=1}^{n_t} X_{t,i}} \middle| \mathcal{F}_{t-1}\right]}. \tag{3.25}$$

with four parameters responsible for different risk premia. The value of two first parameters, ν_c and ν_j, determines the level of variance risk premium. The following two parameters, μ_c and μ_j, correspond to the level of equity risk premium. Component $\mu_c\sqrt{\mathrm{CRV}_t}\epsilon_t$ takes into account the risk related to continuous directional changes in price, whereas component $\mu_j \sum_{i=1}^{n_t} X_{t,i}$ captures the risk related to abrupt and large changes in price.

Transform (3.25) has to guarantee absence of arbitrage opportunity in order to be a stochastic discount factor in our model.

Proposition 8. *If the dynamics of the underlying price is described by (3.12)-(3.18), then the Esscher transform (3.25) is a stochastic discount factor if, and only if the following conditions are satisfied*

$$\mu_c = \lambda_c \quad \textit{and} \quad \mu_j = \frac{1}{2} + \frac{\Lambda + (\lambda_j - \eta)\left(\Lambda^2 + \delta^2\right)\left(1 + 2\nu_j\delta^2\right)}{\delta^2}. \tag{3.26}$$

Proof: See Appendix B.2.

From relation (3.26) one can see that the no-arbitrage condition fixes the value of the parameters μ_c and μ_j, while parameters ν_c and ν_j remain free and they have 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. Performing the change of measure by means of the Esscher transform (3.25) provides the risk-neutral dynamics which governs the return process.

Proposition 9. *Under the risk-neutral measure $\mathbb{Q}$ - corresponding to the SDF specification given by (3.25) - the log-return dynamics for LHARG-ARJ model is governed by equations (3.12)-(3.18) with parameters λ_c^*, λ_j^*, κ^*, θ^*, β_d^*, β_w^*, β_m^*,*

α_d^*, α_w^*, α_m^*, γ^*, Λ^*, δ^*, $\bar{\omega}^*$, ξ^* and ζ^*. The mapping among the starred and the physical parameters is provided in Appendix B.3 (equations (46) and (50)).

Proof: See Appendix B.3.

Given the dynamics under $\mathbb{Q}$, the risk-neutral moment generating function is a straightforward consequence of Proposition 7.

Corollary 4. *Under $\mathbb{Q}$, the MGF for the LHARG-ARJ model has the same form as in (3.20)-(3.24) with parameters as in (46) and (50).*

The existence of analytic expression for MGF under risk-neutral measure allows to proceed variance risk premium estimation with very efficient and reliable numerical methods based on Fourier transform.

3.4.3 Model estimation

The estimation procedure for LHARG-ARJ-RV model inherits the common feature among models based on realized volatility: straightforwardness. It basically follows the general scheme described in Section 4.1.2, consisting in building the realized volatility series from observed returns and successively disentangle the continuous and discontinuous components involving a jump detection test. Nevertheless, the majority of the test statistics to identify jumps in the observed returns series are based on a daily scale giving information if at least one extreme event occurred during the trading day. As can be seen from equation (3.18), the jump dynamics is completely specified at day t if the total number of jumps at day $t-1$ is known. So the choice of the jump detection method has to be done among that ones allowing in addition the identification of the total intra-day discontinuities.

After these passages, the estimation of the parameters proceeds involving the Maximum Likelihood Estimator. According to the model, the log-likelihood function is simply given by the sum of three contributions one related to the daily log-return process and two related to the continuous and discontinuous contribution to the realized variance process. As can be seen from equation (3.12) the jump-adjusted log-return variable $\tilde{y}_t = y_t - \sum_{i=1}^{N(t)} X_i$ is distributed normally with

$\mathcal{N}\left(r+\left(\lambda_c-\frac{1}{2}\right)\mathrm{CRV}_t+(\lambda_j-\eta)\left(\Lambda^2+\delta^2\right)n_t,\mathrm{CRV}_t\right)$ conditionally to CRV_t and n_t. Its log-likelihood contribution is expressed by the following,

$$L^r\left(\lambda_c,\lambda_j,\Lambda,\delta\right)=-\sum_{t=1}^{T}\left[\frac{\left(\tilde{y}_t-\left(r+\left(\lambda_c-\frac{1}{2}\right)\mathrm{CRV}_t+(\lambda_j-\eta)\left(\Lambda^2+\delta^2\right)n_t\right)\right)^2}{2\mathrm{CRV}_t}+\log\left(\sqrt{2\pi\mathrm{CRV}_t}\right)\right]. \tag{3.27}$$

As concerns the continuous component of realized variance, we model it as a random variable sampled from the non-central gamma distribution $\bar{\gamma}(\kappa,\Theta(\mathbf{CRV}_t,\mathbf{L}_t),\theta)$ as described in Section 3.4.1. The corresponding log-likelihood has the form

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

The last log-likelihood contribution takes account of the jump component of realized variance, which, as defined in equation (3.16), is distributed according to a non-central chi-squared conditionally to the number of observed jumps n_t. Since the variable n_t is Poisson distributed, the log-likelihood is given by the following

$$L^{\mathrm{JRV}}\left(\bar{\omega},\xi,\zeta,\Lambda,\delta\right)=\sum_{t=1}^{T}\log\left(e^{-\omega_t}\frac{\omega_t^{n_t}}{n_t!}\sum_{m=0}^{\infty}e^{-\frac{n_t\Lambda^2}{2\delta^2}}\frac{\left(\frac{n_t\Lambda^2}{2\delta^2}\right)^m}{m!}\frac{\mathrm{JRV}_t^{\frac{n_t+2m}{2}-1}e^{-\frac{\mathrm{JRV}_t}{2\delta^2}}}{\Gamma\left(\frac{n_t+2m}{2}\right)\left(2\delta^2\right)^{\frac{n_t+2m}{2}}}\right). \tag{3.28}$$

Finally, the estimation of the total set of parameters characterizing the LHARG-ARJ process is realized by maximizing the whole log-likelihood function $L=L^r+L^{\mathrm{CRV}}+L^{\mathrm{JRV}}$. We observe that ω_t, in the expression (3.28), is a latent process. So the optimization of the L function includes a built-in filtering procedure in which at every step the ω_t series is obtained as a function of the parameters given the set of observed historical values of n_t according to equation (3.18).

In order to reduce the dimensionality of the space of parameters, we fix two of

them by variance targeting with the aim to obtain the exact match of the observed sample mean of the realized variance. We choose as target parameters κ and $\bar{\omega}$ and compute them using the following expressions for the unconditional mean of CRV_t and JRV_t, respectively

$$\mathbb{E}\left[\mathrm{CRV}_t\right] = \frac{\theta\kappa}{1-\theta\left(\sum\limits_{i=d,w,m}\beta_i\right)} \tag{3.29}$$

$$\mathbb{E}\left[\mathrm{JRV}_t\right] = \frac{\bar{\omega}}{1-(\xi+\zeta)}\left(\Lambda^2+\delta^2\right). \tag{3.30}$$

3.5 Variance risk premium

As broadly recognized in literature, stock volatilities are stochastically time-varying. Following the common interpretation of volatility as a measure of the degree of uncertainty of stock returns, it could be observed that investors face, beyond the uncertainty of returns, the uncertainty about the volatility itself. Dealing with this additional source of riskiness represents an important topic for investors in order to effectively manage risk and return profile of large portfolios containing derivatives.

Due to the presence of such uncertainties, asset prices embed what are economically indicated as risk premia which reflect the compensation that market agents demand for bearing different types of risks related to their investments. In particular, variance risk premium (VRP) is a measure of the reward demanded for bearing risk related with unknown future level of return variance. This is usually intended as the difference or spread between the conditional expectation of variance under physical measure and risk-neutral measure.

It is well known that on average, the implied volatility in option prices exceeds the realized volatility resulting in significantly non-zero and time-varying variance risk premium (see for example Carr and Wu (2009) and Bakshi et al. (2010b)). For example, the VIX index, which represents a non-parametric option implied estimate of the volatility of S&P 500 Index averages about 19% per year, while the unconditional realized volatility is only about 16%. According to this difference, a one month maturity at-the-money option priced at 19% volatility is

about 18% more expensive than one priced at 16% volatility. So, big spreads between the two volatilities are translated into high fees for issuing options and could suggest the possibility for arbitrage profits. However, higher initial gains from "selling" volatility are usually associated by higher risk of temporal swings of returns variance.

An historical example of the consequences resulting from underestimating the role of the variance risk premium is the fail of Long-Term Capital Management in 1998. At that time, LTCM was highly involved in shorting options at prices that implied a market volatility of 19%. Despite the increase of option prices, indicating a rise in the perceived market volatility, the fund kept selling. LTCM was neglecting the variance risk and ended with a net vega exposure of −$40 million in U.S. market and an equivalent amount in Europe, gaining the nickname of "Central Bank of Volatility" from the competitor Morgan Stanley (Lowenstein (2000)). When realized volatility rose significantly (to a level of nearly 45% by mid-September 1998), this strategy caused catastrophic losses estimated in $1.3 billion contributing to the final downturn of the hedge fund.

After LTCM collapse, volatility derivatives like variance swap contracts began to be widely traded in OTC markets. The payoff of these contracts depends on the difference between realized variance and a fixed variance level, or strike quoted at inception, of an underlying asset, like a stock index, over a predefined period. This feature makes such contracts attractive both for hedgers as mutual funds and portfolio managers who seek insurance against volatility swings and for speculators wanting pure volatility exposure. As direct consequence, the interest in technical aspects of these contracts has increased in the recent literature together with the need to understand and define a measure of the difference between volatilities under risk-neutral measure and physical one.

Following Bollerslev and Todorov (2011), we define the variance risk premium as

$$\mathrm{VRP}\,(t,T) = \frac{1}{T-t}\left(\mathbb{E}_t^{\mathbb{P}}\left[QV_{t,T}\right] - \mathbb{E}_t^{\mathbb{Q}}\left[QV_{t,T}\right]\right) \tag{3.31}$$

where $QV_{t,T}$ is the quadratic variation of the price of an asset over the time interval $[t,T]$. We note that according to this definition the premium depends on

both physical and risk-neutral measure since it is built as the mean difference between conditional expectations of the quadratic variation over a time horizon $T-t$ computed under the two different measures. Typically, a market measure of the VRP uses the information coming from implied volatilities embedded in the quoted VIX to build a proxy for expected volatility under risk-neutral measure. On the other hand, the expectation of the physical volatility process is usually estimated involving a reduced form model for the realized volatility process. The HAR model by Corsi (2009) has been widely adopted (see Bollerslev et al. (2009b), Bollerslev and Todorov (2011), Bormetti et al. (2015b)). Nevertheless, this "market" proxy has the inconvenience to represent a measure for VRP only over a monthly horizon since its dependence on VIX. This fact implies the impossibility to extract from market quotes the possible dependence of VRP on different time horizons also said its term structure.

Adopting the proposed JLHARG-RV and LHARG-ARJ-RV processes, we use RV_t as estimator of daily variance under physical measure while the risk-neutral estimator indicated as RV_t^* follows the equivalent processes with parameters mapped according to Propositions 6 and 9, respectively. So the expression for variance risk premium (3.31) is translated in our framework as

$$\mathrm{VRP}\,(t,T) = \frac{1}{T-t}\left(\mathbb{E}_t^{\mathbb{P}}\left[\sum_{k=1}^{T-t}\mathrm{RV}_{t+k}\right] - \mathbb{E}_t^{\mathbb{Q}}\left[\sum_{k=1}^{T-t}\mathrm{RV}_{t+k}^*\right]\right). \tag{3.32}$$

An important result is that this quantity can be computed analytically through the recursive formulas determining the conditional expectation of realized variance over a time horizon $T-t$. We state the following propositions for both the proposed models:

Proposition 10. *The conditional expectation of the realized variance over a time horizon T, given an information set up to time t, $\mathbb{E}_t^{\mathbb{P}}\left[\mathrm{RV}_{t+T}\right]$ for a JLHARG process is given by*

$$\mathbb{E}_t^{\mathbb{P}}\left[\mathrm{RV}_{t+T}\right] = a_t + \sum_{i=1}^{22} b_{t,\,i}\mathrm{RV}_{t+i-1}^c + \sum_{j=1}^{22} c_{t,\,j}\ell_{t+j-1} + \tilde{\Theta}\tilde{\theta}\tilde{\delta} \tag{3.33}$$

where

$$
\begin{aligned}
a_s &= a_{s+1} + \theta\delta b_{s+1,1} + \left(1 + \gamma^2\left(\theta\delta + \tilde{\Theta}\tilde{\theta}\tilde{\delta}\right)\right) c_{s+1,1} \\
b_{s,i} &= \begin{cases} b_{s+1,i+1} + \beta_i\left(b_{s+1,1} + \gamma^2 c_{s+1,1}\right) & for\ 1 \le i \le p-1 \\ \beta_i\left(b_{s+1,1} + \gamma^2 c_{s+1,1}\right) & for\ i = p \end{cases} \\
\mathrm{c}_{s,j} &= \begin{cases} \mathrm{c}_{s+1,j+1} + \alpha_j\left(b_{s+1,1} + \gamma^2 c_{s+1,1}\right) & for\ 1 \le j \le q-1 \\ \alpha_j\left(b_{s+1,1} + \gamma^2 c_{s+1,1}\right) & for\ j = q \end{cases}
\end{aligned}
\tag{3.34}
$$

Proof: see Appendix A.4.

Proposition 11. *For a LHARG-ARJ process, the conditional expectation of the realized variance over a time horizon $T-t$, given an information set up to time t, is given by*

$$
\mathbb{E}_t^{\mathbb{P}}\left[\mathrm{RV}_T\right] = \mathrm{a}_t + \sum_{i=1}^{p} \mathrm{b}_{t,i}\mathrm{CRV}_{t+1-i} + \sum_{j=1}^{q} \mathrm{c}_{t,j}\ell_{t+1-j} + \mathrm{d}_{t+1}\left(\xi\omega_t + \zeta n_t\right) \tag{3.35}
$$

where

$$
\begin{aligned}
\mathrm{a}_s &= \mathrm{a}_{s+1} + \theta(\kappa+d)\mathrm{b}_{s+1,1} + \left(1 + \gamma^2\theta(\kappa+d)\right)\mathrm{c}_{s+1,1} + \bar{\omega}\mathrm{d}_{s+1} \\
\mathrm{b}_{s,i} &= \begin{cases} \mathrm{b}_{s+1,i+1} + \beta_i\left(\mathrm{b}_{s+1,1} + \gamma^2\mathrm{c}_{s+1,1}\right) & for\ 1 \le i \le p-1 \\ \beta_i\left(\mathrm{b}_{s+1,1} + \gamma^2\mathrm{c}_{s+1,1}\right) & for\ i = p \end{cases} \\
\mathrm{c}_{s,j} &= \begin{cases} \mathrm{c}_{s+1,j+1} + \alpha_j\left(\mathrm{b}_{s+1,1} + \gamma^2\mathrm{c}_{s+1,1}\right) & for\ 1 \le j \le q-1 \\ \alpha_j\left(\mathrm{b}_{s+1,1} + \gamma^2\mathrm{c}_{s+1,1}\right) & for\ j = q \end{cases} \\
\mathrm{d}_{s+1} &= \left(\xi + \zeta\right)\mathrm{d}_{s+2}
\end{aligned}
\tag{3.36}
$$

with ZM-LHARG parameters mapped according to the equivalent P-LHARG process.

Proof: see Appendix B.4.

We observe that the conditional expectation is considered under $\mathbb{P}$ measure in the previous propositions, but the same formula holds under $\mathbb{Q}$ since we proved that, after the change of measure, the risk-neutral process is still a JLHARG or LHARG-ARJ process re-parametrized according to the mappings in Propositions 6 and 9, respectively.

Chapter 4

Pricing results and stock market returns predictability

4.1 Model Estimation and calibration on real data

This section gathers the results of parameters estimation and calibration procedure of the novel realized volatility models proposed in this thesis on real financial data.

4.1.1 JLHARG case

The first step is the construction of the RV time series from a high-frequency returns dataset. For our empirical study, we use tick-by-tick data for the S&P500 Index Futures, from January 1, 1990 to December 31, 2007. Since the two types of models specify the jump components of the dynamics in two different ways, we consequently build the empirical realized volatility measures for each type, respectively. Our estimation procedure for the continuous and jump component is the following:

(i) we estimate the realized variance of the log-prices using the Two-Scale estimator introduced by Zhang et al. (2005) which provides a robust measure

removing the bias due to the microstructure noise;

(ii) we identify the discontinuous component using the Threshold Bipower variation method by Corsi et al. (2010) which detects the spikes in RV time series and separates it from the continuous component.

Doing so, we note that the RV series is built from open-to-close (trading day) data, thus neglecting the overnight contribution. We adjust our RV estimator by rescaling the time series so to match the unconditional mean of the squared daily returns (close-to-close).
We stress that the adopted jump detection method, according to point (ii) of our procedure, represents a formal statistical test based on asymptotic theory. This is important to statistically identify days with jumps and subsequently associate the most extreme intra-day price movements to jump events.
After constructing the two time series for the RV components, we can estimate the parameters under $\mathbb{P}$ of the JLHARG-RV using the Maximum Likelihood Estimators as defined in Section 3.3.4. We choose the Fed Funds rate as proxy for the risk-free rate r.

In Table 4.1 the values of the estimated parameters under $\mathbb{P}$ are reported. We show the results for three different models JHARG, P-JLHARG and ZM-JLHARG together with their standard deviations and the values of the log-likelihood function. Our results confirm that the impact of past RV on the current one decreases with the increase of the aggregation horizon. The same evidence has been documented by Corsi (2009), Corsi and Renò (2012) and Majewski et al. (2015).

In order to derive the risk-neutral dynamics, the values of risk premium parameters for each proposed model need to be identified. According to Propositions 5, some of them are fixed by the no-arbitrage conditions, while the variance risk premia remain as free parameters to be calibrated on option prices. This is a direct consequence of the fact that we are dealing with an incomplete market since the models include variance as a further source of randomness.
Our data set consists of Plain Vanilla OTM options on S&P500 Index for each Wednesday from January 1, 1996 to December 31, 2004. We first apply a standard

	Model		
Parameter	JHARG	P-JLHARG	ZM-JLHARG
λ		2.6	
		(1.5)	
θ	9.75e-06	9.08e-06	9.43e-06
	(9e-08)	(8e-08)	(8e-08)
δ	1.358	1.252	1.89
	-	-	-
β_d	4.67e+04	3.49e+04	3.90e+04
	(8e+02)	(8e+02)	(5e+02)
β_w	2.9e+04	3.16e+04	3.07e+04
	(1e+03)	(1.6e+03)	(3e+02)
β_m	1.19e+04	1.37e+04	1.549e+04
	(9e+02)	(7e+02)	(1.4e+02)
α_d	-	0.278	0.412
		(0.004)	(0.008)
α_w	-	0.067	0.435
		(0.013)	(0.019)
α_m	-	3.4e-06	0.50
		(4e-07)	(0.07)
γ	-	174	121
		(7)	(7)
$\tilde{\theta}$		4.7e-05	
		(3e-06)	
$\tilde{\delta}$		1.152	
		-	
$\tilde{\Theta}$		0.299	
		(0.009)	
ν_c	739	-1431	-2257
ν_j	-12428	-10251	-7938
Log-likelihood	-24648	-24543	-24476
Persistence	0.8536	0.8229	0.8030

Table 4.1: Maximum likelihood estimates, robust standard errors, and log-likelihood values. The historical data for the JHARG, P-JLHARG and ZM-JLHARG models are given by the daily RV computed on tick-by-tick data for the S&P500 Index futures. For all three models, the estimation period ranges from the period 1990-2007.

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)). Using K/S_t as definition of moneyness, we filter out DOTM 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 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).

Calibration procedure is based on the unconditional minimization of 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 – as described above – obtaining a 5×4 moneyness-maturity grid. Then, for each subset, we compute the unconditional mean of the market implied volatilities. In this way, as shown in Table 4.2, we obtain a 20-point discrete representation

	Maturity			
Moneyness	$\tau \leq 50$	$50 < \tau \leq 90$	$90 < \tau \leq 160$	$160 < \tau$
	Implied Volatility			
$0.7 \leq m \leq 0.9$	0.3564	0.3056	0.2866	0.2662
$0.9 < m \leq 0.98$	0.2353	0.2269	0.2232	0.2230
$0.98 < m \leq 1.02$	0.1958	0.2023	0.2059	0.2108
$1.02 < m \leq 1.1$	0.1767	0.1790	0.1849	0.1923
$1.1 < m \leq 1.3$	0.2317	0.1946	0.1836	0.1842

Table 4.2: Mean market implied volatilities of S&P500 Index options on each Wednesday from January 1,1996 to December 31, 2004 (46066 observations) sorted by moneyness and maturity. Moneyness is defined as $m = K/S_t$, where K and S_t are the strike and the underlying price, respectively. Maturity is measured in calendar days.

of the implied volatility surface. Finally, we compute the same discrete grid for the model implied volatility and we identify the optimal values of (ν_c, ν_j) which minimize the distance between the two grids, i.e.

$$\arg\min_{(\nu_c,\nu_j)} \{f_{\text{obj}}(\nu_c, \nu_j)\}.$$

The objective function $f_{\text{obj}}(\nu_c, \nu_j)$ is defined as

$$f_{\text{obj}}(\nu_c, \nu_j) = \sqrt{\sum_{i=1}^{5}\sum_{j=1}^{4}\left(\text{IV}_{ij}^{\text{mod}}(\nu_c, \nu_j) - \text{IV}_{ij}^{\text{mkt}}\right)^2}, \tag{4.1}$$

and represents the quadratic distance between the model implied volatility surface and the market one, whose elements are $\text{IV}_{ij}^{\text{mod}}(\nu_c, \nu_j)$ and $\text{IV}_{ij}^{\text{mkt}}$, respectively. In order to compute the option prices – and associated implied volatilities – we employ a numerical scheme introduced by Fang and Oosterlee (2008), termed the COS method. This method, based on Fourier-cosine expansions, efficiently evaluates the price of Plain Vanilla options from the characteristic function of log-returns.

At the bottom of Table 4.1 we report the calibrated variance risk premia for JHARG models. It is worth recalling that the presence of a positive or a negative value of the risk premium reduces or amplifies the unconditional mean of realized variance, respectively. Moreover, negative premia have the genuine effect to induce more skew in the distribution of returns. The risk premium, ν_c, associated with the continuous component varies from the positive value of the JHARG model to a large negative value for the ZM-JLHARG model. The risk premia, ν_j, associated with the jump component, are all negative and increasing (decreasing in absolute terms) when a better specified form of the leverage is adopted. The most negative jump premium corresponds to the JHARG model and decreases for the P-JLHARG with heterogeneous parabolic leverage. It reaches the highest value (smallest in absolute terms) for the ZM-JLHARG where the heterogeneous leverage is centred. The compensation taking place between ν_c and ν_j is due to the fact that large negative innovations in the price rise the future variance through the leverage term. Then, a better specification of the leverage component reduces the relative weight of the jump premium in favour of the premium of the continuous component.

4.1.2 LHARG-ARJ case

As concerns LHARG-ARJ estimation procedure on real data, the starting point is the same as the one introduced in the previous subsection for JLHARG models. The first step consists in the construction of the empirical measures of realized volatility. The second concerns the identification and separation of its continuous and jump components. To accomplish the first one, we involve the Two-Scale estimator while the second one is based on the adoption of a jump detection test. In this case we use a dataset of HF returns of the S&P500 Index Futures from 1 July 1990 to 31 June 2011. We stress the fact that the model assumes the presence of jump with autoregressive intensity dynamics depending on the past realized number of jumps occurring during a trading day. Nevertheless jump tests like the Threshold Bipower Variation method allows to spot the days on which at least one jump occurs but does not give information about the number of intraday jumps actually occurred. So the complete information can not be retrieved from application of this sole test.

In order to identify the intraday events we employ a combination of the Threshold Bipower Variation method (Tz-test) and procedure introduced by Andersen et al. (2010). We start by constructing the series of intraday 5-min returns. Then we run the Tz-test according to the prescriptions reported in Section 1.4. Once we find a day with at least one jump, we remove the largest 5-min return in that day and substitute it with the average return for that day, then we run again the Tz-test for this new adjusted intraday series. If the test does not reject the null, we conclude that only one jump occurred. If the test rejects the null, the procedure is repeated. At every run we identify a further jump. Finally, when the Tz-test does not reject, we have the series of intraday 5-min jump returns from tick-by-tick prices. With this information, we are able to reconstruct the daily contribution of jumps to the total daily return and determine the jump-adjusted return series by simply subtracting this contribution.

After separating the continuous and the discontinuous components of the dynamics (Figure 4.1), we estimate the parameters under physical measure of the LHARG-ARJ process using the Maximum Likelihood Estimator.

