A Jump and Smile Ride:
Jump and Variance Risk Premia in Option Pricing

Dario Alitab[a,*], Giacomo Bormetti[b], Fulvio Corsi[c,d],
and Adam A. Majewski[e]

September 19, 2018

[a] *Mediobanca S.p.A, Piazzetta E. Cuccia 1, 20121 Milano, Italy*
[b] *University of Bologna, Piazza di Porta San Donato 5, 40126 Bologna, Italy*
[c] *University of Pisa, via Ridolfi 10, 56100 Pisa, Italy*
[d] *City University London, Northampton Square, London EC1V 0HB, United Kingdom*
[e] *Capital Fund Management, 23 Rue de l'Université, 75007 Paris, France*

Abstract

We introduce a discrete-time model for log-return dynamics with observable volatility and jumps. Our proposal extends the class of Realized Volatility heterogeneous autoregressive gamma (HARG) processes adding a jump component with time-varying intensity. The model is able to reproduce the temporary increase in the probability of occurrence of a jump immediately after an abrupt large movement of the asset price. Belonging to the class of exponentially affine models, the moment generating function under the physical measure is available in closed-form. Thanks to a flexible specification of the pricing kernel compensating for equity, volatility, and jump risks, the generating function

*Corresponding author: dario.alitab@gmail.com. The opinions expressed here are solely those of the authors and do not represent in any way those of their employers.

under the risk-neutral measure inherits analytical tractability too. An application of the leveraged HARG model with dynamic jump intensity to the pricing of a large sample of S&P500 index options assesses its superior performances with respect to state-of-the-art benchmark models.

Keywords: High-frequency, Realized volatility, HARG, Option pricing, Variance risk premium, Jumps
JEL Classification: G12, G13

1 Introduction

It has been well documented that even in the most liquid financial markets, assets feature abrupt price movements. Whether these events are true jump events attributed to the discontinuous component of the price semi-martingale process could be debated at length, see for instance Aït-Sahalia (2004); Cont and Tankov (2004); Maheu and McCurdy (2004); Barndorff-Nielsen and Shephard (2007); Bollerslev et al. (2008); Lee and Mykland (2008); Aït-Sahalia and Jacod (2009); Bollerslev et al. (2009). Recent achievements suggest alternative hypotheses. Christensen et al. (2014) argue that jump events are often spurious detections resulting from the aggregation of returns at larger time scales. Christensen et al. (2016) 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. Whatever the origin of extreme movements and despite their infrequent appearance, a typical observed phenomenon is their clustering behaviour. An unexpected jump event triggers subsequent jumps for a given asset as discussed in Christoffersen et al. (2012); Chen and Poon (2013); Christoffersen et al. (2015); Bormetti et al. (2015) or sparks