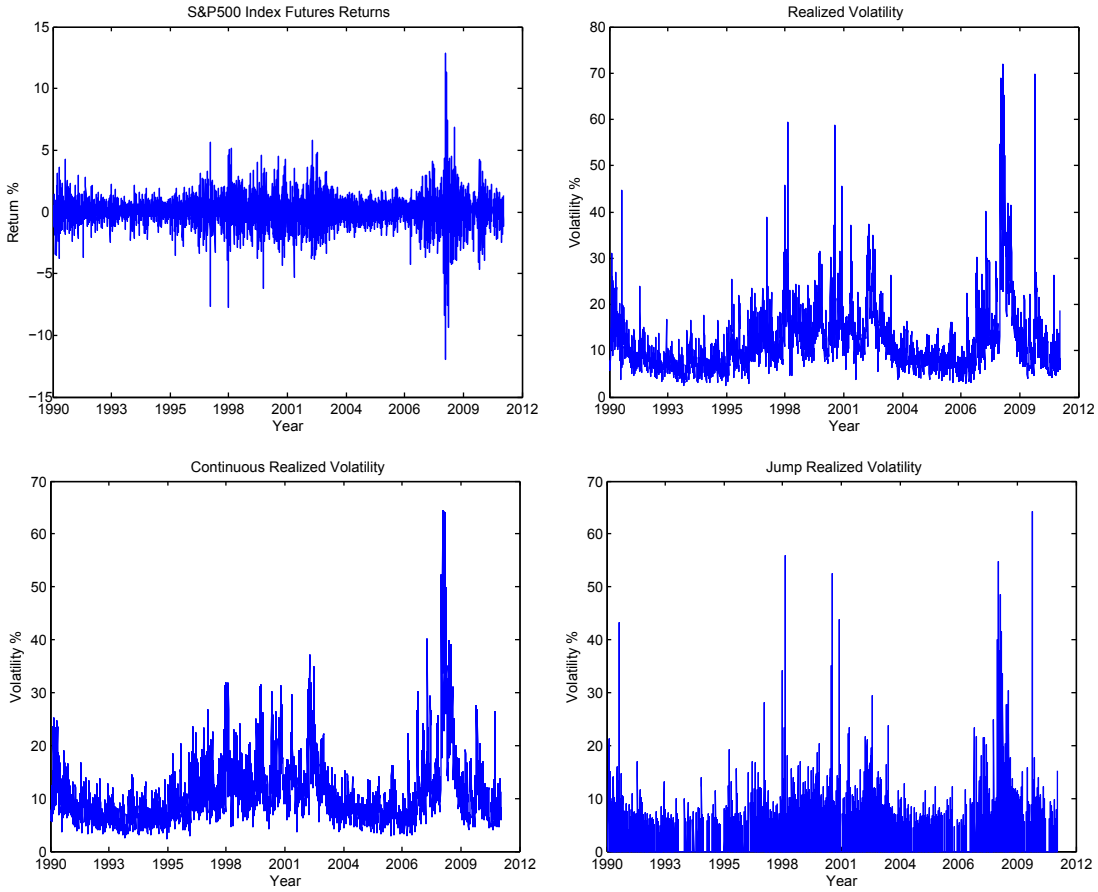

Figure 4.1: Daily returns and realized volatility time series for SP&500 Index Futures. Realized Volatility is computed using the Two-Scale method by Zhang et al. (2005). The sample starts July 3, 1990 and ends June 28, 2011.

LHARG-ARJ model estimation results are summarized in Table 4.3. We point out that the presence of the financial crisis of 2007/2008 implies a sort of structural break of the statistical characteristics of the empirical distributions of the realized variance from the previous to the following time periods. Figure 4.2 represents the densities of the empirical distributions of the continuous contribution of realized volatility for two sub-periods obtained from the original one: a pre-crisis period from July 1990 to June 2007 and a crisis/post-crisis period from July 2007 to June 2011. We observe that the two samples are quite different in particular for extreme values where the distribution related to the crisis/post-crisis period shows a clear fatter right tail. Considering the first two moments of the two samples, we obtain values of 11.54 and 33.31, respectively mean and variance for

	LHARG-ARJ	
Parameter	Data period 1990-2007	Data period 2007-2011
λ_c	1.3	-2
	(1.9)	(2)
λ_j	0.5	3
	(1.6)	(10)
θ	8.7e-06	2.6e-05
	(5e-07)	(2e-06)
κ	1.43	1.238
	-	-
β_d	4.2e+04	1.9e+04
	(4e+03)	(3e+03)
β_w	3.4e+04	1.2e+04
	(4e+03)	(3e+03)
β_m	1.7e+04	9e+02
	(3e+03)	(1.5e+03)
α_d	1.3e-01	5.7e-02
	(3e-02)	(1.5e-02)
α_w	1.1e-01	6.7e-01
	(2e-02)	(1.8e-02)
α_m	1.1e-01	5e-03
	(4e-02)	(4e-02)
γ	3.5e+02	4.2e+02
	(6e+01)	(9e+01)
$\bar{\omega}$	2.9e-03	5.0e-04
	-	-
ξ	9.7e-01	9.80e-01
	(1.0e-02)	(1.9e-02)
ζ	2.9e-02	2.0e-02
	(7e-03)	(1.5e-02)
Λ	-4.4e-04	-1.5e-04
	(1.6e-04)	(4e-04)
δ	5.0e-03	5.9e-03
	(2e-04)	(3e-04)
ν_c	-7.94e+03	-1.49e+03
ν_j	-8.29e+03	-2.29e+03
Persistence CRV_t	0.812	0.831
Persistence ω_t	0.994	0.999

Table 4.3: Maximum likelihood estimates (standard errors in parentheses) for LHARG-ARJ on S&P500 Index Futures on two periods: from July 1990 to June 2007 and from July 2007 to June 2011.

CRV series from 1990 to 2007, while the same quantities for the subsequent period are 17.48 and 159.70, respectively. Furthermore, the range of the most frequent

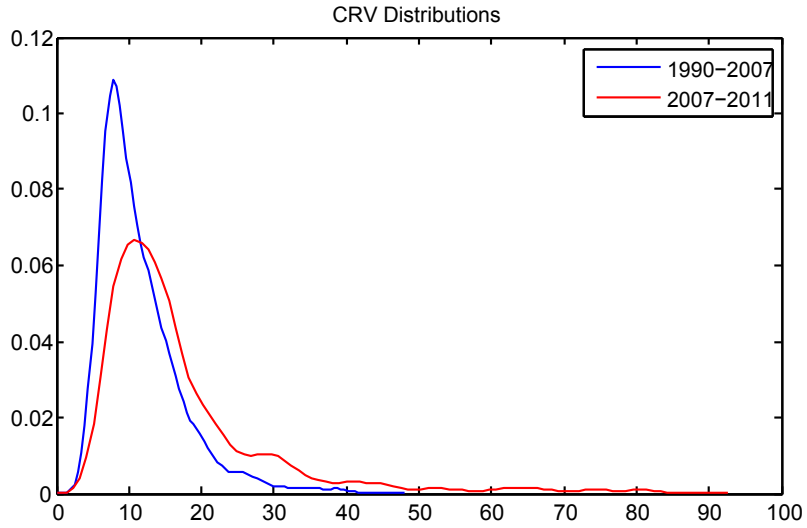

Figure 4.2: Empirical densities of the continuous component of realized volatility (CRV). Blue line refers to the pre-crisis period from July 1990 to June 2007, red line to the crisis/post-crisis period from July 2007 to June 2011.

values in the dataset for the crisis/post-crisis is larger than the pre-crisis one but less likely in comparison, translating in a higher dispersion for the sample. Due to these important differences in the statistical properties of the real data, we choose to estimate the LHARG-ARJ model over the two sub-periods.

After estimating parameters under physical measure, the second step to make the model available for pricing purposes is the calibration procedure to identify the values of risk premia allowing for risk-neutralization of the process$((\mu_c, \mu_j, \nu_c, \nu_j))$in our notation). According to Proposition 8, in our framework μ_c and μ_j are fixed by the no-arbitrage condition, while ν_c and ν_j remain undetermined parameters to be fixed from calibration on real option quotes. For the LHARG-ARJ case we use a dataset of Plain Vanilla put options on S&P500 Index for each Wednesday from January 1, 1996 to June 28, 2011. 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\$. Using K/S_t as definition of moneyness, we filter out deep in-the-money options with moneyness larger than 1.3 and deep out-of-the-money than 0.7. This choice yields a total number of 76331 observations. For our purposes, we will identify put options as deep out-of-the-money (DOTM) if their moneyness is between $0.8 \leq m \leq 0.9$ and

out-of-the-money (OTM) if $0.9 < m \leq 0.98$. On the other hand, we will indicate them as deep-in-the-money (DITM) if $1.1 < m \leq 1.2$ and in-the-money ITM 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).
For the calibration procedure, we adopt the same method based on the unconditional minimization of the distance between the market implied and the model implied volatility surface as introduced for JLHARG. So, we divide our dataset in different intervals of moneyness and maturity – as described in the previous paragraph – obtaining a 5×4 moneyness-maturity grid obtaining a 20-point discrete representation of the implied volatility surface. The optimal values of (ν_c, ν_j) are obtained by finding the minimum of the objective function of the form given in equation (4.1).

4.2 Option pricing benchmarking

This section is dedicated to the presentation of pricing results with JLHARG and LHARG-ARJ models. We can summarize the option pricing procedure in four steps: (i) estimation of the parameters under the physical measure $\mathbb{P}$; (ii) unconditional calibration of the parameter vector (ν_c, ν_j); (iii) mapping of parameter values from $\mathbb{P}$ to $\mathbb{Q}$ using expressions (3.8); (iv) numerical computation of option prices through COS method using the MGF recursive formulas in (16).
As benchmark approaches to assess the pricing performance of JHARG models we use the standard Heston and Nandi (2000) affine GARCH model and the BiPower Jump Variation Model (BPJVM) introduced by Christoffersen et al. (2015).
The BPJVM model is a state-of-the-art approach incorporating a GARCH structure for the latent volatility and jump intensity where bipower and jump variations play a prominent role as idiosyncratic components. The novelty of its contribution is to use directly the observable realized diffusive volatility and realized jump variation to model the dynamics of the two latent volatility components.

The dynamics of log-returns in BPJVM is as follows

$$y_t = r + \left(\lambda_c - \frac{1}{2}\right) h_{z,t} + (\lambda_j - \eta) h_{y,t} + \sqrt{h_{z,t}}\epsilon_{1,t} + \sum_{i=0}^{n_t} X_{t,i},$$

where $\epsilon_{1,t} \sim \mathcal{N}(0,1)$ are i.i.d innovations and jumps' sizes $X_{t,i}$ are independent random variables with normal distribution $\mathcal{N}(\theta, \delta^2)$. $h_{z,t}$ denotes the diffusive variance and the number of jumps n_t has Poisson distribution with intensity $h_{y,t-1}$. The parameter η is set to $\exp(\theta + 0.5\delta^2) - 1$ in order to ensure that the conditional expected total return is $\mathbb{E}_t[\exp(R_{t+1})] = \exp(r + \lambda_z h_{z,t} + \lambda_y h_{y,t})$. Each day, the dynamics of $h_{z,t}$ and $h_{y,t}$ are specified incorporating new information carried by realized bipower and jump variation measures

$$h_{z,t} = \omega_z + b_z h_{z,t-1} + a_z \mathrm{RBV}_t$$
$$h_{y,t} = \omega_y + a_y h_{y,t-1} + b_y \mathrm{RJV}_t$$

where RBV_t is the daily realized bipower variation computed from intraday returns via equation (1.3) and RJV_t is the daily realized jump variation constructed as the simple difference between the total quadratic variation RV_t (built using a market microstructure robust measure based on averaging five overlapping realized variance series from squared 5-minute returns) and the bipower variation. In order to ensure that $\mathrm{RJV}_t \geq 0$, the following definitions are used,

$$\mathrm{RBV}_t = \min(\mathrm{RV}_t, \mathrm{RBV}_t) \tag{4.2}$$
$$\mathrm{RJV}_t = \mathrm{RV}_t - \mathrm{RBV}_t. \tag{4.3}$$

Since RBV_t is measured with error, the model is then completed specifying the following measurement equation

$$\mathrm{RBV}_t = h_{z,t-1} + \sigma\left[\left(\epsilon_{2,t} - \gamma\sqrt{h_{z,t-1}}\right)^2 - (1 + \gamma h_{z,t-1})\right], \tag{4.4}$$

where $\epsilon_{2,t} \sim \mathcal{N}(0,1)$ are i.i.d. random variables and have correlation ρ with the idiosyncratic shocks introduced in the return equation. The innovation term inside square brackets in equation (4.4) is designed to have zero mean in order to

imply that $\mathbb{E}_{t-1}[\mathrm{RBV}_t] = h_{z,t-1}$. On the other hand, the realized jump variation measure is related to the sum of squared daily jump returns as follows

$$\mathrm{RJV}_t = \sum_{i=0}^{n_t} (X_{t,i})^2,$$

which implies that $\mathbb{E}_{t-1}[\mathrm{RJV}_t] = (\theta^2 + \delta^2)\, h_{y,t-1}$. Contrary to RBV_t series, it is assumed that RJV_t is free of error for simplicity.
Risk-neutralization is obtained by employing a three dimensional Esscher transform with two parameters to be calibrated on option data (χ, ν_3).

Table 4.4 summarizes estimates for parameters of the competitor models. Differently from the methodology adopted in Christoffersen et al. (2015), we estimate the BPJVM model via quasi-maximum likelihood method using physical information and afterwards calibrate free parameters χ, ν_3 in the same way as described in the previous section[1].
This choice is coherent with the scheme adopted in this thesis which first estimates the physical parameters from historical realizations of returns and variances and, then, separately assesses the impact on the risk neutral dynamics of the premia embedded in option prices.
However, Christoffersen and co-authors adopt a definition of the jump component which does not involve a test to statistically assess the significance of each event – as typically done in the jump literature. As a result the realized jump series they consider contains at least one jump per day, which is at odd with the generally accepted view of jumps as extreme and rare events. Secondly, the measurement equation describing the dynamics of the realized variance implies the sampling of random values which can be negative with significant likelihood (up to 20% of occurrences over 10^6 Monte Carlo simulated samples) depending on the level of the latent diffusive volatility. This fact poses issues of consistency with the definition of variance which has to be a positive number and consequently the realism of the model for pricing purposes is troublesome. Finally, although the proposed model has the desired property of reproducing the high

[1]For practical implementation, we refer to the updated version available on SSRN including some corrections to the published version. Link: http://ssrn.com/abstract=2494379.

Parameter	GARCH	BPJVM
λ_z	1.1	1
	(1.4)	(4)
λ_y	-	4e-05
	-	(8e-05)
γ	1.22e+02	1.44e+04
	(1.1e+01)	(1.7e+02)
ω_z	6.41e-09	2.48e-08
	-	-
ω_y	-	0.0423
	-	-
σ	-	1.86e-07
	-	(2e-09)
θ	-	1e-05
	-	(5e-05)
δ	-	1.278e-03
	-	(1.5e-05)
ρ	-	3.3e-01
	-	(2e-02)
b_z	8.7e-01	6.50e-01
	(1e-02)	(3e-02)
b_y	-	9.51e-01
	-	(1.2e-02)
a_z	5.5e-06	3.5e-01
	(5e-07)	(3e-02)
a_y	-	2.2e+04
	-	(5e+03)
χ	-	-2.85
ν_3	-	2.5e-04
Persist_z	0.9481	0.9999
Persist_y	-	0.9865

Table 4.4: Maximum likelihood estimation of competitor models on daily S&P 500 futures returns for the period 1990-2007.

persistence of volatility, estimates on real data indicate persistence values close to one. Such an extreme value could imply that the model is not sufficiently flexible to fit the whole term structure of implied ATM volatilities, in particular facing miss-valuation for long-term maturities.

As customary in the literature (Renault (1997), Corsi et al. (2013), Majewski et al. (2015)), we employ the Root Mean Square Error (RMSE) on the percentage

IV as performance measure, i.e.

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

where N is the number of options, and IV^{mod} and IV^{mkt} are the model and the market implied volatility, respectively.

Before involving the chosen benchmark models, we perform a pricing performance comparison among the different specifications of JLHARG models involved in our study.

Implied Volatility RMSE

	Moneyness	
Model	$0.9 < m < 1.1$	$0.8 < m < 1.2$
JHARG	4.919	6.646
P-JLHARG/JHARG	0.908	0.928
ZM-JLHARG/JHARG	0.833	0.862
ZM-JLHARG/P-JLHARG	0.918	0.929

Table 4.5: Global option pricing performance for the JLHARG class of models on S&P500 options from January 1, 1996 to December 31, 2004, computed with the RV measure estimated from 1990 to 2007. We use the maximum likelihood parameter estimates from Table 4.1.

In Table 4.5 we report the global comparison of the option pricing performances between models belonging to the JHARG class. We build ratios between the RMSE of each couple of models. The table shows that – in terms of RMSE – the performance improves for models accounting for the leverage effect, as expected. Specifically, fixing as benchmark the JHARG model with no leverage, performances in the range of moneyness $0.9 < m < 1.1$ improve by nearly 9% for P-JLHARG and by 17% for ZM-JLHARG. In the range of moneyness $0.8 < m < 1.2$ the improvements are by 7% and by 14% for P-JLHARG and ZM-JLHARG, respectively. These results confirm the well established fact that the inclusion of a leverage component is essential for the correct option pricing. Moreover, ZM-JLHARG always outperforms P-JLHARG independently on

the range of moneyness. In accordance with Majewski et al. (2015) this finding reaffirms that the zero-mean leverage shows more flexibility with respect to the parabolic leverage.

Implied Volatility RMSE

	Moneyness	
Model	$0.9 < m < 1.1$	$0.8 < m < 1.2$
ZM-JLHARG	4.098	5.732
ZM-JLHARG/GARCH	0.661	0.680
ZM-JLHARG/ZM-LHARG	1.045	0.987
ZM-JLHARG/BPJVM	0.848	0.867

Table 4.6: Global option pricing performance for the ZM-JLHARG model on S&P500 options from January 1, 1996 to December 31, 2004 and comparison with Heston-Nandi GARCH, ZM-LHARG, and BPJVM as benchmark models.

Table 4.6 shows the global pricing performance of the ZM-JLHARG model, which is the best performing model within the JHARG class, benchmarked with the Heston-Nandi GARCH model, the ZM-LHARG model by Majewski et al. (2015), and the BPJVM. All models incorporating the high frequency information embedded in realized measures outperforms the GARCH model. This result might be expected and a comparison with the BPJVM model is more appropriate. We observe that all models with an heterogeneous autoregressive volatility outperform the BPJVM model. Gains in performance for the ZM-JLHARG model vary from 15.2%, in the range of moneyness close to ATM, to 13.3%, when more extreme moneyness are included. The result for the central region of volatility surface allows to reconfirm that the heterogeneous structure is a parsimonious and effective way to provide a very good description of the ATM implied volatility dynamics. As will be stressed more in Table 4.7, this is true for short maturities and becomes even more evident for longer horizons. When wider area of moneyness is considered, as a consequence of the autoregressive structure of jump intensity, BPJVM partially recovers the loss of performance.

	Maturity			
Moneyness	$\tau \leq 50$	$50 < \tau \leq 90$	$90 < \tau \leq 160$	$160 < \tau$
Panel A	ZM-JLHARG Implied Volatility RMSE			
$0.7 \leq m \leq 0.9$	12.046	7.632	6.233	5.101
$0.9 < m \leq 0.98$	4.058	3.650	3.794	4.103
$0.98 < m \leq 1.02$	3.518	3.802	4.053	4.484
$1.02 < m \leq 1.1$	4.204	4.654	4.807	4.893
$1.1 < m \leq 1.3$	4.660	3.870	4.503	4.982
Panel B	ZM-JLHARG/BPJVM Implied Volatility RMSE			
$0.7 \leq m \leq 0.9$	0.866	0.972	1.080	0.952
$0.9 < m \leq 0.98$	0.684	0.740	0.802	0.727
$0.98 < m \leq 1.02$	0.992	0.857	0.861	0.768
$1.02 < m \leq 1.1$	1.091	1.125	1.086	0.854
$1.1 < m \leq 1.3$	0.718	0.707	0.737	0.683

Table 4.7: Panel A: Percentage implied volatility root mean squared error ($RMSE_{IV}$) of the ZM-JLHARG model sorted by moneyness and maturity. Panel B: $RMSE_{IV}$ ratios computated using BPJVM as benchmark model.

Nonetheless, the global over-performance of the ZM-JLHARG model is still significant. Finally, it is worth commenting on the performance of the ZM-JLHARG with respect to the same model without the jump component. In the ATM region the ZM-LHARG model performs very well and better than ZM-JLHARG by a factor of 1.045. Once more, this result confirms the crucial role played by the HARG structure in the dynamics of the continuous component. Then, consistently with the role expected to be played by discontinuous events, ZM-JLHARG over-performs the competitor model in the extreme regions of moneyness where an increment of probability mass is essential.

As anticipated, in Table 4.7 we choose to focus on a more detailed comparison between the ZM-JLHARG model and the competitor BPJVM. Dividing the entire dataset of options according to the grid used for model calibration (see Section 4.1.1), we observe that for short maturities $\tau \leq 50$ the two models price with almost the same accuracy in the at-the-money region $0.98 < m < 1.02$. BPJVM increases the pricing performance for OTM call options but for DOTM

calls the ZM-JLHARG performs better by a factor of 0.718. In the region of short-maturity puts the ZM-JLHARG model consistently over-performs the competitor BPJVM. This result is confirmed with slightly different percentages for options with medium-short maturity $50 < \tau \leq 90$ noting a worsening of BPJVM performance in the ATM region with respect to ZM-JLHARG. As concerns the medium-long maturity region $90 < \tau \leq 160$ BPJVM maintains a higher performance in pricing OTM calls and shows a better valuation of DOTM puts than the ZM-JLHARG. Finally, valuation of long maturity options exhibits the consistent over-performance of the ZM-JLHARG model for both puts and calls covering all ranges of moneyness under study. The smaller error of ZM-JLHARG for long maturity options suggests that this model has more flexibility than BPJVM to reproduce a realistic term structure of implied volatilities. A possible reason for the rigidity of BPJVM could be the extremely high persistence of both volatility and jump intensity processes, as reported in Table 4.4. High persistence is a crucial feature to reproduce the long-memory property of the volatility process, nevertheless an extreme level could have the side effect of systematical miss-valuing options – either over-pricing or under-pricing depending on the prevailing high or low level of volatility, respectively.

Separately from pricing performances regarding JLHARG models, we present the results of another pricing benchmarking exercise involving the other model introduced in this thesis: the LHARG-ARJ. In this case we use the ZM-LHARG model by Majewski et al. (2015) and the BPJVM model as terms of comparison considered, to the best of our knowledge, the state-of-the-art for the modelling framework based on realized measure of volatilities (see Table 4.8 and 4.9 for the estimated parameters of this competitor models).

Tables 4.10 and 4.11 show that LHARG-ARJ improves option pricing performance with respect to ZM-LHARG and BPJVM for both region of moneyness, before and after financial crisis of 2008. The improvement ranges from 8.5% to 7.6% of relative $RMSE$ with respect to ZM-LHARG for the pre-crisis period,

	ZM-LHARG	
Parameter	Data period 1990-2007	Data period 2007-2011
λ_c	3.77	-1.56
	-	-
θ	8.5e-06	2.6e-05
	(4e-07)	(2e-06)
κ	1.37	1.238
	-	-
β_d	4.3e+04	1.9e+04
	(4e+03)	(3e+03)
β_w	3.6e+04	1.2e+04
	(4e+03)	(3e+03)
β_m	1.7e+04	1.0e+03
	(3e+03)	(1.7e+03)
α_d	1.6e-01	4e-02
	(3e-02)	(1e-02)
α_w	1.0e-01	5.3e-02
	(2e-02)	(1.9e-02)
α_m	8e-03	5e-03
	(1.3e-02)	(1.3e-02)
γ	3.0e+02	4.7e+02
	(5e+01)	(1.1e+02)
ν_c	-8.94e+03	-1.66e+03
Persistence CRV_t	0.824	0.834

Table 4.8: Maximum likelihood estimates (standard errors in parentheses) for ZM-LHARG on S&P500 Index on two periods: from July 1990 to June 2007 and from July 2007 to June 2011.

while the improvements lessen in the crisis/post-crisis period to 2%. This difference indicates that in period of great financial distress option prices seem to be influenced in large measure by the high level of the diffusive component of the underlying price which prevails the jump contribution of its dynamics. Consequently, the performance of a model accounting for jumps reduces to that of the equivalent model which consider only the continuous stochastic evolution.

Analysing the detailed comparison between the two models by dividing the entire dataset of options as described in Section 4.1.2, we observe that for options belonging to the per-crisis period LHARG-ARJ model's best performance concentrates on pricing contracts with short-maturity less then 50 days with RMSE up to 15% smaller than those obtained for the ZM-LHARG (see Table 4.12).

	BPJVM	
Parameter	Data period 1990-2007	Data period 2007-2011
λ_z	1	-6
	(4)	(5)
λ_y	4e-05	2.7e-04
	(8e-05)	(1.9e-04)
γ	1.44e+04	1.45e+04
	(1.7e+02)	(3e+02)
ω_z	2.48e-08	1.79e-09
	-	-
ω_y	4.23e-02	8.53e-02
	-	-
σ	1.86e-07	2.43e-07
	(2e-09)	(5e-09)
θ	1e-05	1.6e-04
	(5e-05)	(1.2e-04)
δ	1.278e-03	2.01e-03
	(1.5e-05)	(6e-05)
ρ	3.3e-01	3e-01
	(2e-02)	(4e-02)
b_z	6.50e-01	6.1e-01
	(3e-02)	(5e-02)
b_y	9.51e-01	9.09e-01
	(1.2e-02)	(1.7e-02)
a_z	3.5e-01	3.9e-01
	(3e-02)	(4e-02)
a_y	2.2e+04	2.2e+04
	(5e+03)	(5e+03)
χ	-5.16	-2.01
ν_3	2.59	1.70
Persistance h_z	0.9999	0.9999
Persistance h_y	0.9865	0.9979

Table 4.9: Maximum likelihood estimates (standard errors in parentheses) for BPJVM on S&P500 Index on two periods: from July 1990 to June 2007 and from July 2007 to June 2011.

These results confirm the well established fact that the inclusion of a jump component is essential for the correct description of the characteristics of the volatility surface implied by short-term options. Indeed, moving towards longer maturities, performances in the ATM region are in favour of the model neglecting the discontinuous dynamics while remaining higher for the LHARG-ARJ pricing of the DOTM, OTM and DITM options which are more sensible to the likelihood of extreme events.

	$RMSE_{IV}$	
	1990-2007	
Model \ Moneyness	$0.9 < m < 1.1$	$0.8 < m < 1.2$
ZM-LHARG	5.600	6.991
BPJVM	6.766	7.964
LHARG-ARJ/ZM-LHARG	0.915	0.924
LHARG-ARJ/BPJVM	0.757	0.811

Table 4.10: Comparison of global option pricing performance on S&P500 Index put options for models LHARG-ARJ, ZM-LHARG and BPJVM from January 1996 to June 2007.

	$RMSE_{IV}$	
	2007-2011	
Model \ Moneyness	$0.9 < m < 1.1$	$0.8 < m < 1.2$
ZM-LHARG	6.244	7.833
BPJVM	7.344	9.169
LHARG-ARJ/ZM-LHARG	0.983	0.979
LHARG-ARJ/BPJVM	0.836	0.837

Table 4.11: Comparison of global option pricing performance on S&P500 Index put options for models LHARG-ARJ, ZM-LHARG and BPJVM from July 2007 to June 2011.

The detailed table referring to the crisis/post-crisis period (Table 4.13) shows results in line with the previous period but the performances are more balanced between the models as consequence of the extraordinarily high level of volatilities typical of that distressed moment for which the impact of possible jumps is negligible.

As far as concerns the comparison between the LHARG-ARJ with the competitor BPJVM model, we observe that the first shows globally higher performances than the second for options traded in the two periods considered (see Tables 4.10 and 4.11). The highest one is obtained for contracts with moneyness between 0.9 and 1.1 (around 24%) in the pre-crisis period. Errors are still smaller of almost 19% considering a wider range of moneyness ($0.8 < m < 1.2$) and are stably lower of almost 14% for the period including and following the crisis. Even in

	Maturity			
Moneyness	$\tau \leq 50$	$50 < \tau \leq 90$	$90 < \tau \leq 160$	$160 < \tau$
Panel A	ZM-LHARG Implied Volatility RMSE			
$0.8 \leq m \leq 0.9$	15.592	8.803	6.049	5.022
$0.9 < m \leq 0.98$	7.512	4.240	3.826	6.191
$0.98 < m \leq 1.02$	4.166	3.561	4.104	7.151
$1.02 < m \leq 1.1$	6.768	3.512	4.020	7.111
$1.1 < m \leq 1.2$	22.776	6.192	3.239	5.981
Panel B	LHARG-ARJ/ZM-LHARG Implied Volatility RMSE			
$0.8 \leq m \leq 0.9$	0.940	0.902	0.919	0.922
$0.9 < m \leq 0.98$	0.848	0.877	0.975	0.925
$0.98 < m \leq 1.02$	0.851	1.087	1.082	0.945
$1.02 < m \leq 1.1$	0.862	1.099	1.117	0.964
$1.1 < m \leq 1.2$	0.932	0.859	1.052	0.992

Table 4.12: Detailed comparison of option pricing performance on S&P500 Index put options for models ZM-LHARG and LHARG-ARJ before the financial crisis: from January 1996 to June 2007.

	Maturity			
Moneyness	$\tau \leq 50$	$50 < \tau \leq 90$	$90 < \tau \leq 160$	$160 < \tau$
Panel A	ZM-LHARG Implied Volatility RMSE			
$0.8 \leq m \leq 0.9$	14.528	7.809	6.019	5.498
$0.9 < m \leq 0.98$	7.696	5.439	5.578	5.752
$0.98 < m \leq 1.02$	5.322	5.610	6.116	6.564
$1.02 < m \leq 1.1$	5.971	6.282	6.824	7.094
$1.1 < m \leq 1.2$	10.838	5.935	6.601	7.689
Panel B	LHARG-ARJ/ZM-LHARG Implied Volatility RMSE			
$0.8 \leq m \leq 0.9$	0.968	0.988	1.003	0.988
$0.9 < m \leq 0.98$	0.932	0.997	0.987	0.955
$0.98 < m \leq 1.02$	0.983	1.047	1.005	0.962
$1.02 < m \leq 1.1$	1.019	1.077	1.026	0.989
$1.1 < m \leq 1.2$	0.972	1.082	1.048	0.990

Table 4.13: Detailed comparison of option pricing performance on S&P500 Index put options for models ZM-LHARG and LHARG-ARJ during and after the financial crisis: from July 2007 to June 2011.

	Maturity			
Moneyness	$\tau \leq 50$	$50 < \tau \leq 90$	$90 < \tau \leq 160$	$160 < \tau$
Panel A	BPJVM Implied Volatility RMSE			
$0.8 \leq m \leq 0.9$	12.986	5.513	4.714	11.606
$0.9 < m \leq 0.98$	6.332	4.007	4.829	11.181
$0.98 < m \leq 1.02$	3.579	3.950	5.249	11.613
$1.02 < m \leq 1.1$	5.692	4.942	6.784	14.213
$1.1 < m \leq 1.2$	18.487	8.055	10.412	21.191
Panel C	LHARG-ARJ/BPJVM Implied Volatility RMSE			
$0.8 \leq m \leq 0.9$	1.014	1.150	0.952	0.411
$0.9 < m \leq 0.98$	0.912	0.914	0.874	0.557
$0.98 < m \leq 1.02$	0.996	1.086	0.986	0.622
$1.02 < m \leq 1.1$	0.928	0.858	0.775	0.518
$1.1 < m \leq 1.2$	0.991	0.452	0.423	0.298

Table 4.14: Detailed comparison of option pricing performance on S&P500 Index put options for models BPJVM and LHARG-ARJ before the financial crisis: from January 1996 to June 2007.

this case, the result for the central region of volatility surface can be ascribed to the good description of the ATM implied volatility dynamics coming from the heterogeneous dependence of different temporal components.

By observing the detailed Tables 4.14 and 4.15, we note that the RMSE are balanced between the models for short maturities and becomes more evidently in favour of LHARG-ARJ when moving to longer horizons. Similarly to JLHARG class, the better LHARG-ARJ pricing of long maturity options suggests a greater flexibility of the proposed framework to reproduce a realistic term structure of implied volatilities with respect to the BPJVM model. Even in this case, the extremely high persistence of both volatility and jump intensity processes evidenced from estimation of the BPJVM model, as reported in Table 4.9, could be the cause of a systematic miss-pricing of the long-term options, contingently on the prevailing high or low level of present volatilities. Figures 4.3 deepen this point. They show the autocorrelation functions for the observed financial data as returns and realized volatilities (total RV_t, continuous CRV_t and jump JRV_t built

	Maturity			
Moneyness	$\tau \leq 50$	$50 < \tau \leq 90$	$90 < \tau \leq 160$	$160 < \tau$
Panel A	BPJVM Implied Volatility RMSE			
$0.8 \leq m \leq 0.9$	14.409	8.148	7.582	9.060
$0.9 < m \leq 0.98$	8.060	7.051	7.531	9.201
$0.98 < m \leq 1.02$	5.400	6.358	7.200	9.375
$1.02 < m \leq 1.1$	5.334	6.640	7.893	12.460
$1.1 < m \leq 1.2$	9.504	9.398	10.851	14.275
Panel B	LHARG-ARJ/BPJVM Implied Volatility RMSE			
$0.8 \leq m \leq 0.9$	0.875	0.790	0.793	0.557
$0.9 < m \leq 0.98$	0.836	0.758	0.726	0.600
$0.98 < m \leq 1.02$	1.003	0.949	0.871	0.675
$1.02 < m \leq 1.1$	1.177	1.048	0.901	0.552
$1.1 < m \leq 1.2$	0.968	0.720	0.686	0.532

Table 4.15: Detailed comparison of option pricing performance on S&P500 Index put options for models BPJVM and LHARG-ARJ during and after the financial crisis: from July 2007 to June 2011.

in LHARG-ARJ framework and bipower variation BPV_t and related jump variation RJV_t involved in BPJVM framework). They also show the filtered latent variables as the diffusive volatilities $h_{z,t}$ and the jump intensity $h_{y,t}$ used in the BPJVM model and the jump intensity ω_t for LHARG-ARJ model. The figures confirm the typical slow decay of the autocorrelation of empirical volatility measures which stay significantly above the confidence interval for high lag order (up to three months in the figures). Between the two, the continuous component has higher autocorrelations with respect to the discontinuous one. Comparing the function of the filtered BPJVM latent volatility $h_{z,t}$ with its observable RBV_t we note that the first assumes greater values (between 0.4 and 1 for the data period 1990-2007 and between 0.35 and 1 for the data period 2007-2011) than the second one (between 0.35 and 0.8 and between 0.35 and 0.85, respectively) consequently showing a higher persistence than the realized one (0.9999 for both periods as implied by the estimated parameters). Similarly, the jump intensity $h_{y,t}$ has highly persistent autocorrelations with values in the range 0.6-1 for the pre-crisis period and 0.45-1 for the crisis period (values of persistence of 0.987 and 0.998,

respectively). This result show that the BPJVM model is able to capture (and even exceed) the high persistence in the observed volatility but with a value of persistence of almost 1 the model has difficulty in reproducing the mean reversion which is another typical characteristic of the volatility process. This translates in a poorly performance in pricing options with long time to maturities.

In conclusion, the novel models introduced in this work, the JLHARG-RV and the LHARG-ARJ models, have been involved in an empirical assessment of pricing replication on real option data evidencing interesting results. In particular, we find that ZM-JLHARG and LHARG-ARJ models shows a general good pricing performance when compared to a standard reference as the Heston-Nandi GARCH model and the state-of-the-art realized volatility models, the ZM-LHARG and the BPJVM model. A detailed analysis confirm that the inclusion of a jump component is important in order to better reproduce the implied volatilities for out-of-the-money options, specifically for short-maturities. Furthermore, the possibility to account for an heterogeneous structure provides a realistic description of the term structure of the volatility surface implying a good option valuation over different maturities.

4.3 U-shaped log-density ratio

This section establishes if the SDF (3.7), combined with the JLHARG-RV dynamics, implies a U-shaped profile for the log-ratio between the risk-neutral and physical probability densities. This interesting property of a multi-dimensional pricing kernel has been recently investigated in Christoffersen et al. (2013) for a class of GARCH models as summarized in Section 2.4. Authors show that a premium for variance explains a number of puzzles concerning the level and dynamics of option implied variance compared with observed time series variance. Among others, they solve the puzzle pointed out by Bates (1996a) and Broadie et al. (2007) that the physical and risk-neutral volatility smiles differ. The key feature of their modelling approach is that the projection of the pricing kernel onto the stock price return is U-shaped. The strong implied volatility smile associated to this non-monotonic relation can be quantified looking at the natural logarithm

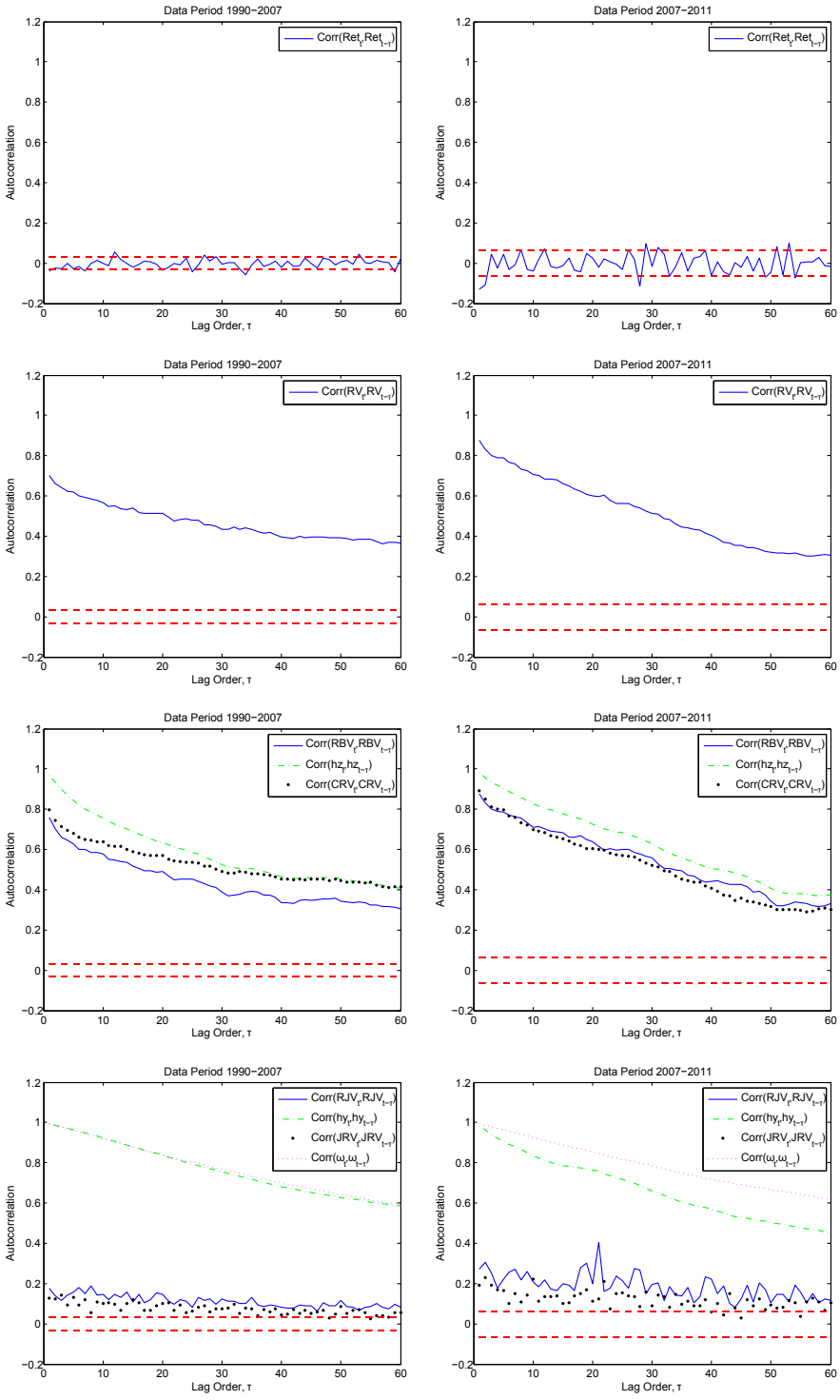

Figure 4.3: Autocorrelations of daily returns and realized volatility measures computed for the two period under study. The last two pairs of figures show the comparison between autocorrelations in the component of realized volatility and the corresponding measures constructed according to Christoffersen et al. (2015).

of the ratio of the risk-neutral and physical conditional densities implied by the model.

We check the non-monotonicity of the log-ratio of the densities implied by the JLHARG class of models. In the GARCH case, the analytical derivation of this property is facilitated by the fact that the variance term at time $t+1$ is an $\mathcal{F}_t$-measurable quantity. Thus, it can be directly projected onto returns at time t. This is clearly not the case for JLHARG models. As it can be seen from (3.4) and (3.6), in the multi-risk premia kernel (3.7) the quantities CRV_{t+1} and JRV_{t+1} are not $\mathcal{F}_t$-measurable, so the analytical projection-onto-returns method by Christoffersen et al. (2013) can not be directly replicated. We proceed as follows. We simulate a large sample of log-returns according to the P-JLHARG dynamics under the physical and risk-neutral measures with variance risk premia fixed by the model calibration. Specifically, we draw 280000 log-returns for four different time-scales – 1 day, 1 month, 3 months, and 6 months – we build the histograms for each time series (Figure 4.4), and we compute their log-ratios. By this method we indirectly enlighten the properties of the pricing kernel in the way it modifies the log-return probability density from the physical to the risk-neutral measure. The blue solid line in Figure 4.5 clearly shows the emergence of a non-monotonic relationship, and confirms the remarkable stability of such U-shaped relation across different time horizons. In Figure 4.5, we also show the results for the density log-ratio in the particular case where both variance risk premia are set equal to zero (green dotted line). We observe that the relation – apart from the noisy behavior in the tail regions – becomes approximately linear. In fact, in the absence of variance risk premia, the effect of the change of measure on the parameter of the RV process is negligible.[2] As a result the risk-neutral dynamics of the realized volatility is very close to the physical one, which is inconsistent with the empirical evidence. These results underline the importance of having a variance dependent SDF in order to reconcile the time series return distribution with that implied by option prices.

[2]This can be seen from formula (3.8) noticing that when $\nu_c = 0$ and $\nu_j = 0$, $\mathrm{y}^{c\,*}$ and $\mathrm{y}^{j\,*}$ are of order one while $\theta \sim 10^{-5}$ so all the denominators are approximately equal to one.

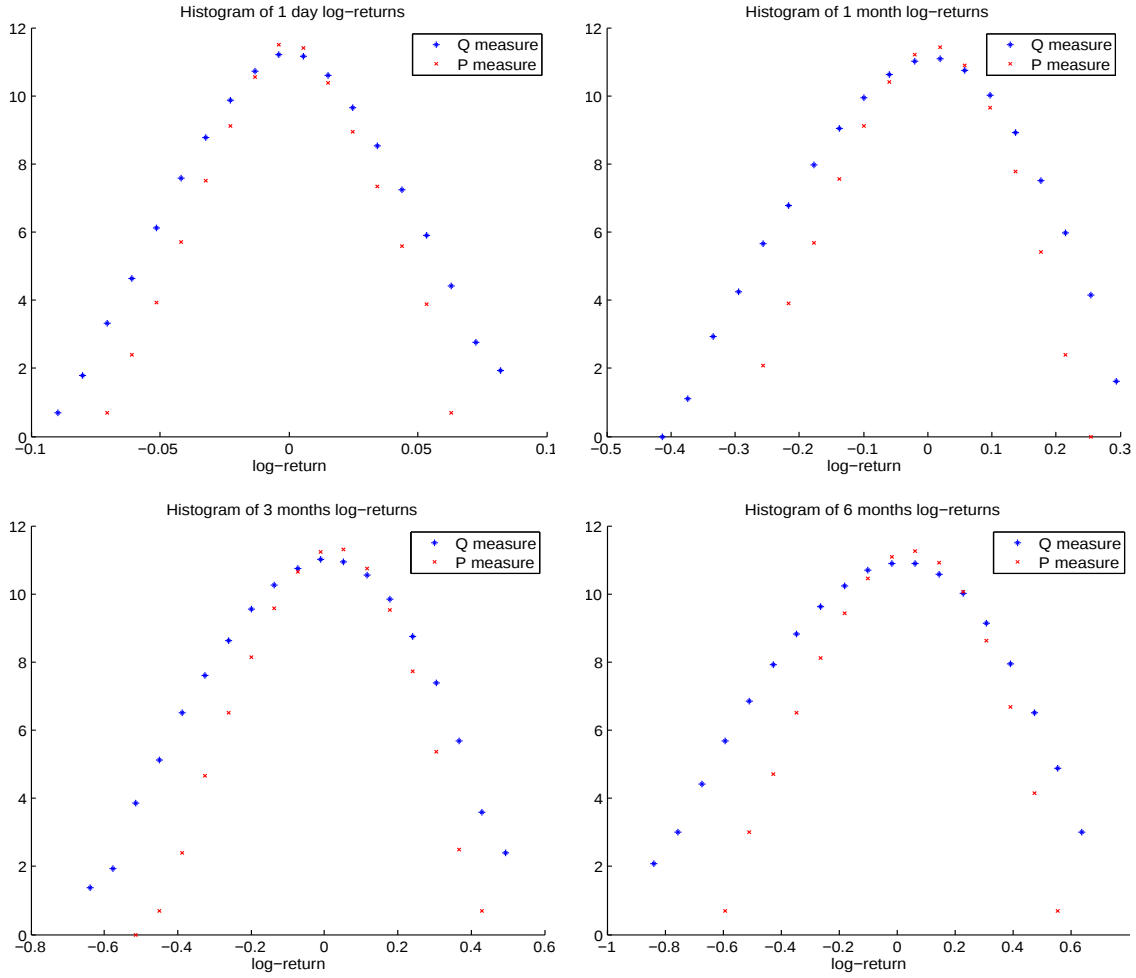

Figure 4.4: Histograms of log-returns over different time horizons. The physical density is obtained by simulating a P-LHARG model with parameters reported in Table 4.1, while the risk-neutral density corresponds to parameter values rescaled according to (3.8). On the y-axis we report log-counts per bin.

4.4 Predictive regressions for stock market returns based on variance risk premium

Several recent studies[3] (Bollerslev et al. (2009b), Drechsler and Yaron (2011), Du and Kapadia (2012), Camponovo et al. (2014), Bollerslev et al. (2015)) have argued that aggregate stock market return is predictable over time horizons ranging from quarterly to annually if the difference between realized and options-implied variation measures, even known as variance risk premium, is considered as ex-

[3]See also Zhou and Yingzi (2009), Zhou (2010), Camponovo et al. (2014), Bekaert and Hoerova (2014), Bandi and Renò (2016), Carr and Wu (2016)

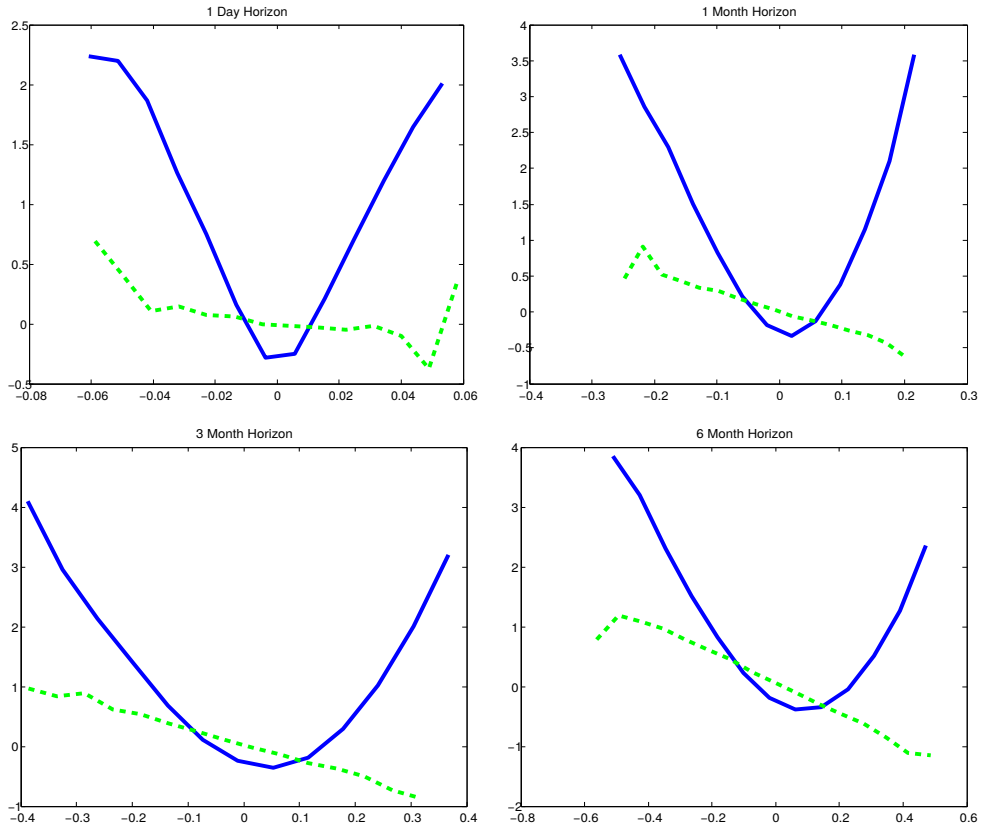

Figure 4.5: Log-ratio of the risk-neutral and physical log-return conditional densities over different time horizons. The physical and risk-neutral densities has been computed proceeding as in Figure 4.4. The blue solid line represents the log-ratio with variance risk premia obtained from calibration, while the green dotted line refers to the case with both premia set equal to zero.

planatory variable. This fact differs from the long-run multi-year return predictability patterns that have been studied extensively[4] and has been typically associated with more traditional valuation measures such as dividend yields, P/E ratios, or consumption-wealth ratios. In this section we advance a test to assess if the proposed LHARG-ARJ-implied VRP, introduced in Section 3.5, is capable to reproduce or improve the predictability of future stock returns relatively to well-established predictors.

We build our model-implied measure of VRP applying the definition in equation (3.32) along with the formulas to calculate expectations of conditional variance contained in Proposition 11. We compute this quantity for a monthly horizon

[4]See among others Fama and French (1988), Campbell and Shiller (1988) and Lettau and Ludvigson (2001)

($T = 22$ days) for last Wednesdays of each month in the period from January 1996 to June 2011. According to the division into two consecutive smaller period (pre-crisis and crisis/post-crisis) of the estimation and calibration on real data as discussed in Section 4.1.2, the model-implied VRP, which we indicate as VRP_{22}^{Mod}, is computed using the two set of parameters and risk-premia (as in Table 4.3) separately on the dates belonging to each sub-period. We compare our newly introduced VRP measure with a "market proxy", or VRP_{22}^{Mkt}, where the conditional expected value of variance over the next month under risk-neutral measure is obtained by the square of the VIX index following Carr and Wu (2006) and Bollerslev et al. (2009b). While for the computation of the market expectation of variance over a monthly horizon under physical measure, we adopt the approach by Bollerslev and Todorov (2011) based on reduced-form time series modelling procedures. For tractability reasons we apply the HAR-RV modelling approach by Corsi (2009) assuming that the model-implied estimates of conditional expectations are reasonable proxy for market-implied expectations. The formal expression of this quantity is given by the following

$$\text{VRP}^{Mkt}(t, t+22) = \frac{1}{22} \sum_{i=1}^{22} \mathbb{E}^{\mathbb{P}} \left[\text{RV}_{t+i}^{(HAR)} \right] - \text{VIX}_t^2. \tag{4.5}$$

Furthermore, we define another measure for the variance risk premium constructed by mixing the market-implied risk-neutral expectation of future variances with the model-implied expectation based on historical or physical information. We indicate it as a semi-parametric variance risk premium $\text{VRP}_{22}^{SemiPar}$ and is expressed by

$$\text{VRP}^{SemiPar}(t, t+22) = \frac{1}{22} \sum_{i=1}^{22} \mathbb{E}^{\mathbb{P}} \left[\text{RV}_{t+i}^{(LHARG-ARJ)} \right] - \text{VIX}^2. \tag{4.6}$$

In Figure 4.6 we present the square of the VIX index and the results of the conditional expected variance under $\mathbb{Q}$ measure computed using the LHARG-ARJ model. We can see that the model (black line) behaves similarly to the squared VIX (red dashed line) but appears to slightly reproduce the dynamics of the market-implied data in particular during period of rapid swings of volatilities.

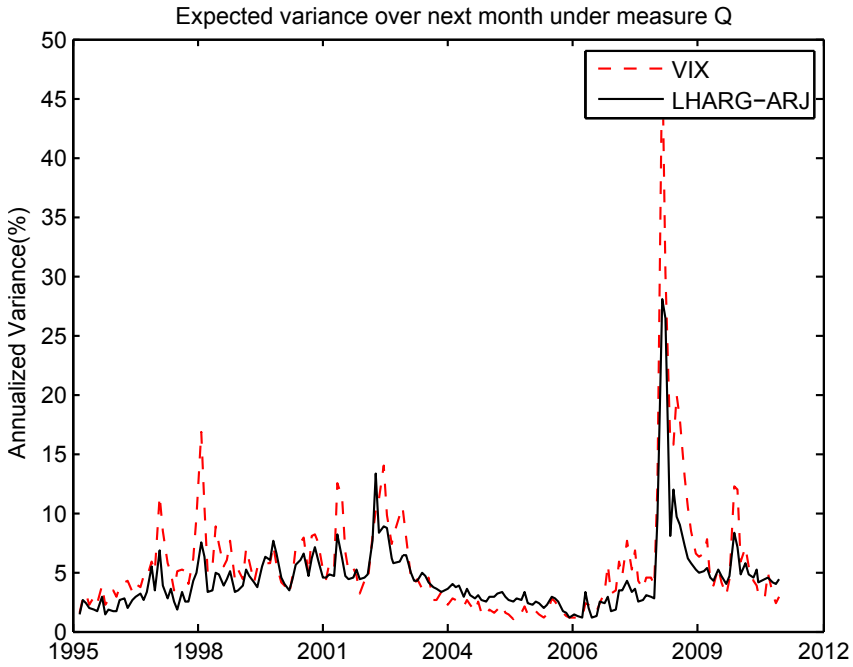

Figure 4.6: Expected variance over next month from 1996 to 2011. The figure compares value of squared VIX index with expected variance over next month under risk-neutral measure resulting from LHARG-ARJ model.

Furthermore our method allows to easily separate the contributions to expected variance under $\mathbb{Q}$ coming from the continuous and jump dynamics of RV. Figure 4.7 shows the time evolution of the two separated components. We note that the continuous contribution appears to be the bigger of the two and carry much of the time variability, while on the other hand the jump contribution is almost stable during the all period. Figure 4.8 collects the results for the market and model conditional expectation of variances over a month under $\mathbb{P}$ measure. It evidences that the modelled quantities follows closely the time evolution of the expected variance obtained from a HAR-RV model, proxy of the real objective dynamics of variance. In Figure 4.9 we show the VRP measurements obtained for each Wednesday in our sample. As we can see the modelled measure is still less dynamical than the market measure with difficulties to significantly reproduce the highest values of the variance premia estimated for particularly turbulent periods for financial market. On the other hand, the comparison between the market-implied VRP and the semi-parametric estimate indicates that these two measure are closer with the second one able to correctly capture the succession of periods of different perceived uncertainty. This result implies that the lack of

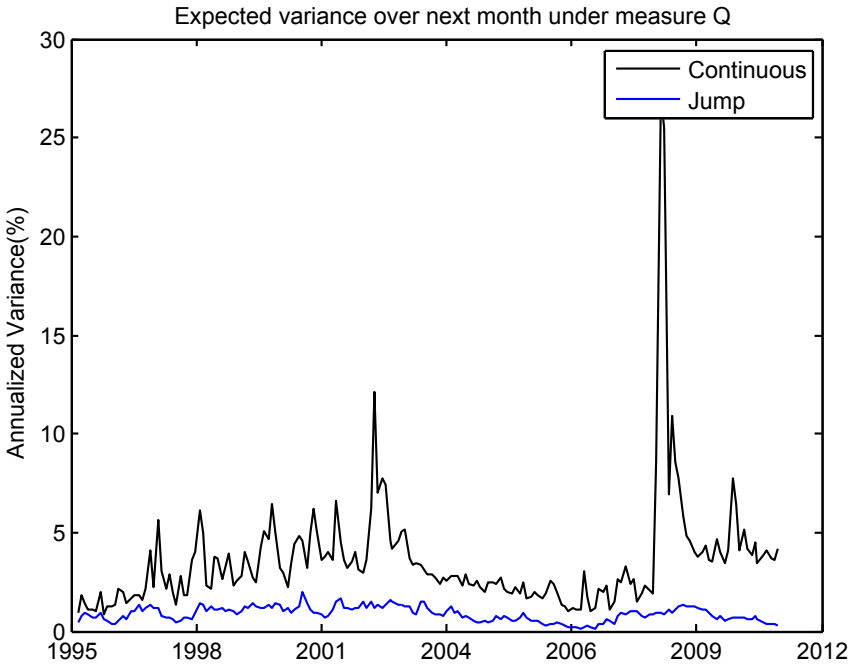

Figure 4.7: Expected variance over next month from 1996 to 2005. Continuous (black line) and jump (blue line) component of the expected variance over next month under risk-neutral measure resulting from LHARG-ARJ model.

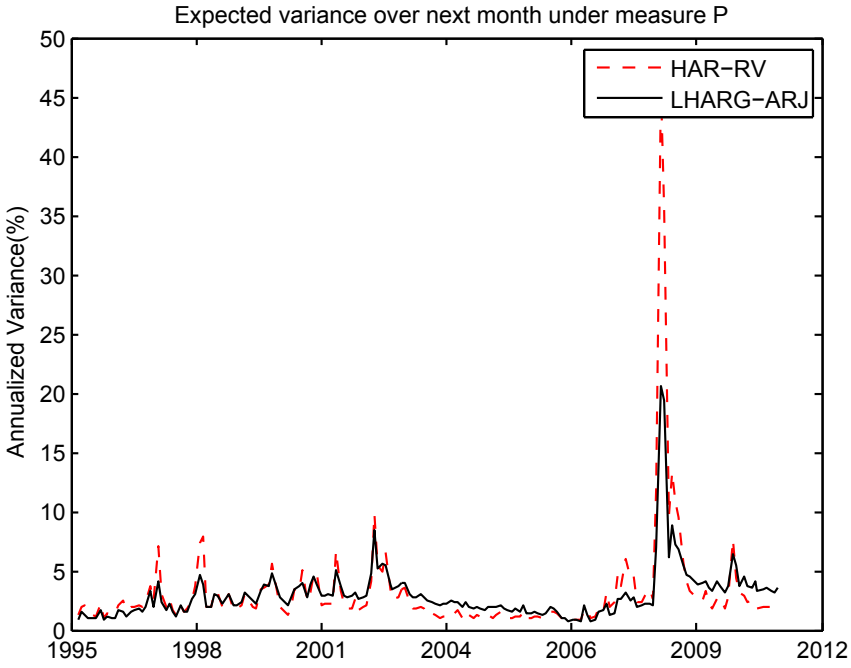

Figure 4.8: Expected variance over next month from 1996 to 2011. The figure compares the expected variance over next month under physical measure resulting from HAR-RV and LHARG-ARJ models.

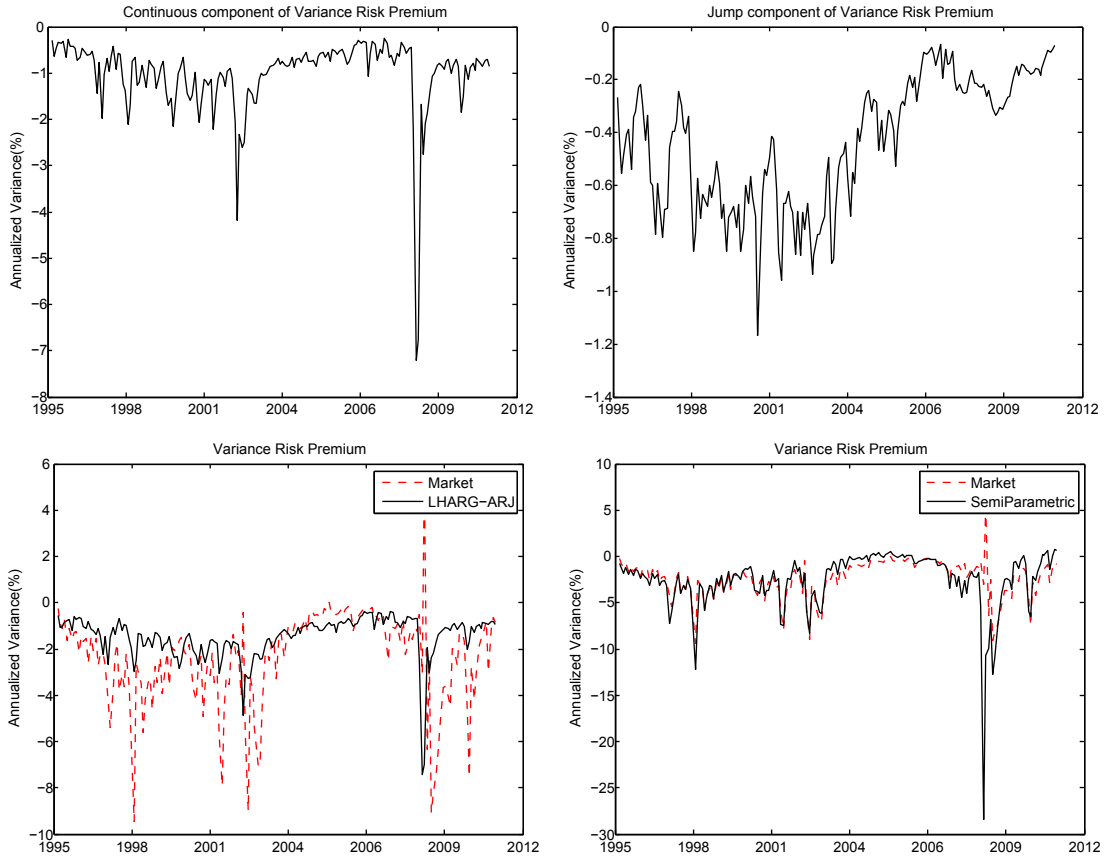

Figure 4.9: Monthly horizon Variance Risk Premium from 1996 to 2011. Top left figure shows the continuous component of VRP computed from the proposed LHARG-ARJ model. Top right figure shows the model implied jump component of VRP. Bottom left figure: The red dashed line represents the proxy chosen for the market VRP while the black solid line indicates the model-implied VRP measure. Bottom right figure: The black solid line represents the semi-parametric measure of VRP.

sufficient dynamics of the model-implied VRP can be attributed to the conditional expected variance under risk-neutral measure which is affected by a rigid change of measure due to static premia in the adopted stochastic discount factor (see equation (3.25)).

Generally, the VRP values are negative meaning that normally the risk neutral expected variances exceed the ones computed from historical information. Top figures in 4.9 contain separately the two contribution to modelled VRP, denoted $\text{VRP}_{22,c}^{Mod}$ and $\text{VRP}_{22,j}^{Mod}$. We observe that the typical order of magnitude of the continuous component is higher (almost a factor 2) than the jump component. Furthermore, it is interesting to observe that during the crisis period the jump

VRP assumes its lowest values while the continuous VRP has its highest pick. This could indicate that during distressed moments the contribution to the premium requested to compensate the perceived uncertainty about future volatilities is largely due to the diffusive dynamics while the other contribution related to extreme unpredictable movements becomes negligible.

In what follows, we present some empirical findings of a study of predictability involving the VRP measures defined above using the market "proxy", the LHARG-ARJ semi and full parametric specification and its decomposition in continuous and jump contributions, along with more traditional predictor variables of returns. This class includes Cyclically Adjusted Price Earning ratio (CAPE) introduced in Schiller (2000) and the slope of Treasury yield curve, also known as Term Spread (TMSP), defined as the difference between the ten-year T-bond and the three month T-bill yields. All of our forecasts are based on simple linear regression of the S&P 500 Index excess log-returns on different sets of predictor variables. We always use monthly observed predictors while we choose two aggregation horizons for returns, quarterly and annually. For all regression coefficient we report Newey-West t-statistics taking into account for heteroskedasticity and serial correlation from the overlap in the regressions. We also report t-statistics computed according to Hodrick (1992) to assess the significance of coefficients in overlapping return regressions with persistent predictor variables.

We choose to study the phenomenon of predictability by running the regressions over two time period according to what previously described in Section 4.1.2. One is referred as "pre-crisis", starting in January 1996 and ending in June 2007, a year before the turmoil of the summer/fall 2008 Financial Crisis. The other one coincides with the period going from July 2007 up to June 2011, which includes the crisis and the following downturn. This choice has been made in order to investigate if statistically significant differences emerges from the data in the behaviour of predictors. The results of the regression for the Pre-crisis period are presented in Tables 4.16-4.18 and 4.19-4.21.

Quarterly Horizon									
regressor:	constant	CAPE	TMSP	VRP_{22}^{Mkt}	$VRP_{22}^{SemiPar}$	VRP_{22}^{Mod}	$VRP_{22,c}^{Mod}$	$VRP_{22,j}^{Mod}$	$\hat{R}^2$
Coefficient	-3.676	-	-	-2.680	-	-	-	-	0.030
H	-0.832	-	-	-1.569	-	-	-	-	0.117
NW	-0.829	-	-	-1.589	-	-	-	-	0.112
Coefficient	-4.633	-	-	-	-3.277	-	-	-	0.060
H	-1.158	-	-	-	-2.291	-	-	-	0.022
NW	-1.078	-	-	-	-2.293	-	-	-	0.022
Coefficient	4.669	-	-	-	-	1.522	-	-	0.002
H	0.643	-	-	-	-	0.314	-	-	0.753
NW	0.840	-	-	-	-	0.386	-	-	0.700
Coefficient	3.235	-	-	-	-	-	0.839	-	0.000
H	0.529	-	-	-	-	-	0.137	-	0.891
NW	0.658	-	-	-	-	-	0.178	-	0.859
Coefficient	7.578	-	-	-	-	-	-	10.087	0.007
H	0.942	-	-	-	-	-	-	0.687	0.492
NW	1.210	-	-	-	-	-	-	0.723	0.470

Table 4.16: Results from regressions of quarterly excess stock market returns on different predictors with data for the "Pre-Crisis" period. 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. VRP_{22}^{Mkt} corresponds to the market proxy obtained with equation (4.5) and VRP_{22}^{Mod} to the model-implied measurement obtained with equation (3.32). $VRP_{22,c}^{Mod}$ and $VRP_{22,j}^{Mod}$ are in order the continuous and jump component of VRP computed isolating the two contribution in (3.32). Below the coefficients of regressions we report the Newey-West (NW) and Hodrick (H) t-statistics. In the last column, we report the adjusted R^2 and for multiple regressions NW and H p-values resulting from the $\chi^2(n)$ $(n \geq 2)$ test that all coefficients are jointly equal to zero.

Quarterly Horizon

regressor:	constant	CAPE	TMSP	VRP_{22}^{Mkt}	$VRP_{22}^{SemiPar}$	VRP_{22}^{Mod}	$VRP_{22,c}^{Mod}$	$VRP_{22,j}^{Mod}$	$\hat{R}^2$
Coefficient	24.952	-0.741	-	-	-	-	-	-	0.029
H	1.527	-1.351	-	-	-	-	-	-	0.177
NW	1.858	-1.688	-	-	-	-	-	-	0.091
Coefficient	25.123	-1.019	-	-3.648	-	-	-	-	0.081
H	1.536	-1.770	-	-2.026	-	-	-	-	0.063
NW	1.866	-2.246	-	-2.118	-	-	-	-	0.025
Coefficient	27.615	-1.129	-	-	-4.246	-	-	-	0.122
H	1.671	-1.949	-	-	-2.795	-	-	-	0.011
NW	2.001	-2.431	-	-	-2.868	-	-	-	0.005
Coefficient	25.057	-0.737	-	-	-	0.144	-	-	0.029
H	1.544	-1.284	-	-	-	0.028	-	-	0.396
NW	1.867	-1.630	-	-	-	0.035	-	-	0.240
Coefficient	24.729	-0.746	-	-	-	-	-0.394	-	0.029
H	1.518	-1.332	-	-	-	-	-0.063	-	0.401
NW	1.821	-1.686	-	-	-	-	-0.079	-	0.239
Coefficient	25.607	-0.694	-	-	-	-	-	4.057	0.030
H	1.588	-1.146	-	-	-	-	-	0.250	0.361
NW	1.954	-1.410	-	-	-	-	-	0.266	0.217

Table 4.17: Continued from Table 4.16.

Quarterly Horizon

regressor:	constant	CAPE	TMSP	VRP_{22}^{Mkt}	$VRP_{22}^{SemiPar}$	VRP_{22}^{Mod}	$VRP_{22,c}^{Mod}$	$VRP_{22,j}^{Mod}$	$\hat{R}^2$
Coefficient	39.169	-1.047	-3.814	-	-	-	-	-	0.051
H	2.089	-1.825	-1.118	-	-	-	-	-	0.163
NW	2.497	-2.222	-1.102	-	-	-	-	-	0.078
Coefficient	46.793	-1.559	-5.802	-4.613	-	-	-	-	0.130
H	2.424	-2.518	-1.642	-2.478	-	-	-	-	0.026
NW	2.836	-2.897	-1.599	-2.591	-	-	-	-	0.017
Coefficient	44.439	-1.507	-4.481	-	-4.444	-	-	-	0.154
H	2.327	-2.472	-1.308	-	-2.912	-	-	-	0.011
NW	2.759	-2.896	-1.321	-	-3.145	-	-	-	0.006
Coefficient	39.951	-1.173	-4.556	-	-	-2.723	-	-	0.056
H	2.116	-1.854	-1.238	-	-	-0.498	-	-	0.278
NW	2.530	-2.256	-1.155	-	-	-0.548	-	-	0.145
Coefficient	39.192	-1.120	-4.255	-	-	-	-2.858	-	0.055
H	2.090	-1.872	-1.207	-	-	-	-0.442	-	0.281
NW	2.507	-2.284	-1.146	-	-	-	-0.513	-	0.148
Coefficient	42.069	-1.265	-5.042	-	-	-	-	-10.386	0.056
H	2.133	-1.751	-1.243	-	-	-	-	-0.539	0.280
NW	2.476	-2.030	-1.104	-	-	-	-	-0.502	0.143

Table 4.18: Continued from Table 4.16 and 4.17.

Annual Horizon

regressor:	constant	CAPE	TMSP	VRP_{22}^{Mkt}	$VRP_{22}^{SemiPar}$	VRP_{22}^{Mod}	$VRP_{22,c}^{Mod}$	$VRP_{22,j}^{Mod}$	$\hat{R}^2$
Coefficient	0.026	-	-	-0.479	-	-	-	-	0.003
H	0.004	-	-	-0.307	-	-	-	-	0.758
NW	0.005	-	-	-0.294	-	-	-	-	0.768
Coefficient	-1.071	-	-	-	-1.015	-	-	-	0.015
H	-0.216	-	-	-	-0.856	-	-	-	0.392
NW	-0.266	-	-	-	-0.914	-	-	-	0.361
Coefficient	5.799	-	-	-	-	3.182	-	-	0.018
H	0.660	-	-	-	-	0.663	-	-	0.507
NW	0.744	-	-	-	-	0.659	-	-	0.510
Coefficient	5.099	-	-	-	-	-	4.144	-	0.019
H	0.749	-	-	-	-	-	0.820	-	0.412
NW	0.765	-	-	-	-	-	0.701	-	0.483
Coefficient	4.633	-	-	-	-	-	-	6.890	0.009
H	0.414	-	-	-	-	-	-	0.368	0.713
NW	0.555	-	-	-	-	-	-	0.449	0.653

Table 4.19: Results from regressions of annual excess stock market returns on different predictors with data for the "Pre-Crisis" period.

Annual Horizon

regressor:	constant	CAPE	TMSP	VRP_{22}^{Mkt}	$VRP_{22}^{SemiPar}$	VRP_{22}^{Mod}	$VRP_{22,c}^{Mod}$	$VRP_{22,j}^{Mod}$	$\hat{R}^2$
Coefficient	28.639	-0.905	-	-	-	-	-	-	0.113
H	1.488	-1.396	-	-	-	-	-	-	0.163
NW	1.979	-1.840	-	-	-	-	-	-	0.066
Coefficient	28.706	-1.015	-	-1.442	-	-	-	-	0.135
H	1.493	-1.576	-	-0.987	-	-	-	-	0.191
NW	2.106	-2.310	-	-0.980	-	-	-	-	0.039
Coefficient	29.858	-1.083	-	-	-1.945	-	-	-	0.165
H	1.554	-1.655	-	-	-1.732	-	-	-	0.080
NW	2.249	-2.488	-	-	-1.892	-	-	-	0.010
Coefficient	29.776	-0.867	-	-	-	1.562	-	-	0.118
H	1.427	-1.407	-	-	-	0.353	-	-	0.371
NW	1.909	-1.828	-	-	-	0.404	-	-	0.181
Coefficient	30.172	-0.870	-	-	-	-	2.706	-	0.121
H	1.471	-1.384	-	-	-	-	0.581	-	0.376
NW	1.962	-1.819	-	-	-	-	0.610	-	0.178
Coefficient	28.464	-0.917	-	-	-	-	-	-1.080	0.114
H	1.380	-1.508	-	-	-	-	-	-0.061	0.319
NW	1.886	-1.878	-	-	-	-	-	-0.077	0.168

Table 4.20: Continued from Table 4.19.

Annual Horizon

regressor:	constant	CAPE	TMSP	VRP_{22}^{Mkt}	$VRP_{22}^{SemiPar}$	VRP_{22}^{Mod}	$VRP_{22,c}^{Mod}$	$VRP_{22,j}^{Mod}$	$\hat{R}^2$
Coefficient	31.719	-0.971	-0.827	-	-	-	-	-	0.116
H	1.144	-1.234	-0.215	-	-	-	-	-	0.373
NW	1.653	-1.723	-0.264	-	-	-	-	-	0.189
Coefficient	34.532	-1.160	-1.560	-1.702	-	-	-	-	0.145
H	1.262	-1.491	-0.423	-1.430	-	-	-	-	0.239
NW	1.959	-2.410	-0.540	-1.385	-	-	-	-	0.015
Coefficient	34.085	-1.178	-1.126	-	-1.994	-	-	-	0.171
H	1.239	-1.509	-0.298	-	-1.888	-	-	-	0.122
NW	1.911	-2.364	-0.374	-	-2.111	-	-	-	0.008
Coefficient	31.358	-0.913	-0.484	-	-	1.257	-	-	0.119
H	1.150	-1.252	-0.136	-	-	0.366	-	-	0.576
NW	1.663	-1.778	-0.161	-	-	0.368	-	-	0.319
Coefficient	31.699	-0.909	-0.449	-	-	-	2.445	-	0.123
H	1.144	-1.206	-0.122	-	-	-	0.648	-	0.566
NW	1.645	-1.695	-0.145	-	-	-	0.594	-	0.334
Coefficient	33.176	-1.081	-1.443	-	-	-	-	-5.214	0.120
H	1.288	-1.570	-0.442	-	-	-	-	-0.372	0.403
NW	1.932	-2.215	-0.508	-	-	-	-	-0.411	0.159

Table 4.21: Continued from Table 4.19 and 4.20.

As concerns the quarterly forecast horizon, we note that the market and the model VRP have not significant predictive power as alone regressor while the semi-parametric measure of monthly VRP has higher $t-$ratios statistically significant below the 5% confidence level. The $\hat{R}^2$ are particularly low for model-implied measures (below 1%) and are slightly higher for VRP_{22}^{Mkt} and $\text{VRP}_{22}^{SemiPar}$ with values of 3% and 6% respectively.

The predictability appears significantly enhanced when the monthly VRP measures are used in a multiple regressions scheme with CAPE as further predictor. The market VRP shows more significant $t-$statistics and higher $\hat{R}^2$ values with respect to the previous regression with associated $p-$values of 6.3% and 2.5% and adjusted $\hat{R}^2$ of 8.1%. Considering the semi-parametric VRP, we note a improvement both in the significance level ($t-$ratios of -2.795 and-2.868) and the forecast power represented by a 12.2% $\hat{R}^2$. On the other hand, nor the model-implied VRP nor its components are significant predictors of future returns.

Finally if both CAPE and TMSP are involved in the regression, all the VRP measures increase their $t-$ratios and $\hat{R}^2$. The results confirm a better performance in terms of greater $\hat{R}^2$, significance of t-statistics and p-value both for the market and semi-parametric measures of VRP. Nevertheless, despite the improvement, the model computed VRP VRP_{22}^{Mod} and its components remain below the threshold of statistical significance for forecasting.

Moving to the annual horizon, we observe that the only schemes that evidence some predictability of returns is that one involving the semi-parametric monthly VRP used in the multiple regression with both CAPE and TMSP where the $t-$ratio (whose value is -2.111) is computed according to Newey-West estimator. In all other regressions, p-values are higher than the confidence level of 5% indicating that generally the predictability of the VRP measures fades away over longer horizons. So, for the Pre-crisis period we discover that the monthly VRP measures, both the market and the semi-parametric ones, can be used as predictors of market returns in some multiple regression schemes together with well-established forecasting quantities such as CAPE and TMSP, while the modelled implied measures appear to have no forecast power in all of the studied regression schemes.

Now we describe the results of the assessment of predictive property of VRP measures applying the same approach used before to data belonging to the time window including the crisis. The 2007-2008 Financial Crisis has been indicated as the biggest postwar era downturn. In similar periods, radical changes can occur in the way market agents perceive and manage financial risks, impacting the compensations, or risk premia, attached to them. In particular the risk premium related to the variance could be significantly affected during a period of distress. Bollerslev et al. (2015) find that a significant VRP predictability comes from what is named as "tail risk" or the premium that market participate request to bare the risk of extreme events. In tables 4.22-4.24 and 4.25-4.27 we report the findings of the multiple regression scheme for predictability.

Quarterly Horizon									
regressor:	constant	CAPE	TMSP	VRP_{22}^{Mkt}	$VRP_{22}^{SemiPar}$	VRP_{22}^{Mod}	$VRP_{22,c}^{Mod}$	$VRP_{22,j}^{Mod}$	$\hat{R}^2$
Coefficient	4.118	-	-	-5.304	-	-	-	-	0.140
H	0.432	-	-	-2.178	-	-	-	-	0.029
NW	0.515	-	-	-2.370	-	-	-	-	0.018
Coefficient	2.758	-	-	-	-4.836	-	-	-	0.179
H	0.290	-	-	-	-2.747	-	-	-	0.006
NW	0.349	-	-	-	-3.438	-	-	-	0.001
Coefficient	0.160	-	-	-	-	-14.292	-	-	0.109
H	0.011	-	-	-	-	-1.710	-	-	0.087
NW	0.014	-	-	-	-	-1.913	-	-	0.056
Coefficient	4.374	-	-	-	-	-	-17.963	-	0.106
H	0.408	-	-	-	-	-	-1.990	-	0.047
NW	0.446	-	-	-	-	-	-1.952	-	0.051
Coefficient	1.843	-	-	-	-	-	-	-33.520	0.058
H	0.080	-	-	-	-	-	-	-0.912	0.362
NW	0.140	-	-	-	-	-	-	-1.389	0.165

Table 4.22: Results from regressions of quarterly excess stock market returns on different predictors with data belonging to the "Post-crisis" period.

Quarterly Horizon

regressor:	constant	CAPE	TMSP	VRP_{22}^{Mkt}	$VRP_{22}^{SemiPar}$	VRP_{22}^{Mod}	$VRP_{22,c}^{Mod}$	$VRP_{22,j}^{Mod}$	$\hat{R}^2$
Coefficient	40.082	-0.681	-	-	-	-	-	-	0.019
H	1.030	-0.539	-	-	-	-	-	-	0.590
NW	1.388	-0.729	-	-	-	-	-	-	0.466
Coefficient	83.545	-2.956	-	-10.595	-	-	-	-	0.373
H	1.981	-1.970	-	-4.053	-	-	-	-	0.000
NW	3.047	-2.978	-	-5.709	-	-	-	-	0.000
Coefficient	63.434	-2.158	-	-	-7.227	-	-	-	0.338
H	1.549	-1.497	-	-	-3.567	-	-	-	0.001
NW	2.373	-2.350	-	-	-5.578	-	-	-	0.000
Coefficient	50.230	-1.907	-	-	-	-22.421	-	-	0.231
H	1.225	-1.167	-	-	-	-1.913	-	-	0.155
NW	1.882	-1.930	-	-	-	-2.677	-	-	0.026
Coefficient	57.020	-1.921	-	-	-	-	-28.535	-	0.229
H	1.361	-1.241	-	-	-	-	-2.365	-	0.055
NW	2.013	-1.871	-	-	-	-	-2.700	-	0.025
Coefficient	33.852	-1.176	-	-	-	-	-	-44.193	0.112
H	0.899	-0.774	-	-	-	-	-	-1.004	0.595
NW	1.293	-1.239	-	-	-	-	-	-1.756	0.199

Table 4.23: Continued from Table 4.22.

Quarterly Horizon

regressor:	constant	CAPE	TMSP	VRP_{22}^{Mkt}	$VRP_{22}^{SemiPar}$	VRP_{22}^{Mod}	$VRP_{22,c}^{Mod}$	$VRP_{22,j}^{Mod}$	$\hat{R}^2$
Coefficient	65.528	-1.268	-7.580	-	-	-	-	-	0.030
H	1.086	-0.777	-0.569	-	-	-	-	-	0.736
NW	1.032	-0.822	-0.444	-	-	-	-	-	0.677
Coefficient	67.067	-2.616	5.329	-10.939	-	-	-	-	0.389
H	1.113	-1.581	0.355	-3.609	-	-	-	-	0.001
NW	2.195	-2.861	0.549	-5.384	-	-	-	-	0.000
Coefficient	42.176	-1.713	6.674	-	-7.583	-	-	-	0.354
H	0.682	-1.047	0.427	-	-3.129	-	-	-	0.003
NW	1.285	-1.923	0.650	-	-5.118	-	-	-	0.000
Coefficient	60.767	-2.139	-3.174	-	-	-22.163	-	-	0.239
H	1.008	-1.204	-0.225	-	-	-1.829	-	-	0.215
NW	1.289	-1.737	-0.254	-	-	-2.628	-	-	0.040
Coefficient	53.615	-1.849	1.052	-	-	-	-28.747	-	0.235
H	0.885	-1.096	0.070	-	-	-	-2.156	-	0.097
NW	1.193	-1.582	0.082	-	-	-	-2.504	-	0.056
Coefficient	75.271	-2.200	-12.532	-	-	-	-	-48.819	0.142
H	1.217	-1.097	-0.925	-	-	-	-	-1.089	0.638
NW	1.256	-1.351	-0.826	-	-	-	-	-2.083	0.219

Table 4.24: Continued from Table 4.22 and 4.23.

Annual Horizon

regressor:	constant	CAPE	TMSP	VRP_{22}^{Mkt}	$VRP_{22}^{SemiPar}$	VRP_{22}^{Mod}	$VRP_{22,c}^{Mod}$	$VRP_{22,j}^{Mod}$	$\hat{R}^2$
Coefficient	20.739	-	-	1.174	-	-	-	-	0.078
H	2.350	-	-	0.647	-	-	-	-	0.517
NW	5.850	-	-	1.423	-	-	-	-	0.155
Coefficient	19.917	-	-	-	0.725	-	-	-	0.046
H	2.409	-	-	-	0.557	-	-	-	0.577
NW	5.330	-	-	-	1.132	-	-	-	0.258
Coefficient	24.076	-	-	-	-	5.082	-	-	0.157
H	2.590	-	-	-	-	1.154	-	-	0.248
NW	5.337	-	-	-	-	1.941	-	-	0.052
Coefficient	22.082	-	-	-	-	-	5.756	-	0.124
H	2.578	-	-	-	-	-	0.941	-	0.347
NW	5.488	-	-	-	-	-	1.676	-	0.094
Coefficient	25.484	-	-	-	-	-	-	15.959	0.149
H	2.716	-	-	-	-	-	-	1.372	0.170
NW	6.725	-	-	-	-	-	-	2.698	0.007

Table 4.25: Results from regressions of annual excess stock market returns on different predictors with data belonging to the "Post-crisis" period.

Annual Horizon

regressor:	constant	CAPE	TMSP	VRP_{22}^{Mkt}	$VRP_{22}^{SemiPar}$	VRP_{22}^{Mod}	$VRP_{22,c}^{Mod}$	$VRP_{22,j}^{Mod}$	$\hat{R}^2$
Coefficient	48.553	-0.977	-	-	-	-	-	-	0.452
H	1.736	-1.210	-	-	-	-	-	-	0.226
NW	6.683	-4.579	-	-	-	-	-	-	0.000
Coefficient	52.384	-1.178	-	-0.934	-	-	-	-	0.496
H	2.020	-1.678	-	-1.187	-	-	-	-	0.010
NW	7.807	-5.848	-	-2.690	-	-	-	-	0.000
Coefficient	50.048	-1.072	-	-	-0.463	-	-	-	0.479
H	1.886	-1.498	-	-	-0.741	-	-	-	0.018
NW	7.234	-5.385	-	-	-2.789	-	-	-	0.000
Coefficient	48.012	-0.912	-	-	-	1.196	-	-	0.471
H	1.731	-1.156	-	-	-	0.439	-	-	0.473
NW	6.814	-4.528	-	-	-	0.734	-	-	0.000
Coefficient	48.257	-0.955	-	-	-	-	0.500	-	0.465
H	1.786	-1.296	-	-	-	-	0.162	-	0.382
NW	6.958	-4.871	-	-	-	-	0.250	-	0.000
Coefficient	49.666	-0.889	-	-	-	-	-	7.895	0.498
H	1.808	-1.022	-	-	-	-	-	0.598	0.270
NW	7.271	-4.436	-	-	-	-	-	1.993	0.000

Table 4.26: Continued from Table 4.25.

Annual Horizon

regressor:	constant	CAPE	TMSP	VRP_{22}^{Mkt}	$VRP_{22}^{SemiPar}$	VRP_{22}^{Mod}	$VRP_{22,c}^{Mod}$	$VRP_{22,j}^{Mod}$	$\hat{R}^2$
Coefficient	39.982	-0.779	2.553	-	-	-	-	-	0.477
H	1.907	-1.457	0.222	-	-	-	-	-	0.213
NW	2.167	-1.728	0.549	-	-	-	-	-	0.000
Coefficient	40.149	-0.926	3.957	-1.189	-	-	-	-	0.540
H	1.917	-1.795	0.337	-1.996	-	-	-	-	0.026
NW	2.813	-2.857	0.958	-2.668	-	-	-	-	0.000
Coefficient	37.934	-0.818	3.803	-	-0.665	-	-	-	0.519
H	1.736	-1.580	0.319	-	-1.235	-	-	-	0.035
NW	2.375	-2.256	0.852	-	-2.162	-	-	-	0.000
Coefficient	40.198	-0.740	2.354	-	-	1.004	-	-	0.496
H	1.874	-1.592	0.199	-	-	0.328	-	-	0.320
NW	2.085	-1.596	0.488	-	-	0.527	-	-	0.000
Coefficient	39.975	-0.779	2.558	-	-	-	-0.016	-	0.490
H	1.818	-1.598	0.213	-	-	-	-0.005	-	0.192
NW	2.126	-1.794	0.524	-	-	-	-0.007	-	0.000
Coefficient	38.149	-0.604	3.485	-	-	-	-	9.182	0.536
H	1.945	-1.268	0.312	-	-	-	-	0.799	0.360
NW	1.959	-1.214	0.690	-	-	-	-	1.666	0.000

Table 4.27: Continued from Table 4.25 and 4.26.

Tables 4.22-4.24 show that for a quarterly horizon forecast most of the regression scheme we involve are predictive. Both the market and the semi-parametric VRP show significant forecasting power as stand-alone regressors (in particular $\text{VRP}_{22}^{SemiPar}$ with $t-$ratios of -2.747 and -3.438 and $\hat{R}^2 = 17.9\%$). Indeed, the scheme with VRP_{22}^{Mod} is barely significant with $p-$values of 8.7% and 5.6% and $\hat{R}^2 = 10.9\%$. By observing its components we note that much of this possible predictability is associated with the continuous $\text{VRP}_{22,c}^{Mod}$ which exhibits $t-$ratios of -1.990 and -1.952 while the jump contribution $\text{VRP}_{22,j}^{Mod}$ is statistically insignificant.

The multi-regression schemes in which the different estimates of the variance risk premium are combined with the traditional predictors confirm a strong degree of predictability for most of them. In combination with CAPE, the market VRP is the most significant with the adjusted $\hat{R}^2$ of 37.3%, $t-$ratios of -4.053 and -5.709 and corresponding $p-$values below the 0.1% level. It follows the semi-parametric $\text{VRP}_{22}^{SemiPar}$ with importantly significance of its regression statistics: $\hat{R}^2$ of 33.8%, $t-$ratios of -3.567 and -5.578 statistically significant at the 0.1% level or better. As concerns the model-implied VRP measure, we find an improvement of its forecasting performance with $t-$ratios of -1.913 and -2.677 and $\hat{R}^2$ equal to 23.1%. Even in this case the enhanced statistics can be explained in terms of the continuous component $\text{VRP}_{22,c}^{Mod}$ while the contribution coming from the extreme movements is still not significant. Finally, by adding the term spread predictor, we obtain results that resemble the previous one with the difference of slightly higher $\hat{R}^2$ due to the presence of one more explanatory variable which however is statistically not significant.

Tables 4.25-4.27 collect results of the predictive regressions applied for annual returns. We see that most of the predictors do not preserve statistical significance. The few regressions which evidence predictability are those involving the market VRP_{22}^{Mkt} along with CAPE and TMSP. With caution, even $\text{VRP}_{22}^{SemiPar}$ shows some forecasting power but only if the Newey-West estimator is involved with $t-$ratios below the significance level of -1.96 (confidence of 5%) while the corresponding ratios computed with Hodrick estimator are above that level.

In conclusion, our empirical findings support in large part what has been recently

advocated by an increasing number of papers in literature: the variance risk premium seems to be a good predictor of stock market returns. This property is particularly evident and stable among different past periods for medium/short time horizons (quarterly) and disappears shifting towards longer horizon (annual). In particular, we find that the predictability emerges even adopting a semi-parametric measure of VRP based on our proposed model LHARG-ARJ used to compute analytically an estimate of the conditional expected variance under physical measure. On the other hand, a full model-implied VRP with both the physical and the risk-neutral expectations of future variances computed according to LHARG-ARJ specification has been proven to not fully evidence property. Nevertheless, some of the performed regressions shows that the continuous component of the model-implied VRP could be a possible predictor. A possible reason for the lack of predictability could rely on the particular choice to assume constant the variance premia in the stochastic discount factor adopted for the change of measure as described in Section 3.4.2. This assumption could result in a not sufficiently dynamical modelling of the risk-neutral expectations of variances, which on the contrary seems to be rapidly mutable in time as evidenced from VIX quotations. Relaxing the assumption and incorporating a dynamics for the variance premia in the SDF could be an attempt to get more insight into the subject for a future research perspective.

Conclusions

This thesis has the purpose to present a novel class of discrete-time models with stochastic volatility and jumps based on realized volatility. The motivation comes from the interest in modelling the connection between the behaviour of the time series of historical returns of financial assets and the pricing of derivatives written on them. Usually, this element is dropped since it is preferred to propose models which are directly specified under the risk-neutral measure. The approach of this thesis is different. The presented models are firstly built in the physical measure in order to encompass the documented statistical properties and stylized facts of assets' return and volatility, and then they are risk-neutralized with the introduction of a stochastic discount factor including multiple risk premia.

First, we introduce the JLHARG-RV class of models which extends the HARG model by Corsi et al. (2013) and the LHARG model by Majewski et al. (2015) by adding a discontinuous dynamic component with constant jump intensity which provides a rapidly moving volatility factor to the autoregressive time evolution of volatility on heterogeneous time scale and leverage effect. Secondly, we detail a model, named LHARG-ARJ, which allows to have a time-varying intensity of extreme events with an autoregressive structure depending on the past number of jump events in order to mimic empirical jump clustering.

They have the advantage to be easily estimated, since they are suited to describe observable variables, and analytically tractable. We demonstrate how to characterize the probability distribution of the log-return process under physical measure by computing the moment generating function with a set of backward recursive formulae. We also provide the framework for the formal change to risk-neutral measure obtained with a flexible exponential affine pricing kernel which identifies different risks and separately compensates for them introducing three

components of risk premia: equity, continuous and jump variance. We prove that the risk-neutral processes are formally equal to the physical ones with the existence of a one-to-one mapping of model parameters depending explicitly on risk premia.

Conducting an empirical assessment of pricing performances, we show the improvements of the novel class of models in reproducing different features of the implied volatility surface compared with the state-of-the-art of pricing models encompassing both continuous and jump dynamics of underlying assets.

Finally, we show that our proposed class of models allows to analytically compute the Variance Risk Premium, defined as the difference between the expected risk-neutral variance and the expected physical variance. In particular, it is possible to compute this quantity for any desired future time horizon and separate the two contributions coming from the continuous and jump dynamics. Our last study investigates the recently evidenced property of the variance risk premium to predict stock market returns. Our results confirm the statistical significance of this predictability of some of the adopted estimates of the variance risk premium. In particular we find a significant predictor alternative to the commonly used market estimate of the variance premium which is defined in a semi-parametric way, depending on both market-implied and model-implied expectations. On the other hand, the complete model-implied measure within the presented framework seems to show predictability only in few cases leaving open the problem to further future researches.

Appendix

A JLHARG-RV proofs and computations

A.1 MGF computations under $\mathbb{P}$ measure

For the ease of computation, the expression (3.5) is rewritten as

$$\Theta(\mathbf{CRV}_t, \mathbf{L}_t) = d + \sum_{i=1}^{22} \beta_i \mathrm{CRV}_{t+1-i} + \sum_{i=1}^{22} \alpha_i \left(\epsilon_{t+1-i} - \gamma\sqrt{\mathrm{CRV}_{t+1-i} + \mathrm{JRV}_{t+1-i}}\right)^2 ,$$

with

$$\beta_i = \begin{cases} \beta_d & \text{for } i = 1 \\ \beta_w/4 & \text{for } 2 \le i \le 5 \\ \beta_w/17 & \text{for } 6 \le i \le 22 \end{cases} \qquad \alpha_i = \begin{cases} \alpha_d & \text{for } i = 1 \\ \alpha_w/4 & \text{for } 2 \le i \le 5 \\ \alpha_w/17 & \text{for } 6 \le i \le 22 \end{cases} . \tag{7}$$

We start showing that JLHARG processes satisfy the affine relation

$$\mathbb{E}\left[e^{zy_{s+1} + \mathbf{b}\cdot\mathbf{RV}_{s+1} + c\ell_{s+1}} | \mathcal{F}_s\right] = e^{\mathcal{A}(z,\mathbf{b},c) + \sum_{i=1}^{p} \boldsymbol{\mathcal{B}}_i(z,\mathbf{b},c)\cdot\mathbf{RV}_{s+1-i} + \sum_{j=1}^{q} \mathcal{C}_j(z,\mathbf{b},c)\ell_{s+1-j}} , \tag{8}$$

for some functions $\mathcal{A} : \mathbb{R}\times\mathbb{R}^2\times\mathbb{R} \to \mathbb{R}$, $\boldsymbol{\mathcal{B}}_i : \mathbb{R}\times\mathbb{R}^2\times\mathbb{R} \to \mathbb{R}^2$, $\boldsymbol{\mathcal{C}}_j : \mathbb{R}\times\mathbb{R}^2\times\mathbb{R} \to \mathbb{R}$, where $\mathbf{RV}_t = (\mathrm{CRV}_t, \mathrm{JRV}_t)$, $\mathbf{b} \in \mathbb{R}^2$, $c \in \mathbb{R}$, and $\cdot$ is the scalar product in $\mathbb{R}^2$. To derive the explicit form of the functions $\mathcal{A}$, $\mathcal{B}_i$, $\mathcal{C}_j$ which allows to characterise

the MGF we show that

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[e^{zy_t+\mathbf{b}\cdot\mathbf{RV}_t+c\ell_t}|\mathcal{F}_{t-1}\right] \\
&= \mathbb{E}^{\mathbb{P}}\left[e^{z(r+\left(\lambda-\frac{1}{2}\right)(\mathrm{CRV}_t+\mathrm{JRV}_t)+\sqrt{\mathrm{CRV}_t+\mathrm{JRV}_t}\epsilon_t)+\mathbf{b}\cdot\mathbf{RV}_t+c\ell_t}|\mathcal{F}_{t-1}\right] \\
&= \mathbb{E}^{\mathbb{P}}\left[e^{z(r+\left(\lambda-\frac{1}{2}\right)(\mathrm{CRV}_t+\mathrm{JRV}_t))+\mathbf{b}\cdot\mathbf{RV}_t}\mathbb{E}^{\mathbb{P}}\left[e^{z\sqrt{\mathrm{CRV}_t+\mathrm{JRV}_t}\epsilon_t+c(\epsilon_t-\gamma\sqrt{\mathrm{CRV}_t+\mathrm{JRV}_t})^2}|\mathbf{RV}_t\right]|\mathcal{F}_{t-1}\right] \\
&= \mathbb{E}^{\mathbb{P}}\left[e^{z(r+\left(\lambda-\frac{1}{2}\right)(\mathrm{CRV}_t+\mathrm{JRV}_t))+b_1\mathrm{CRV}_t+b_2\mathrm{JRV}_t-\frac{1}{2}\ln(1-2c)+\left(\frac{\frac{z^2}{2}+\gamma^2c-2c\gamma z}{1-2c}\right)(\mathrm{CRV}_t+\mathrm{JRV}_t)}|\mathcal{F}_{t-1}\right] \\
&= \mathbb{E}^{\mathbb{P}}\left[e^{zr-\frac{1}{2}\ln(1-2c)+\left(z\left(\lambda-\frac{1}{2}\right)+b_1+\frac{\frac{z^2}{2}+\gamma^2c-2c\gamma z}{1-2c}\right)\mathrm{CRV}_t+\left(z\left(\lambda-\frac{1}{2}\right)+b_2+\frac{\frac{z^2}{2}+\gamma^2c-2c\gamma z}{1-2c}\right)\mathrm{JRV}_t}|\mathcal{F}_{t-1}\right] \\
&= e^{zr-\frac{1}{2}\ln(1-2c)}\mathbb{E}^{\mathbb{P}}\left[e^{\left(z\left(\lambda-\frac{1}{2}\right)+b_1+\frac{\frac{z^2}{2}+\gamma^2c-2c\gamma z}{1-2c}\right)\mathrm{CRV}_t}|\mathcal{F}_{t-1}\right]\mathbb{E}^{\mathbb{P}}\left[e^{\left(z\left(\lambda-\frac{1}{2}\right)+b_2+\frac{\frac{z^2}{2}+\gamma^2c-2c\gamma z}{1-2c}\right)\mathrm{JRV}_t}|\mathcal{F}_{t-1}\right].
\end{aligned}
\tag{9}
$$

In the third line we have used the result that if $Z \sim \mathcal{N}(0,1)$ then

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

For a noncentred gamma random variable, from Gourieroux and Jasiak (2006) we know that

$$\mathbb{E}^{\mathbb{P}}\left[e^{x_1\mathrm{CRV}_t}|\mathcal{F}_{t-1}\right] = \exp\left(-\delta\mathcal{W}\left(x_1,\theta\right)+\mathcal{V}\left(x_1,\theta\right)\left(d+\sum_{i=1}^{p}\beta_i\mathrm{RV}^c_{s-i}+\sum_{j=1}^{q}\alpha_j\ell_{s-j}\right)\right),$$

where

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

and

$$x(z,b,c) = z\left(\lambda-\frac{1}{2}\right)+b+\frac{\frac{1}{2}z^2+\gamma^2c-2c\gamma z}{1-2c}, \quad x_1 = x(z,b_1,c). \tag{10}$$

For the computation of the last expectation in the final line of (9), we use the property that if Z_t is a compound Poisson process with rate ω and *i.i.d.* jump

sizes D_i, then

$$\mathbb{E}\left[e^{xZ_t}|\mathcal{F}_{t-1}\right] = \exp\left(\omega\left(M_D(x)-1\right)\right), \tag{11}$$

where $M_D(x)$ is the MGF of the jump size random variable D. Since sizes are distributed according to a gamma distribution, we have

$$M_D(x) = \frac{1}{\left(1-x\tilde{\theta}\right)^{\tilde{\delta}}}. \tag{12}$$

From expressions (11) and (12) we obtain

$$\mathbb{E}^{\mathbb{P}}\left[e^{x_2 \mathrm{JRV}_t}|\mathcal{F}_{t-1}\right] = \exp\left(\tilde{\Theta}\mathcal{J}\left(x_2,\tilde{\theta},\tilde{\delta}\right)\right),$$

where

$$\mathcal{J}(x,\tilde{\theta},\tilde{\delta}) = \frac{1-(1-\tilde{\theta}x)^{\tilde{\delta}}}{(1-\tilde{\theta}x)^{\tilde{\delta}}} \quad \text{and} \quad x_2 = x(z,b_2,c).$$

Gathering all the previous results, we finally conclude

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

where we have introduced two functions $x_1 = x(z,b_1,c)$ and $x_2 = x(z,b_2,c)$, while the expression for x is given by (10). The direct comparison of the last expression with (8) allows to derive the following explicit expressions

$$\mathcal{A}(z,\mathbf{b},c) = zr-\frac{1}{2}\ln(1-2c)-\delta\mathcal{W}(x_1,\theta)+d\mathcal{V}(x_1,\theta)+\tilde{\Theta}\mathcal{J}(x_2,\theta,\tilde{\delta}), \tag{13}$$

$$\mathcal{B}_i(z,b_1,c) = \mathcal{V}(x_1,\theta)\beta_i, \tag{14}$$

$$\mathcal{C}_j(z,b_1,c) = \mathcal{V}(x_1,\theta)\alpha_j. \tag{15}$$

As shown in Majewski et al. (2015), once we have above expressions we obtain

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

where

$$\begin{aligned}
\mathrm{a}_s &= \mathrm{a}_{s+1} + zr - \frac{1}{2}\log(1-2\mathrm{c}_{s+1,1}) + d\mathcal{V}(\mathrm{x}^c_{s+1},\theta) - \delta\mathcal{W}(\mathrm{x}^c_{s+1},\theta) + \tilde{\Theta}\mathcal{J}(\mathrm{x}^j_{s+1},\tilde{\theta}) \\
\mathrm{b}_{s,i} &= \begin{cases} \mathrm{b}_{s+1,i} + \mathcal{V}(\mathrm{x}^c_{s+1},\theta)\beta_i & for\ 1 \le i \le p-1 \\ \mathcal{V}(\mathrm{x}^c_{s+1},\theta)\beta_i & for\ i = p \end{cases} \\
\mathrm{c}_{s,i} &= \begin{cases} \mathrm{c}_{s+1,i} + \mathcal{V}(\mathrm{x}^c_{s+1},\theta)\alpha_i & for\ 1 \le i \le q-1 \\ \mathcal{V}(\mathrm{x}^c_{s+1},\theta)\alpha_i & for\ i = q \end{cases}
\end{aligned} \tag{16}$$

with

$$\mathrm{x}^c_{s+1} = z\left(\lambda - \frac{1}{2}\right) + \mathrm{b}_{s+1,1} + \frac{\frac{1}{2}z^2 + \gamma^2\mathrm{c}_{s+1,1} - 2\mathrm{c}_{s+1,1}\gamma z}{1-2\mathrm{c}_{s+1,1}}, \tag{17}$$

$$\mathrm{x}^j_{s+1} = z\left(\lambda - \frac{1}{2}\right) + \frac{\frac{1}{2}z^2 + \gamma^2\mathrm{c}_{s+1,1} - 2\mathrm{c}_{s+1,1}\gamma z}{1-2\mathrm{c}_{s+1,1}}. \tag{18}$$

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

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

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

A.2 No-arbitrage condition

The no-arbitrage conditions are

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}}\left[M_{s,s+1}|\mathcal{F}_s\right] &= 1 \text{ for } s \in \mathbb{N}, \\
\mathbb{E}^{\mathbb{P}}\left[M_{s,s+1}e^{y_{s+1}}|\mathcal{F}_s\right] &= e^r \text{ for } s \in \mathbb{N}.
\end{aligned} \tag{19}$$

The first relation is satisfied by definition of $M_{s,s+1}$. From a general result in Majewski et al. (2015), condition (19) is satified if, and only if

$$\begin{aligned}
&\mathcal{A}(1-\nu_y, -\boldsymbol{\nu}, 0) = r + \mathcal{A}(-\nu_y, -\boldsymbol{\nu}, 0)\,,\\
&\mathcal{B}_i(1-\nu_y, -\boldsymbol{\nu}, 0) = \mathcal{B}_i(-\nu_y, -\boldsymbol{\nu}, 0)\,,\\
&\mathcal{C}_j(1-\nu_y, -\boldsymbol{\nu}, 0) = \mathcal{C}_j(-\nu_y, -\boldsymbol{\nu}, 0)\,,
\end{aligned}$$

with $\boldsymbol{\nu} = (\nu_c, \nu_j)$. From the last two relations, using the explicit expressions of $\mathcal{B}_i$ and $\mathcal{C}_j$ given in (14) and (15) we obtain

$$\mathcal{V}(x(1-\nu_y, -\nu_c, 0), \theta) = \mathcal{V}(x(-\nu_y, -\nu_c, 0), \theta)\,,$$

which is equivalent to

$$x(1-\nu_y, -\nu_c, 0) = x(-\nu_y, -\nu_c, 0). \tag{20}$$

Simple computations show that the latter equation fixes the value of the equity premium

$$\nu_y = \lambda.$$

It is worth noticing that the result holding for the equity premium does not constrain the value of the variance risk premia. From the condition on $\mathcal{A}$ it follows that

$$\begin{aligned}
&d\mathcal{V}(x(1-\nu_y, -\nu_c, 0), \theta) - \delta\mathcal{W}(x(1-\nu_y, -\nu_c, 0), \theta) + \tilde{\Theta}\mathcal{J}(x(1-\nu_y, -\nu_j, 0), \tilde{\theta})\\
&= d\mathcal{V}(x(-\nu_y, -\nu_c, 0), \theta) - \delta\mathcal{W}(x(-\nu_y, -\nu_c, 0), \theta) + \tilde{\Theta}\mathcal{J}(x(-\nu_y, -\nu_j, 0), \tilde{\theta})\,,
\end{aligned}$$

which – in light of the relation (20) – reduces to

$$\tilde{\Theta}\mathcal{J}(x(1-\nu_y, -\nu_j, 0), \tilde{\theta}) = \tilde{\Theta}\mathcal{J}(x(-\nu_y, -\nu_j, 0), \tilde{\theta}). \tag{21}$$

Equation (21) is identically satisfied if

$$x(1-\nu_y, -\nu_j, 0) = x(-\nu_y, -\nu_j, 0)\,,$$

which holds for any value of ν_j. In conclusion, the no-arbitrage conditions fix the level of the equity risk premium, while both the continuous and discontinuous variance risk premia remain free parameters to be calibrated on option data.

A.3 Risk-neutral dynamics

JLHARG models imply a risk-neutral MGF for log-returns whose exponential affine terms can be re-parametrized in order to obtain an expression formally equivalent to the physical MGF. Firstly we observe that the risk-neutral MGF can be expressed with a recursive set of expressions, involving a combination of the functions $\mathcal{A}$, $\mathcal{B}_i$, $\mathcal{C}_j$. Then, recalling the results given in Majewski et al. (2015), the MGF for JLHARG model under measure $\mathbb{Q}$ has the following form

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

where

$$\begin{aligned}
\mathrm{a}^*_s &= \mathrm{a}^*_{s+1} + zr - \frac{1}{2}\log(1-2\mathrm{c}^*_{s+1,1}) + d\mathcal{V}(\mathrm{x}^{c\,*}_{s+1},\theta) - d\mathcal{V}(\mathrm{y}^{c\,*}_{s+1},\theta)\\
&- \delta\mathcal{W}(\mathrm{x}^{c\,*}_{s+1},\theta) + \delta\mathcal{W}(\mathrm{y}^{c\,*}_{s+1},\theta) + \tilde{\Theta}\mathcal{J}(\mathrm{x}^{j\,*}_{s+1},\tilde{\theta}) - \tilde{\Theta}\mathcal{J}(\mathrm{y}^{j\,*}_{s+1},\tilde{\theta})\\
\mathrm{b}^*_{s,i} &= \begin{cases} \mathrm{b}^*_{s+1,i} + \left(\mathcal{V}(\mathrm{x}^{c\,*}_{s+1},\theta) - \mathcal{V}(\mathrm{y}^{c\,*}_{s+1},\theta)\right)\beta_i & for\ 1\le i\le p-1\\ \left(\mathcal{V}(\mathrm{x}^{c\,*}_{s+1},\theta) - \mathcal{V}(\mathrm{y}^{c\,*}_{s+1},\theta)\right)\beta_i & for\ i=p \end{cases}\\
\mathrm{c}^*_{s,j} &= \begin{cases} \mathrm{c}^*_{s+1,j} + \left(\mathcal{V}(\mathrm{x}^{c\,*}_{s+1},\theta) - \mathcal{V}(\mathrm{y}^{c\,*}_{s+1},\theta)\right)\alpha_j & for\ 1\le j\le q-1\\ \left(\mathcal{V}(\mathrm{x}^{c\,*}_{s+1},\theta) - \mathcal{V}(\mathrm{y}^{c\,*}_{s+1},\theta)\right)\alpha_j & for\ j=q \end{cases}
\end{aligned} \tag{22}$$

where

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

with $l = c, j$ 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 (22) with (16). 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

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

Then, the following relations have to hold

$$\delta\left(\mathcal{W}\left(\mathrm{x}_{s+1}^{c\,*}, \theta\right) - \mathcal{W}\left(\mathrm{y}^{c\,*}, \theta\right)\right) = \delta^* \mathcal{W}\left(\mathrm{x}_{s+1}^{c\,**}, \theta^*\right) \tag{23}$$
$$\beta_i\left(\mathcal{V}\left(\mathrm{x}_{s+1}^{c\,*}, \theta\right) - \mathcal{V}\left(\mathrm{y}^{c\,*}, \theta\right)\right) = \beta_i^* \mathcal{V}\left(\mathrm{x}_{s+1}^{c\,**}, \theta^*\right) \tag{24}$$
$$\alpha_j\left(\mathcal{V}\left(\mathrm{x}_{s+1}^{c\,*}, \theta\right) - \mathcal{V}\left(\mathrm{y}^{c\,*}, \theta\right)\right) = \alpha_j^* \mathcal{V}\left(\mathrm{x}_{s+1}^{c\,**}, \theta^*\right) \tag{25}$$
$$\tilde{\Theta}\left(\mathcal{J}\left(\mathrm{x}_{s+1}^{j\,*}, \tilde{\theta}\right) - \mathcal{J}\left(\mathrm{y}^{j\,*}, \tilde{\theta}\right)\right) = \tilde{\Theta}^* \mathcal{J}\left(\mathrm{x}_{s+1}^{j\,**}, \tilde{\theta}^*\right) \tag{26}$$

with $\mathrm{y}^{c\,*} = -\lambda^2/2 - \nu_c$ and $\mathrm{y}^{j\,*} = -\lambda^2/2 - \nu_j$.
Equation (23) 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

$$\begin{aligned} \delta^* &= \delta \\ \theta^* &= \frac{\theta}{1 - \theta \mathrm{y}^{c\,*}} \\ \mathrm{x}_{s+1}^{c\,**} &= \mathrm{x}_{s+1}^{c\,*} - \mathrm{y}^{c\,*}. \end{aligned} \tag{27}$$

It can be easily verified that the last condition (27) is satisfied by substituting

$$\begin{aligned} \lambda^* &= 0\,, \\ \gamma^* &= \gamma + \lambda. \end{aligned}$$

The equation (24) 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 (25) allows to obtain the condition on α_i^*,

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

Relation (7) gives us the expressions for β_d^*, β_w^*, β_m^*, α_d^*, α_w^* and α_m^*. Finally, equation (26) 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

$$\begin{aligned}
\tilde{\delta}^* &= \tilde{\delta}\,,\\
\tilde{\Theta}^* &= \frac{\tilde{\Theta}}{\left(1-\tilde{\theta}\mathrm{y}^{j\,*}\right)^{\tilde{\delta}}}\,,\\
\tilde{\theta}^* &= \frac{\tilde{\theta}}{1-\tilde{\theta}\mathrm{y}^{j\,*}}\,,\\
\mathrm{x}_{s+1}^{j\,**} &= \mathrm{x}_{s+1}^{j\,*} - \mathrm{y}^{j\,*}\,.
\end{aligned} \tag{28}$$

As it can be seen the last condition (28) is redundant when compared to the condition (27).

A.4 Conditional expectation of variance

Here we derive the recursive relations that allow to compute the expectation of realized variance over a time horizon of length T conditioned to information up to t. Using the tower law of conditional expectation we obtain

$$\begin{aligned}\mathbb{E}_t^{\mathbb{P}}\left[\mathrm{RV}_{t+T}\right] &= \mathbb{E}_t^{\mathbb{P}}\left[\mathbb{E}_{t+T-1}^{\mathbb{P}}\left[\mathrm{RV}_{t+T}\right]\right] = \mathbb{E}_t^{\mathbb{P}}\left[\mathbb{E}_{t+T-1}^{\mathbb{P}}\left[\mathrm{RV}_{t+T}^{c} + \mathrm{RV}_{t+T}^{j}\right]\right] \\ &= \mathbb{E}_t^{\mathbb{P}}\left[\theta\delta + \sum_{i=1}^{22}\theta\beta_i \mathrm{CRV}_{(t+T-1)+i-1} + \sum_{j=1}^{22}\theta\alpha_j \ell_{(t+T-1)+j-1} + \tilde{\Theta}\tilde{\theta}\tilde{\delta}\right].\end{aligned}$$

We rewrite the last expression in this way

$$\begin{aligned}&\mathbb{E}_t^{\mathbb{P}}\left[a_{t+T-1} + \sum_{i=1}^{22} b_{t+T-1}\mathrm{CRV}_{(t+T-1)+i-1} + \sum_{j=1}^{22} c_{t+T-1}\ell_{(t+T-1)+j-1} + \tilde{\Theta}\tilde{\theta}\tilde{\delta}\right] \\ &= \mathbb{E}_t^{\mathbb{P}}\left[\begin{array}{l} a_{t+T-1} + \sum_{i=2}^{22} b_{t+T-1,\,i}\mathrm{CRV}_{(t+T-1)+i-1} + \sum_{j=2}^{22} c_{t+T-1,\,j}\ell_{(t+T-1)+j-1} + \\ + \mathbb{E}_{t+T-2}^{\mathbb{P}}\left[b_{t+T-1,\,1}\mathrm{RV}_{t+T-1}^{c} + c_{t+T-1,\,1}\ell_{t+T-1}\right] + \tilde{\Theta}\tilde{\theta}\tilde{\delta}\end{array}\right] \\ &= \mathbb{E}_t^{\mathbb{P}}\left[\begin{array}{l} a_{t+T-1} + \sum_{i=2}^{22} b_{t+T-1,\,i}\mathrm{CRV}_{(t+T-1)+i-1} + \sum_{j=2}^{22} c_{t+T-1,\,j}\ell_{(t+T-1)+j-1} \\ + b_{t+T-1,\,1}\left(\theta\delta + \sum_{i=1}^{22}\theta\beta_i\mathrm{CRV}_{(t+T-1)+i-1} + \sum_{j=1}^{22}\theta\alpha_j\ell_{(t+T-2)+j-1}\right) \\ + c_{t+T-1,\,1}\left(1+\gamma^2\left(\theta\delta + \sum_{i=1}^{22}\theta\beta_i\mathrm{CRV}_{(t+T-1)+i-1} + \sum_{j=1}^{22}\theta\alpha_j\ell_{(t+T-2)+j-1}\right) + \gamma^2\tilde{\Theta}\tilde{\theta}\tilde{\delta}\right) \\ + \tilde{\Theta}\tilde{\theta}\tilde{\delta}\end{array}\right].\end{aligned}$$

Iterating the procedure above we finally obtain the expression

$$\mathbb{E}_t^{\mathbb{P}}\left[\mathrm{RV}_{t+T}\right] = a_t + \sum_{i=1}^{22} b_{t,\,i}\mathrm{CRV}_{t+i-1} + \sum_{j=1}^{22} c_{t,\,j}\ell_{t+j-1} + \tilde{\Theta}\tilde{\theta}\tilde{\delta} \tag{29}$$

where the coefficients a_t, $b_{t,\,i}$, $c_{t,\,j}$ are are given at each iteration by formulas (3.34) in Proposition 10, with initial condition $a_{t+T-1} = \theta\delta$, $b_{t+T-1,\,i} = \theta\beta_i$ and $c_{t+T-1,\,j} = \theta\alpha_j$.

B LHARG-ARJ-RV proofs and computations

B.1 MGF computation under $\mathbb{P}$ measure

Firstly we apply the property that a ZM-LHARG process with zero mean leverage can be mapped in a P-LHARG process with parabolic leverage as shown in Majewski et al. (2015), and we rewrite $\Theta(\mathbf{CRV}_t, \mathbf{L}_t)$ in (3.14), as

$$\Theta(\mathbf{CRV}_t, \mathbf{L}_t) = d + \sum_{i=1}^{22} \beta_i \mathrm{CRV}_{t+1-i} + \sum_{j=1}^{22} \alpha_j \ell_{t+1-j}\,, \tag{30}$$

with $d = -(\alpha_d + \alpha_w + \alpha_m)$, $\beta_{d,w,m} = \beta^o_{d,w,m} - \gamma^2 \alpha_{d,w,m}$, where $\beta^o_{d,w,m}$ refer to the original autoregressive parameters for ZM-LHARG in (3.14), and $\ell_t = \left(\epsilon_t - \gamma\sqrt{\mathrm{CRV}_t}\right)^2$ is the mapped parabolic leverage. We further define

$$\beta_i = \begin{cases} \beta_d & \text{for } i = 1 \\ \beta_w/4 & \text{for } 2 \le i \le 5 \\ \beta_m/17 & \text{for } 6 \le i \le 22 \end{cases} \qquad \alpha_j = \begin{cases} \alpha_d & \text{for } j = 1 \\ \alpha_w/4 & \text{for } 2 \le j \le 5 \\ \alpha_m/17 & \text{for } 6 \le j \le 22 \end{cases}. \tag{31}$$

Then we compute one time-step MGF by applying tower law of conditional expectation

$$\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{zy_t + b\mathrm{CRV}_t + c\ell_t} | \mathcal{F}_{t-1}\right] = \mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{z(r + (\lambda_c - \frac{1}{2})\mathrm{CRV}_t + (\lambda_j - \eta)\left(\Lambda^2 + \delta^2\right) n_t + \sqrt{\mathrm{CRV}_t}\epsilon_t + \sum_{i=1}^{n_t} X_{t,i}) + b\mathrm{CRV}_t + c\ell_t} | \mathcal{F}_{t-1}\right]$$

$$= \mathbb{E}^{\mathbb{P}}\left[\begin{array}{l} \mathrm{e}^{z(r + (\lambda_c - \frac{1}{2})\mathrm{CRV}_t + (\lambda_j - \eta)\left(\Lambda^2 + \delta^2\right) n_t + \sum_{i=1}^{n_t} X_{t,i}) + b\mathrm{CRV}_t} \\ \\ \times\, \mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{z\sqrt{\mathrm{CRV}_t}\epsilon_t + c(\epsilon_t - \gamma\sqrt{\mathrm{CRV}_t})^2} | \mathrm{CRV}_t\right] | \mathcal{F}_{t-1} \end{array}\right].$$

Applying the following property of normally distributed random variable $Z \sim \mathcal{N}(0,1)$,

$$\mathbb{E}\left[\exp\left(x(Z+y)^2\right)\right] = \exp\left(-\frac{1}{2}\ln(1-2x) + \frac{xy^2}{1-2x}\right), \tag{32}$$

we obtain

$$\mathbb{E}^{\mathbb{P}}\left[e^{zy_t+b\mathrm{CRV}_t+c\ell_t}|\mathcal{F}_{t-1}\right]=\mathbb{E}^{\mathbb{P}}\left[\begin{array}{l}\mathrm{e}^{z(r+(\lambda_c-\frac{1}{2})\mathrm{CRV}_t+(\lambda_j-\eta)\left(\Lambda^2+\delta^2\right)n_t+\sum_{i=1}^{n_t}X_{t,i})+b\mathrm{CRV}_t} \\ \times\,\mathrm{e}^{-\frac{1}{2}\ln(1-2c)+\frac{\frac{z^2}{2}+\gamma^2c-2c\gamma z}{1-2c}\mathrm{CRV}_t}\end{array}|\mathcal{F}_{t-1}\right] \tag{33}$$

Since n_t and $X_{t,i}\sim\mathcal{N}(\Lambda,\delta)$ are independent we have

$$\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{z(\lambda_j-\eta)\left(\Lambda^2+\delta^2\right)n_t+z\sum_{i=1}^{n_t}X_{t,i}}|n_t\right]=\exp\left(\left(\frac{z^2\delta^2}{2}+\Lambda z+(\lambda_j-\eta)\left(\Lambda^2+\delta^2\right)z\right)n_t\right).$$

Introducing

$$v(z)=\frac{z^2\delta^2}{2}+\Lambda z+(\lambda_j-\eta)\left(\Lambda^2+\delta^2\right)z \tag{34}$$

and

$$x(z,b,c)=z\left(\lambda_c-\frac{1}{2}\right)+b+\frac{\frac{z^2}{2}+\gamma^2c-2c\gamma z}{1-2c}, \tag{35}$$

we can rewrite (33) as

$$\mathbb{E}^{\mathbb{P}}\left[e^{zy_t+b\mathrm{CRV}_t+c\ell_t}|\mathcal{F}_{t-1}\right]=\mathbb{E}^{\mathbb{P}}\left[e^{zr-\frac{1}{2}\ln(1-2c)+x(z,b,c)\mathrm{CRV}_t+v(z)n_t}|\mathcal{F}_{t-1}\right].$$

Using the independence of the random variables, CRV_t and n_t, we can factorize the previous expression

$$\mathbb{E}^{\mathbb{P}}\left[e^{zy_t+b\mathrm{CRV}_t+c\ell_t}|\mathcal{F}_{t-1}\right]=e^{zr-\frac{1}{2}\ln(1-2c)}\mathbb{E}^{\mathbb{P}}\left[e^{x(z,b,c)\mathrm{CRV}_t}|\mathcal{F}_{t-1}\right]\mathbb{E}^{\mathbb{P}}\left[e^{v(z)n_t}|\mathcal{F}_{t-1}\right].$$

For the brevity of the notation, in the following part of the text we will write x for $x(z,b,c)$ and v for $v(z)$. Using eq.s (8)-(9) from Gourieroux and Jasiak (2006) we obtain

$$\mathbb{E}^{\mathbb{P}}\left[e^{x\mathrm{CRV}_t}|\mathcal{F}_{t-1}\right]=\exp\left(-\kappa\mathcal{W}(x,\theta)+\mathcal{V}(x,\theta)\left(d+\sum_{i=1}^{22}\beta_i\mathrm{CRV}_{t-i}+\sum_{j=1}^{22}\alpha_j\ell_{t-j}\right)\right),$$

where

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

Since n_t has Poisson distribution with intensity ω_t we have

$$\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{vn_t}|\mathcal{F}_{t-1}\right]=\exp\left(\omega_t\left(\mathrm{e}^{v}-1\right)\right).$$

Finally, we collect the previous results and obtain an exponentially affine form for the physical MGF,

$$\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{zy_t+b\mathrm{CRV}_t+c\ell_t}|\mathcal{F}_{t-1}\right]=\exp\left(\begin{array}{l} zr-\dfrac{1}{2}\ln(1-2c)-\kappa\mathcal{W}\left(x,\theta\right)+d\mathcal{V}\left(x,\theta\right)\\ +\mathcal{V}\left(x,\theta\right)\left(\displaystyle\sum_{i=1}^{22}\beta_i\mathrm{CRV}_{t-i}+\sum_{j=1}^{22}\alpha_j\ell_{t-j}\right)\\ +\omega_t\left(\mathrm{e}^{v}-1\right)\end{array}\right) \tag{36}$$

where ω_t is $\mathcal{F}_{t-1}$-measurable.

Now we are going to compute the MGF for the log-return $y_{t,T}=\sum_{i=t}^{T}y_i$ between t and T. We involve a repeated use of the tower law of expectation,

$$\begin{aligned}\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{zy_{t,T}}|\mathcal{F}_t\right]&=\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{zy_{t,T-1}}\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{zy_T}|\mathcal{F}_{T-1}\right]|\mathcal{F}_t\right]\\&=\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{zy_{t,T-1}+\mathrm{a}_{T-1}+\sum_{i=1}^{22}\mathrm{b}_{T-1,i}\mathrm{CRV}_{T-i}+\sum_{j=1}^{22}\mathrm{c}_{T-1,j}\ell_{T-j}+\mathrm{d}_T(\xi\omega_{T-1}+\zeta n_{T-1})}|\mathcal{F}_t\right],\end{aligned}$$

where we introduce a set of coefficients defined as follows

$$\begin{aligned}\mathrm{a}_{T-1}&=zr-\frac{1}{2}\ln(1-2c)-\kappa\mathcal{W}\left(x(z,0,0),\theta\right)+d\mathcal{V}\left(x(z,0,0),\theta\right)+\mathrm{d}_T\bar{\omega},\\ \mathrm{b}_{T-1,i}&=\beta_i\mathcal{V}\left(x(z,0,0),\theta\right),\\ \mathrm{c}_{T-1,j}&=\alpha_j\mathcal{V}\left(x(z,0,0),\theta\right),\\ \mathrm{d}_T&=\mathrm{e}^{v}-1.\end{aligned} \tag{37}$$

Moving a time step backward, we isolate all the random variables at time $T-1$

and compute the expectation conditioning to information up to $T-2$,

$$\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{zy_{t,T}}|\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{P}}\left[\begin{array}{l} e^{zy_{t,T-2}+\mathrm{a}_{T-1}+\sum_{i=2}^{22}\mathrm{b}_{T-1,i}\mathrm{CRV}_{T-i}+\sum_{j=2}^{22}\mathrm{c}_{T-1,j}\ell_{T-j}+\mathrm{d}_T\xi\omega_{T-1}} \\ \times\, \mathbb{E}^{\mathbb{P}}\left[e^{zy_{T-1}+\mathrm{b}_{T-1,1}\mathrm{CRV}_{T-1}+\mathrm{c}_{T-1,1}\ell_{T-1}+\mathrm{d}_T\zeta n_{T-1}}\middle|\mathcal{F}_{T-2}\right] \end{array}\middle|\mathcal{F}_t\right]$$

$$= \mathbb{E}^{\mathbb{P}}\left[e^{zy_{t,T-2}+\mathrm{a}_{T-2}+\sum_{i=1}^{22}\mathrm{b}_{T-2,i}\mathrm{CRV}_{T-1-i}+\sum_{j=1}^{22}\mathrm{c}_{T-2,j}\ell_{T-1-j}+\mathrm{d}_{T-1}(\xi\omega_{T-2}+\zeta n_{T-2})}|\mathcal{F}_t\right]$$

and continuing this reasoning further

$$\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{zy_{t,T}}|\mathcal{F}_t\right] = \exp\left(\mathrm{a}_t + \sum_{i=1}^{22}\mathrm{b}_{t,i}\mathrm{CRV}_{t+1-i} + \sum_{j=1}^{22}\mathrm{c}_{t,j}\ell_{t+1-j} + \mathrm{d}_{t+1}\left(\xi\omega_t + \zeta n_t\right)\right).$$

where a set of recursive relations is as follows:

$$\begin{aligned}
\mathrm{a}_s &= \mathrm{a}_{s+1} + zr - \frac{1}{2}\ln(1-2c_{s+1,1}) - \kappa\mathcal{W}\left(x_{s+1},\theta\right) + d\mathcal{V}\left(x_{s+1},\theta\right) + \mathrm{d}_{s+1}\bar{\omega}, \\
\mathrm{b}_{s,i} &= \begin{cases} \mathrm{b}_{s+1,i} + \beta_i\mathcal{V}\left(x_{s+1},\theta\right) & \text{for } 1 \le i < 22 \\ \beta_i\mathcal{V}\left(x_{s+1},\theta\right) & \text{for } i = 22 \end{cases}, \\
\mathrm{c}_{s,j} &= \begin{cases} \mathrm{c}_{s+1,j} + \alpha_j\mathcal{V}\left(x_{s+1},\theta\right) & \text{for } 1 \le j < 22 \\ \alpha_j\mathcal{V}\left(x_{s+1},\theta\right) & \text{for } j = 22 \end{cases}, \\
\mathrm{d}_s &= e^{v+\zeta\mathrm{d}_{s+1}} - 1 + \xi\mathrm{d}_{s+1},
\end{aligned}$$

with

$$x_{s+1} = z\left(\lambda_c - \frac{1}{2}\right) + b_{s+1,1} + \frac{\frac{z^2}{2} + \gamma^2 c_{s+1,1} - 2c_{s+1,1}\gamma z}{1-2c_{s+1,1}}, \tag{38}$$

and terminal conditions are following $\mathrm{d}_T = \mathrm{e}^v - 1$ and $\mathrm{a}_T = \mathrm{b}_{T,i} = \mathrm{c}_{T,j} = 0$ for $i = 1,\ldots,p$ and $j = 1,\ldots,q$.

B.2 No-arbitrage condition

The no-arbitrage constraint

$$\mathbb{E}^{\mathbb{P}}\left[M_{t-1,t}\mathrm{e}^{y_t}|\mathcal{F}_{t-1}\right] = \mathrm{e}^r \tag{39}$$

with stochastic discount factor as defined in (3.25) reads

$$\frac{\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{y_t-\nu_c\mathrm{CRV}_t-\nu_j\mathrm{JRV}_t-\mu_c\sqrt{\mathrm{CRV}_t}\epsilon_t-\mu_j\sum_{i=1}^{n_t}X_{t,i}}\middle|\mathcal{F}_{t-1}\right]}{\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{r-\nu_c\mathrm{CRV}_t-\nu_j\mathrm{JRV}_t-\mu_c\sqrt{\mathrm{CRV}_t}\epsilon_t-\mu_j\sum_{i=1}^{n_t}X_{t,i}}\middle|\mathcal{F}_{t-1}\right]}=1. \tag{40}$$

Firstly we compute the denominator of the expression of the SDF above:

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[e^{r-\nu_c\mathrm{CRV}_t-\nu_j\mathrm{JRV}_t-\mu_c\sqrt{\mathrm{CRV}_t}\epsilon_t-\mu_j\sum_{i=1}^{n_t}X_{t,i}}\middle|\mathcal{F}_{t-1}\right]=\\
&=\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{r-\nu_c\mathrm{CRV}_t}\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{-\mu_c\sqrt{\mathrm{CRV}_t}\epsilon_t}\middle|\mathrm{CRV}_t\right]\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{-\nu_j\mathrm{JRV}_t-\mu_j\sum_{i=1}^{n_t}X_{t,i}}\middle|n_t\right]\middle|\mathcal{F}_{t-1}\right]\\
&=\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{r+\left(\frac{\nu_2^2}{2}-\nu_c\right)\mathrm{CRV}_t}\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{-\nu_j\sum_{i=1}^{n_t}X_{t,i}^2-\mu_j\sum_{i=1}^{n_t}X_{t,i}}\middle|n_t\right]\middle|\mathcal{F}_{t-1}\right]\\
&=\exp\left(r-\kappa\mathcal{W}\left(\bar{x},\theta\right)+\mathcal{V}\left(\bar{x},\theta\right)\left(d+\sum_{i=1}^{22}\beta_i\mathrm{CRV}_{t-i}+\sum_{j=1}^{22}\alpha_j\ell_{t-j}\right)+\omega_t\left(\mathrm{e}^{\bar{v}}-1\right)\right)
\end{aligned}$$

where

$$\begin{aligned}
\bar{x}&=-\nu_c+\frac{\mu_c^2}{2},\\
\bar{v}&=-\frac{1}{2}\ln(1+2\nu_j\delta^2)+\frac{\left(\mu_j+2\Lambda\nu_j\right)^2\delta^2}{2(1+2\nu_j\delta^2)}-\Lambda(\mu_j+\nu_j\Lambda).
\end{aligned} \tag{41}$$

Proceeding analogous computations for the numerator of the no-arbitrage condition gives

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{-\nu_c\mathrm{CRV}_t-\nu_j\mathrm{JRV}_t-\mu_c\sqrt{\mathrm{CRV}_t}\epsilon_t-\mu_j\sum_{i=1}^{n_t}X_{t,i}+y_t}\middle|\mathcal{F}_{t-1}\right]=\\
&=\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{-\nu_c\mathrm{CRV}_t-\nu_j\mathrm{JRV}_t-\mu_c\sqrt{\mathrm{CRV}_t}\epsilon_t-\mu_j\sum_{i=1}^{n_t}X_{t,i}+r+\left(\lambda_c-\frac{1}{2}\right)\mathrm{CRV}_t+(\lambda_j-\eta)\left(\Lambda^2+\delta^2\right)n_t+\sqrt{\mathrm{CRV}_t}\epsilon_t+\sum_{i=1}^{n_t}X_{t,i}}\middle|\mathcal{F}_{t-1}\right]\\
&=\exp\left(r-\kappa\mathcal{W}\left(\tilde{x},\theta\right)+\mathcal{V}\left(\tilde{x},\theta\right)\left(d+\sum_{i=1}^{22}\beta_i\mathrm{CRV}_{t-i}+\sum_{j=1}^{22}\alpha_j\ell_{t-j}\right)+\omega_t\left(\mathrm{e}^{\tilde{v}}-1\right)\right)
\end{aligned}$$

where

$$\tilde{x} = \left(\lambda_c - \frac{1}{2}\right) - \nu_c + \frac{(1-\mu_c)^2}{2}$$
$$\tilde{v} = -\frac{1}{2}\ln(1+2\nu_j\delta^2) + \frac{(1-\mu_j-2\Lambda\nu_j)^2\,\delta^2}{2(1+2\nu_j\delta^2)} + \Lambda(1-\mu_j-\nu_j\Lambda) + (\lambda_j-\eta)\left(\Lambda^2+\delta^2\right).$$

Putting together numerator and denominator, the no-arbitrage condition (40) becomes

$$\left(\mathcal{V}(\tilde{x},\theta) - \mathcal{V}(\bar{x},\theta)\right)\left(d + \sum_{i=1}^{22}\beta_i \mathrm{CRV}_{t-i} + \sum_{j=1}^{22}\alpha_j\ell_{t-j}\right) + \omega_t\left(e^{\tilde{v}} - e^{\bar{v}}\right) - \kappa\left(\mathcal{W}(\tilde{x},\theta) - \mathcal{W}(\bar{x},\theta)\right) = 0$$

In order to be satisfied, the above equation implies the further following relations,

$$\begin{aligned}\mathcal{W}(\tilde{x},\theta) - \mathcal{W}(\bar{x},\theta) &= 0\\ \mathcal{V}(\tilde{x},\theta) - \mathcal{V}(\bar{x},\theta) &= 0\\ \mathrm{e}^{\tilde{v}} - \mathrm{e}^{\bar{v}} &= 0,\end{aligned}$$

which are satisfied if $\tilde{x} = \bar{x}$ and $\tilde{v} = \bar{v}$. From $\tilde{x} = \bar{x}$, we have

$$\mu_c = \lambda_c$$

and $\tilde{v} = \bar{v}$ implies

$$\mu_j = \frac{1}{2} + \frac{\Lambda + (\lambda_j - \eta)\left(\Lambda^2+\delta^2\right)\left(1+2\nu_j\delta^2\right)}{\delta^2}.$$

B.3 Risk-neutral dynamics

First we derive moment generating function of vector $(y_t, \mathrm{CRV}_t, \ell_t)$ conditioned on $\mathcal{F}_{t-1}$under risk-neutral measure $\mathbb{Q}$.

$$\mathbb{E}^{\mathbb{Q}}\left[\mathrm{e}^{zy_t + b^*\mathrm{CRV}_t + c^*\ell_t}\middle|\mathcal{F}_{t-1}\right] = \mathbb{E}^{\mathbb{P}}\left[M_{t-1,t}\mathrm{e}^{zy_t + b^*\mathrm{CRV}_t + c^*\ell_t}\middle|\mathcal{F}_{t-1}\right] =$$

$$= \mathbb{E}^{\mathbb{P}}\left[\frac{\mathrm{e}^{-\nu_c\mathrm{CRV}_t - \nu_j\mathrm{JRV}_t - \mu_c\sqrt{\mathrm{CRV}_t}\epsilon_t - \mu_j\sum_{i=1}^{n_t}X_{t,i} + zy_t + b^*\mathrm{CRV}_t + c^*\ell_t}}{\mathbb{E}^{\mathbb{P}}\left[e^{-\nu_c\mathrm{CRV}_t - \nu_j\mathrm{JRV}_t - \mu_c\sqrt{\mathrm{CRV}_t}\epsilon_t - \mu_j\sum_{i=1}^{n_t}X_{t,i}}\middle|\mathcal{F}_{t-1}\right]}\middle|\mathcal{F}_{t-1}\right]$$

Proceeding analogous computation as in Appendix B.1 we obtain

$$\mathbb{E}^{\mathbb{Q}}\left[\mathrm{e}^{zy_t+b^*\mathrm{CRV}_t+c^*\ell_t}|\mathcal{F}_{t-1}\right] = \\ = \exp\left(\begin{array}{l} zr-\frac{1}{2}\ln(1-2c^*)-\kappa\left(\mathcal{W}\left(\hat{x},\theta\right)-\mathcal{W}\left(\bar{x},\theta\right)\right)+d\left(\mathcal{V}\left(\hat{x},\theta\right)-\mathcal{V}\left(\bar{x},\theta\right)\right) \\ +\left(\mathcal{V}\left(\hat{x},\theta\right)-\mathcal{V}\left(\bar{x},\theta\right)\right)\left(\sum_{i=1}^{22}\beta_i\mathrm{CRV}_{t-i}+\sum_{j=1}^{22}\alpha_j\ell_{t-j}\right) \\ +\omega_t\mathrm{e}^{\bar{v}}\left(e^{\hat{v}-\bar{v}}-1\right) \end{array}\right) \tag{42}$$

where $\bar{x}$, $\bar{v}$ are given in (41) and we introduce notation

$$\begin{aligned} \hat{x} &= z\left(\lambda_c-\frac{1}{2}\right)+b^*+\frac{\frac{(z-\mu_c)^2}{2}+\gamma^2c^*-2c^*\gamma\left(z-\mu_c\right)}{1-2c^*}, \\ \hat{v} &= -\frac{1}{2}\ln\left(1+2\nu_j\delta^2\right)+\frac{\left(z-\mu_j-2\Lambda\nu_j\right)^2\delta^2}{2(1+2\nu_j\delta^2)}+\Lambda(z-\mu_j-\nu_j\Lambda)+z\left(\lambda_j-\eta\right)\left(\Lambda^2+\delta^2\right). \end{aligned} \tag{43}$$

Comparison of (42) with equivalent expression under physical measure $\mathbb{P}$, given by (36) yields the following relations

$$\begin{aligned} \kappa\left(\mathcal{W}\left(\hat{x},\theta\right)-\mathcal{W}\left(\bar{x},\theta\right)\right) &= \kappa^*\mathcal{W}\left(x^*,\theta^*\right) \\ d\left(\mathcal{V}\left(\hat{x},\theta\right)-\mathcal{V}\left(\bar{x},\theta\right)\right) &= d^*\mathcal{V}\left(x^*,\theta^*\right) \\ \alpha_j\left(\mathcal{V}\left(\hat{x},\theta\right)-\mathcal{V}\left(\bar{x},\theta\right)\right) &= \alpha_j^*\mathcal{V}\left(x^*,\theta^*\right) \\ \beta_i\left(\mathcal{V}\left(\hat{x},\theta\right)-\mathcal{V}\left(\bar{x},\theta\right)\right) &= \beta_i^*\mathcal{V}\left(x^*,\theta^*\right) \\ \omega_t e^{\bar{v}} &= \omega_t^* \\ \hat{v}-\bar{v} &= v^*, \end{aligned} \tag{44}$$

where

$$\begin{aligned} x^* &= z\left(\lambda^*-\frac{1}{2}\right)+b^*+\frac{\frac{z^2}{2}+(\gamma^*)^2c^*-2c^*\gamma^*z}{1-2c^*} \\ v^* &= \frac{z^2(\delta^*)^2}{2}+z\Lambda^*+z\left(\lambda_j^*-\eta^*\right)\left(\left(\Lambda^*\right)^2+\left(\delta^*\right)^2\right). \end{aligned} \tag{45}$$

We observe that the parameters Λ^* and δ^* account for the reparameterization of the mean and the standard deviation of the random variable of jump size that in physical measure is modelled by a normal distribution $\mathcal{N}(\Lambda, \delta^2)$ transforming in a normal $\mathcal{N}(\Lambda^*, (\delta^*)^2)$ under risk-neutral measure.

The first four relations in (44) togehter with no-arbitrage condition imply the following reparameterization of the noncentered gamma process:

$$\begin{array}{lll} \beta_d^* = \frac{1}{1-\theta \mathrm{y}^*}\beta_d\,, & \beta_w^* = \frac{1}{1-\theta \mathrm{y}^*}\beta_w\,, & \beta_m^* = \frac{1}{1-\theta \mathrm{y}^*}\beta_m\,, \\ \alpha_d^* = \frac{1}{1-\theta \mathrm{y}^*}\alpha_d\,, & \alpha_w^* = \frac{1}{1-\theta \mathrm{y}^*}\alpha_w\,, & \alpha_m^* = \frac{1}{1-\theta \mathrm{y}^*}\alpha_m\,, \\ \theta^* = \frac{1}{1-\theta \mathrm{y}^*}\theta\,, & \delta^* = \delta\,, & \gamma^* = \gamma + \lambda_c\,, \\ d^* = \frac{1}{1-\theta \mathrm{y}^*}d\,, & \lambda_c^* = 0, & \end{array} \tag{46}$$

where $\mathrm{y}^* = -\nu_c + \lambda^2/2$.

The last two relations in (44) are related to the reparameterization of the jump component of the process. However, to get the complete characteristics of jump component of the log-return dynamics under risk-neutral measure we derive the MGF of sum of jumps over 2 days.

$$\begin{aligned} &\mathbb{E}^{\mathbb{Q}}\left[\mathrm{e}^{z\left(\lambda_j^*-\eta^*\right)\left((\Lambda^*)^2+(\delta^*)^2\right)(n_{t+1}+n_{t+2})+z\sum_{i=t+1}^{t+2}\sum_{k=1}^{n_i^*}X_{i,k}^*}\middle|\mathcal{F}_t\right] \\ &= \mathbb{E}^{\mathbb{Q}}\left[\begin{array}{l} \mathrm{e}^{z\left(\lambda_j^*-\eta^*\right)\left((\Lambda^*)^2+(\delta^*)^2\right)n_{t+1}+z\sum_{j=1}^{n_{t+1}^*}X_{t+1,j}^*} \\ \\ \times\,\mathbb{E}^{\mathbb{Q}}\left[\mathrm{e}^{z\left(\lambda_j^*-\eta^*\right)\left((\Lambda^*)^2+(\delta^*)^2\right)n_{t+2}+z\sum_{k=1}^{n_{t+2}^*}X_{t+2,k}^*}\middle|\mathcal{F}_{t+1}\right] \end{array}\middle|\mathcal{F}_t\right], \\ &= \mathbb{E}^{\mathbb{Q}}\left[\mathrm{e}^{z\left(\lambda_j^*-\eta^*\right)\left((\Lambda^*)^2+(\delta^*)^2\right)n_{t+1}+z\sum_{k=1}^{n_{t+1}^*}X_{t+1,k}^*+\omega_{t+2}^*\left(\mathrm{e}^{v^*}-1\right)}\middle|\mathcal{F}_t\right] \end{aligned}$$

where v^* is given in (45). Assuming that ω^* has under measure $\mathbb{Q}$ dynamics given

by equation (3.18), we can write

$$\mathbb{E}^{\mathbb{Q}}\left[\mathrm{e}^{z\left(\lambda_j^*-\eta^*\right)\left((\Lambda^*)^2+(\delta^*)^2\right)(n_{t+1}+n_{t+2})+z\sum_{i=t+1}^{t+2}\sum_{k=1}^{n_i^*}X_{i,k}^*}|\mathcal{F}_t\right]$$
$$=\mathbb{E}^{\mathbb{Q}}\left[\exp\left(\begin{array}{l} z\left(\lambda_j^*-\eta^*\right)\left((\Lambda^*)^2+(\delta^*)^2\right)n_{t+1}+\left(\bar{\omega}^*+\xi^*\omega_{t+1}^*+\zeta^* n_{t+1}^*\right)\left(\mathrm{e}^{v^*}-1\right) \\ +z\sum_{k=1}^{n_{t+1}^*}X_{t+1,k}^* \end{array}\right)|\mathcal{F}_t\right].$$

Finally, considering that n_{t+1}^* is Poisson and $X_{t+1,j}^*$ are i.i.d. normally distributed, we obtain

$$\begin{aligned}&\mathbb{E}^{\mathbb{Q}}\left[\mathrm{e}^{z\left(\lambda_j^*-\eta^*\right)\left((\Lambda^*)^2+(\delta^*)^2\right)(n_{t+1}+n_{t+2})+z\sum_{i=t+1}^{t+2}\sum_{k=1}^{n_i^*}X_{i,k}^*}|\mathcal{F}_t\right]= \\ &=\exp\left(\left(\bar{\omega}^*+\xi^*\omega_{t+1}^*\right)\left(\mathrm{e}^{v^*}-1\right)+\omega_{t+1}^*\left(\mathrm{e}^{u^*}-1\right)\right),\end{aligned} \tag{47}$$

where

$$u^*=\frac{z^2(\delta^*)^2}{2}+z\Lambda^*+z\left(\lambda_j^*-\eta^*\right))\left((\Lambda^*)^2+(\delta^*)^2\right)+\zeta^*\left(\mathrm{e}^{v^*}-1\right).$$

We are going to compare (47) with

$$\mathbb{E}^{\mathbb{P}}\left[\prod_{i=t}^{t+1}M_{i,i+1}\mathrm{e}^{z(\lambda_j-\eta)\left(\Lambda^2+\delta^2\right)n_{i+1}+z\sum_{k=1}^{n_{i+1}}X_{i+1,k}}|\mathcal{F}_t\right]$$
$$=\mathbb{E}^{\mathbb{P}}\left[\frac{\mathrm{e}^{-\nu_c\mathrm{CRV}_{t+2}-\mu_c\sqrt{\mathrm{CRV}_{t+2}}\epsilon_{t+2}}}{\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{-\nu_c\mathrm{CRV}_{t+2}-\mu_c\sqrt{\mathrm{CRV}_{t+2}}\epsilon_{t+2}}|\mathcal{F}_{t+1}\right]}|\mathcal{F}_t\right]$$
$$\times\mathbb{E}^{\mathbb{P}}\left[\begin{array}{l} M_{t,t+1}\mathrm{e}^{z(\lambda_j-\eta)\left(\Lambda^2+\delta^2\right)n_{t+1}+z\sum_{k=1}^{n_{t+1}}X_{t+1,k}} \\ \times\mathbb{E}^{\mathbb{P}}\left[\frac{\mathrm{e}^{-\nu_j\mathrm{JRV}_{t+2}-\mu_j\sum_{k=1}^{n_{t+2}}X_{t+2,k}+z(\lambda_j-\eta)\left(\Lambda^2+\delta^2\right)n_{t+2}+z\sum_{k=1}^{n_{t+2}}X_{t+2,k}}}{\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{-\nu_j\mathrm{JRV}_{t+2}-\mu_j\sum_{k=1}^{n_{t+2}}X_{t+2,k}}|\mathcal{F}_{t+1}\right]}|\mathcal{F}_{t+1}\right] \end{array}|\mathcal{F}_t\right]$$
$$=\mathbb{E}^{\mathbb{P}}\left[M_{t,t+1}\exp\left(z\left(\lambda_j-\eta\right)\left(\Lambda^2+\delta^2\right)n_{t+1}+z\sum_{k=1}^{n_{t+1}}X_{t+1,k}+\omega_{t+2}\left(\mathrm{e}^{v}-\mathrm{e}^{\bar{v}}\right)\right)|\mathcal{F}_t\right]$$

where

$$\bar{v} = -\frac{1}{2}\ln(1+2\nu_j\delta^2) + \frac{(\mu_j+2\Lambda\nu_j)^2\,\delta^2}{2(1+2\nu_j\delta^2)} - \Lambda(\mu_j+\nu_j\Lambda),$$

$$v = -\frac{1}{2}\ln(1+2\nu_j\delta^2) + \frac{(z-\mu_j-2\Lambda\nu_j)^2\,\delta^2}{2(1+2\nu_j\delta^2)} + \Lambda(z-\mu_j-\nu_j\Lambda) + z\,(\lambda_j-\eta)\left(\Lambda^2+\delta^2\right).$$

Applying the dynamics of the intensity ω given by equation (3.18) we get

$$\begin{aligned}&\mathbb{E}^{\mathbb{P}}\left[\prod_{i=t}^{t+1} M_{i,i+1}\mathrm{e}^{z(\lambda_j-\eta)\left(\Lambda^2+\delta^2\right)n_{i+1}+z\sum_{k=1}^{n_{i+1}}X_{i+1,k}}|\mathcal{F}_t\right] = \\ &= \exp\left(\mathrm{e}^{\bar{v}}\left(\bar{\omega}+\xi\omega_{t+1}\right)\left(\mathrm{e}^{v-\bar{v}}-1\right)+\omega_{t+1}\mathrm{e}^{\bar{v}}\left(\mathrm{e}^{u-\bar{v}}-1\right)\right)\end{aligned} \tag{48}$$

where

$$u = -\frac{1}{2}\ln(1+2\nu_j\delta^2)+\frac{(z-\mu_j-2\Lambda\nu_j)^2\,\delta^2}{2(1+2\nu_j\delta^2)}+\Lambda(z-\mu_j-\nu_j\Lambda)+z\,(\lambda_j-\eta)\left(\Lambda^2+\delta^2\right)+\zeta e^{\bar{v}}\left(e^{v-\bar{v}}-1\right).$$

Comparing (47) with (48) yields

$$\begin{aligned}\bar{\omega}e^{\bar{v}}\left(e^{v-\bar{v}}-1\right) &= \bar{\omega}^*\left(e^{v}-1\right)\\ \xi\omega_{t+1}e^{\bar{v}}\left(e^{v-\bar{v}}-1\right)+\omega_{t+1}e^{\bar{v}}\left(e^{u-\bar{v}}-1\right) &= \xi^*\omega^*_{t+1}\left(e^{v^*}-1\right)+\omega^*_{t+1}\left(e^{u^*}-1\right).\end{aligned}$$

The above expressions yield the following relation between physical and risk-neutral intensity of jumps

$$\omega^*_{t+1} = \omega_{t+1}e^{\bar{v}} \tag{49}$$

and the following mapping of the parameters

$$\begin{aligned}
\Lambda^* &= \Lambda - (\mu_j + 2\nu_j\Lambda)\,\frac{\delta^2}{(1+2\nu_j\delta^2)},\\
(\delta^*)^2 &= \frac{\delta^2}{(1+2\nu_j\delta^2)},\\
\lambda_j^* &= \frac{(\lambda_j-\eta)\,(\Lambda^2+\delta^2)}{(\Lambda^*)^2+(\delta^*)^2}+\eta^*,\\
\bar{\omega}^* &= e^{\bar{v}}\bar{\omega},\\
\xi^* &= \xi,\\
\zeta^* &= e^{\bar{v}}\zeta
\end{aligned} \tag{50}$$

with $\eta^* = \frac{\Lambda^*+\frac{1}{2}(\delta^*)^2}{(\Lambda^*)^2+(\delta^*)^2}$.

B.4 Conditional expectation of variance

Here we derive the recursive relations that allow to compute the expectation of realized variance over a time horizon of length T conditioned to information up to t. We apply the same mapping from ZM-LHARG to P-LHARG concerning the CRV_t dynamics as described at the beginning of Appendix B.1. Using the tower law of conditional expectation we obtain

$$\begin{aligned}
&\mathbb{E}_t^{\mathbb{P}}\left[\mathrm{RV}_T\right] = \mathbb{E}_t^{\mathbb{P}}\left[\mathbb{E}_{T-1}^{\mathbb{P}}\left[\mathrm{RV}_T\right]\right] = \mathbb{E}_t^{\mathbb{P}}\left[\mathbb{E}_{T-1}^{\mathbb{P}}\left[\mathrm{CRV}_T + \mathrm{JRV}_T\right]\right]\\
&= \mathbb{E}_t^{\mathbb{P}}\left[\theta(d+\kappa) + \sum_{i=1}^{p}\theta\beta_i\mathrm{CRV}_{T-i} + \sum_{j=1}^{q}\theta\alpha_j\ell_{T-j} + (\delta^2+\Lambda^2)\,\omega_T\right]\\
&= \mathbb{E}_t^{\mathbb{P}}\left[\begin{array}{l}\theta(d+\kappa) + \left(\delta^2+\Lambda^2\right)\bar{\omega} + \sum_{i=1}^{p}\theta\beta_i\mathrm{CRV}_{T-i} + \sum_{j=1}^{q}\theta\alpha_j\ell_{T-j}\\ \\ \quad + \left(\delta^2+\Lambda^2\right)(\xi\omega_{T-1} + \zeta n_{T-1})\end{array}\right].
\end{aligned}$$

We rewrite the last expression in this way

$$
\begin{aligned}
&\mathbb{E}_t^{\mathbb{P}}\left[\mathrm{a}_{T-1}+\sum_{i=1}^{p}\mathrm{b}_{T-1,i}\mathrm{CRV}_{T-i}+\sum_{j=1}^{q}\mathrm{c}_{T-1,j}\ell_{T-j}+\mathrm{d}_T\left(\xi\omega_{T-1}+\zeta n_{T-1}\right)\right]\\
&=\mathbb{E}_t^{\mathbb{P}}\left[\begin{array}{l}\mathrm{a}_{T-1}+\sum_{i=2}^{p}\mathrm{b}_{T-1,i}\mathrm{CRV}_{T-i}+\sum_{j=2}^{q}\mathrm{c}_{T-1,j}\ell_{T-j}+\\ +\mathbb{E}_{T-2}^{\mathbb{P}}\left[\mathrm{b}_{T-1,1}\mathrm{CRV}_{T-1}+\mathrm{c}_{T-1,1}\ell_{T-1}\right]+\mathrm{d}_T\xi\omega_{T-1}+\mathbb{E}_{T-2}^{\mathbb{P}}\left[\mathrm{d}_T\zeta n_{T-1}\right]\end{array}\right]\\
&=\mathbb{E}_t^{\mathbb{P}}\left[\begin{array}{l}\mathrm{a}_{T-1}+\sum_{i=2}^{p}\mathrm{b}_{T-1,i}\mathrm{CRV}_{T-i}+\sum_{j=2}^{q}\mathrm{c}_{T-1,j}\ell_{T-j}\\ +\mathrm{b}_{T-1,1}\left(\theta(d+\kappa)+\sum_{i=1}^{p}\theta\beta_i\mathrm{CRV}_{T-1-i}+\sum_{j=1}^{q}\theta\alpha_j\ell_{T-1-j}\right)\\ +\mathrm{c}_{T-1,1}\left(1+\gamma^2\left(\theta(d+\kappa)+\sum_{i=1}^{p}\theta\beta_i\mathrm{CRV}_{T-1-i}+\sum_{j=1}^{q}\theta\alpha_j\ell_{T-1-j}\right)\right)\\ +\mathrm{d}_T\xi\omega_{T-1}+\mathrm{d}_T\zeta\omega_{T-1}\end{array}\right]\\
&=\mathbb{E}_t^{\mathbb{P}}\left[\mathrm{a}_{T-2}+\sum_{i=1}^{p}\mathrm{b}_{T-2,i}\mathrm{CRV}_{T-1-i}+\sum_{j=1}^{q}\mathrm{c}_{T-2,j}\ell_{T-1-j}+\mathrm{d}_{T-1}\left(\xi\omega_{T-2}+\zeta n_{T-2}\right)\right].
\end{aligned}
\tag{51}
$$

Iterating the procedure above we finally obtain the expression

$$
\mathbb{E}_t^{\mathbb{P}}\left[\mathrm{RV}_T\right]=\mathrm{a}_t+\sum_{i=1}^{p}\mathrm{b}_{t,i}\mathrm{CRV}_{t+1-i}+\sum_{j=1}^{q}\mathrm{c}_{t,j}\ell_{t+1-j}+\mathrm{d}_{t+1}\left(\xi\omega_t+\zeta n_t\right)\tag{52}
$$

where the coefficients a_s, $\mathrm{b}_{s,i}$, $\mathrm{c}_{s,j}$, d_{s+1}, with $t\leq s\leq T-1$, are are given at each iteration by formulas (3.36) in Proposition 11, with initial condition $\mathrm{a}_{T-1}=\theta(d+\kappa)$, $\mathrm{b}_{T-1,i}=\theta\beta_i$, $\mathrm{c}_{T-1,j}=\theta\alpha_j$ and $\mathrm{d}_T=(\delta^2+\Lambda^2)$.

Bibliography

Aït-Sahalia, Y., Cacho-Diaz, J., Laeven, R. J., 2015. Modeling financial contagion using mutually exciting jump processes. Journal of Financial Economics 117 (3), 585 – 606.

Aït-Sahalia, Y., Jacod, J., 2014. High-Frequency Financial Econometrics. Princeton University Press.

Aït-Sahalia, Y., Mykland, P. A., Zhang, L., 2005. How often to sample a continuous-time process in the presence of market microstructure noise. Review of Financial Studies 18 (2), 351–416.

Aït-Sahalia, Y., Mykland, P. A., Zhang, L., 2011. Ultra high frequency volatility estimation with dependent microstructure noise. Journal of Econometrics 160 (1), 160–175.

Alitab, D., Bormetti, G., Corsi, F., Majewski, A. A., 2016. A Jump and Smile Ride: Continuous and Jump Variance Risk Premia in Option Pricing. Available at SSRN 2631155.

Alitab, D., Bormetti, G., Corsi, F., Majewski, A. A., 2017. Option pricing with realized volatility and persistent jump intensity. Preprint available upon request.

Andersen, T., Bollerslev, T., 1998. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. International Economic Review 39 (4), 885–905.

Andersen, T. G., Bollerslev, T., Diebold, F., 2007a. Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility. Review of Economics and Statistics 89, 701–720.

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

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

Andersen, T. G., Bollerslev, T., Dobrev, D., 2007b. No-arbitrage semi-martingale restrictions for continuous-time volatility models subject to leverage effects, jumps and i.i.d. noise: Theory and testable distributional implications. Journal of Financial Econometrics 138 (1), 125–180.

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.

Andersen, T. G., Dobrev, D., Schaumburg, E., 2012. Jump-robust volatility estimation using nearest neighbor truncation. Journal of Econometrics 169 (1), 75 – 93.

Bachelier, L., 1900. Théorie de la spéculation. Annales de l'Ecole Normale Supérieure 17, 21–86, English translation in Davis and Etheridge (2006).

Baillie, R., Bollerslev, T., Mikkelsen, H., 1996. Fractionally integrated generalized autoregressive conditional heteroskedasticity. Journal of econometrics 74 (1), 3–30.

Bakshi, C., Cao, C., Chen, Z., 1997. Empirical performance of alternative option pricing models. Journal of Finance 52, 1223–1258.

Bakshi, G., Kapadia, N., 2003. Delta-hedged gains and the negative market volatility risk premium. Review of Financial Studies 16 (2), 527.

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

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

Bandi, F. M., Renò, R., 2016. Price and volatility co-jumps. Journal of Financial Economics 119 (1), 107–146.

Bandi, F. M., Russell, J. R., 2011. Market microstructure noise, integrated variance estimators, and the accuracy of asymptotic approximations. Journal of Econometrics 160 (1), 145 – 159.

Bansal, R., Yaron, A., 2004. Risks for the long run: A potential resolution of asset pricing puzzles. The Journal of Finance 59 (4), 1481–1509.

Barndorff-Nielsen, O., 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 (Statistical Methodology) 63 (2), 167–241.

Barndorff-Nielsen, O., 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., Hansen, P. R., Lunde, A., Shephard, N., 2008. Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. Econometrica 76 (6), 1481–1536.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., Shephard, N., 2011a. Multivariate realised kernels: Consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. Journal of Econometrics 162 (2), 149 – 169.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., Shephard, N., 2011b. Subsampling realised kernels. Journal of Econometrics 160 (1), 204 – 219.

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., 2004. Power and bipower variation with stochastic volatility and jumps. Journal of Financial Econometrics 2 (1), 1–37.

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.

Barndorff-Nielsen, O. E., Shephard, N., 2006. Econometrics of testing for jumps in financial economics using bipower variation. Journal of Financial Econometrics 4 (1), 1–30.

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

Bates, D., 1996b. Testing Option Pricing Models, in Handbook of Statistics, Statistical Methods in Finance, G.S. Maddala and C.R. Rao (eds.), 567–611. Amsterdam: Elsevier.

Bates, D., 2000. Post-'87 crash fears in the S&P 500 futures option market. Journal of Econometrics 94 (1-2), 181–238.

Bates, D. S., 2008. The market for crash risk. Journal of Economic Dynamics and Control 32 (7), 2291–2321.

Bekaert, G., Hoerova, M., 2014. The VIX, the variance premium and stock market volatility. Journal of Econometrics 183 (2), 181–192.

Bertholon, H., Monfort, A., Pegoraro, F., 2008. Econometric asset pricing modelling. Journal of Financial Econometrics 6, 407–458.

Bibby, B. M., Sørensen, M., 2003. Hyperbolic processes in finance. Handbook of heavy tailed distributions in finance, 211–248.

Black, F., 1976. Studies of stock price volatility changes. In: Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economics Statistics Section. pp. 177–181.

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. The Journal of Political Economy, 637–654.

Bollerslev, T., 1986. Generalize autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–327.

Bollerslev, T., Tauchen, G., Zhou, H., 2009a. Expected stock returns and variance risk premia. Review of Financial Studies 22 (11), 4463–4492.

Bollerslev, T., Tauchen, G., Zhou, H., 2009b. 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., 2015. Tail risk premia and return predictability. Journal of Financial Economics 118 (1), 113 – 134.

Bormetti, G., Calcagnile, L. M., Treccani, M., Corsi, F., Marmi, S., Lillo, F., 2015a. Modelling systemic price cojumps with Hawkes factor models. Quantitative Finance 15 (7), 1137–1156.

Bormetti, G., Corsi, F., Majewski, A., A., 2015b. Term structure of variance risk premium in multi-component GARCH models. Preprint available at SSRN: http://papers.ssrn.com/sol3/papers.cfm?abstract id=2619278.

Breeden, D. T., 1979. An intertemporal asset pricing model with stochastic consumption and investment opportunities. Journal of Financial Economics 7 (3), 265–296.

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.

Brown, D. P., Jackwerth, J. C., 2012. The pricing kernel puzzle: Reconciling index option data and economic theory. Derivative securities pricing and modelling, 155–183.

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 (2), 171–186.

Campbell, J. Y., Cochrane, J. H., 1999. By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior. The Journal of Political Economy 107 (2), 205–251.

Campbell, J. Y., Shiller, R. J., 1988. The dividend-price ratio and expectations of future dividends and discount factors. The Review of Financial Studies 1 (3), 195.

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

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.

Carr, P., Wu, L., 2016. Analyzing volatility risk and risk premium in option contracts: A new theory. Journal of Financial Economics 120, 1–20.

Chabi-Yo, F., Garcia, R., Renault, E., 2008. State dependence can explain the risk aversion puzzle. Review of Financial Studies 21 (2), 973–1011.

Chen, K., Poon, S.-H., 2013. Variance swap premium under stochastic volatility and self-exciting jumps. Available at SSRN: https://ssrn.com/abstract=2200172 or http://dx.doi.org/10.2139/ssrn.2200172.

Christensen, K., Kinnebrock, S., Podolskij, M., 2010. Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data. Journal of Econometrics 159 (1), 116–133.

Christensen, K., Oomen, R. C., Podolskij, M., 2014. Fact or friction: Jumps at ultra high frequency. Journal of Financial Economics 114 (3), 576–599.

Christensen, K., Oomen, R. C., Renò, R., 2016. The drift burst hypothesis. Available at SSRN: ssrn.com/abstract=2842535.

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., Feunou, B., Jeon, Y., 2015. Option valuation with observable volatility and jump dynamics. Journal of Banking & Finance 61, S101–S120.

Christoffersen, P., Heston, S., Jacobs, K., 2013. Capturing option anomalies with a variance-dependent pricing kernel. Review of Financial Studies 26 (8), 1962–2006.

Christoffersen, P., Jacobs, K., Ornthanalai, C., 2012. Dynamic jump intensities and risk premia: Evidence from S&P 500 returns and options. Journal of Financial Economics 106 (3), 447–472.

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.

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

Cont, R., 1998. Statistical Finance: Emprical and theoretical approaches to the statistical modelling of price variations in speculative markets. Doctoral Thesis, Université de Paris XI.

Cont, R., 2001. Empirical properties of asset returns: Stylized facts and statistical issues. Quantitative Finance 1 (2), 223–236.

Cont, R., Potters, M., Bouchaud, J.-P., 1997. Scaling in stock market data: Stable laws and beyond. In: B. Dubrulle, F. Graner, D. Sornette (eds.) Scale Invariance and Beyond. Springer.

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

Corsi, F., Audrino, F., Renò, R., 2012. HAR Modeling for Realized Volatility Forecasting. Handbook of Volatility Models and Their Applications. John Wiley & Sons, Inc., pp. 363–382.

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., Ross, S. A., Rubinstein, M., 1979. Option pricing: A simplified approach. Journal of Financial Economics 7 (3), 229 – 263.

Darolles, S., Gourieroux, C., Jasiak, J., 2006. Structural Laplace transform and compound autoregressive models. Journal of Time Series Analysis 27, 477–503.

Davis, M., Etheridge, A., 2006. Louis Bachelier's Theory of Speculation: The Origins of Modern Finance. Princeton University Press.

Derman, E., Kani, I., 1994. Riding on a smile. Risk 7 (2), 32–39.

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., 2004. Jump starting GARCH: Pricing and hedging options with jumps in returns and volatilities. Unpublished working paper. University of Toronto, Rotman School, Toronto, Canada.

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., Filipović, D., Schachermayer, W., 08 2003. Affine processes and applications in finance. Ann. Appl. Probab. 13 (3), 984–1053.

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

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

Ebens, H., 1999. Realized stock index volatility. Working Paper No. 420. Department of Economics, Johns Hopkins University, Baltimore.

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

Engle, R., Lee, G., 1999. A permanent and transitory component model of stock return volatility, in ed. R. Engle and H. White Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive W. J. Granger.

Engle, R., Ng, V. K., 1993. Measuring and testing the impact of news on volatility. Journal of Finance 48 (5), 1749–78.

Eraker, B., 2008. Affine general equilibrium models. Management Science 54 (12), 2068–2080.

Esscher, F., 1932. On the probability function in the collective theory of risk. Scandinavian Actuarial Journal 1932 (3), 175–195.

Fama, E. F., 1965. The behavior of stock-market prices. The Journal of Business 38 (1), 34–105.

Fama, E. F., French, K. R., 1988. Dividend yields and expected stock returns. Journal of Financial Economics 22 (1), 3 – 25.

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.

Gagliardini, P., Gouriéroux, C., Renault, E., 2011. Efficient derivative pricing by the extended method of moments. Econometrica 79 (4), 1181–1232.

Garcia, R., Renault, E., Semenov, A., 2006. Disentangling risk aversion and intertemporal substitution through a reference level. Finance Research Letters 3 (3), 181 – 193.

Gatheral, J., 2006. The volatility surface. John Wiley & Sons, Inc.

Gatheral, J., Jaisson, T., Rosenbaum, M., 2014. Volatility is rough. arXiv preprint arXiv:1410.3394.

Gerber, H. U., Shiu, E. S., 1994. Option pricing by Esscher transforms. Transactions of the Society of Actuaries 46 (99), 140.

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.

Gourieroux, C., Monfort, A., Polimenis, V., 2002. Affine Term Structure Models. CREST Working Papers 2002-49.

Gourieroux, C., Monfort, A., Polimenis, V., 2006. Affine models for credit risk analysis. Journal of Financial Econometrics 4 (3), 494–530.

Gourieroux, C., Monfort, A., Sufana, R., 2010. International money and stock market contingent claims. Journal of International Money and Finance 29 (8), 1727–1751.

Hagan, P. S., Kumar, D., Lesniewski, A. S., Woodward, D. E., 2002. Managing smile risk. Wilmott Magazine 1, 84–108.

Hansen, P., Large, J., Lunde, A., 2008. Moving average-based estimators of integrated variance. Econometric Reviews 27 (1-3), 79–111.

Hautsch, N., Podolskij, M., 2013. Preaveraging-based estimation of quadratic variation in the presence of noise and jumps: theory, implementation, and empirical evidence. Journal of Business & Economic Statistics 31 (2), 165–183.

Heston, 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.

Huang, X., Tauchen, G., 2005. The relative contribution of jumps to total price variance. Journal of Financial Econometrics 3 (4), 456–499.

Hull, J., White, A., 1987. The pricing of options on assets with stochastic volatilities. The Journal of Finance 42 (2), 281–300.

Jackwerth, J., 2000. Recovering risk aversion from option prices and realized returns. Review of Financial Studies 13 (2), 433–51.

Jacod, J., Li, Y., Mykland, P. A., Podolskij, M., Vetter, M., 2009. Microstructure noise in the continuous case: the pre-averaging approach. Stochastic processes and their applications 119 (7), 2249–2276.

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.

Lee, S. S., Mykland, P. A., 2008. Jumps in financial markets: A new nonparametric test and jump dynamics. Review of Financial Studies 21 (6), 2535–2563.

Lettau, M., Ludvigson, S., 2001. Consumption, aggregate wealth, and expected stock returns. The Journal of Finance 56 (3), 815–849.

Liu, Y., Cizeau, P., Meyer, M., Peng, C.-K., Stanley, E. H., 1997. Correlations in economic time series. Physica A 245 (3), 437–440.

Lowenstein, R., 2000. When Genius Failed: The Rise and Fall of Long-Term Capital Management. Random House Publishing Group.

Lucas, R. E., 1978. Asset prices in an exchange economy. Econometrica 46 (6), 1429–1445.

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 (2), 521 – 531.

Mancini, C., 2009. Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. Scandinavian Journal of Statistics 36 (2), 270–296.

Mandelbrot, B., 1963. The variation of certain speculative prices. The Journal of Business 36, 394–419.

Mankiw, N. G., Shapiro, M. D., 1986. Risk and return: Consumption beta versus market beta. The Review of Economics and Statistics 68 (3), 452–459.

McNees, S. K., 1979. The forecasting record for the 1970s. New England Economic Review. September/October 1979, 33–53.

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.

Muller, 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.

Poteshman, A. M., 2001. Underreaction, overreaction, and increasing misreaction to information in the options market. The Journal of Finance 56 (3), 851–876.

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

Rendleman, R. J., Bartter, B. J., 1979. Two-state option pricing. The Journal of Finance 34 (5), 1093–1110.

Rubinstein, M., 1976. The valuation of uncertain income streams and the pricing of options. The Bell Journal of Economics 7 (2), 407–425.

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

Stein, E. M., Stein, J. C., 1991. Stock price distributions with stochastic volatility: an analytic approach. Review of financial Studies 4 (4), 727–752.

Stein, J., 1989. Overreactions in the Options Market. Journal of Finance 44 (4), 1011–23.

Stentoft, L., 2008. Option pricing using realized volatility, working Paper at CREATES, University of Copenhagen.

Veronesi, P., 2001. Belief-dependent utilities, aversion to state-uncertainty and asset prices. CRSP Working Paper.

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.

Zhou, G., Yingzi, Z., 2009. A long-run risks model with long- and short-run volatilities: Explaining predictability and volatility risk premium. Working paper, Olin School of Business.

Zhou, H., 2010. Variance risk premia, asset predictability puzzles, and macroeconomic uncertainty. Working paper, Federal Reserve Board.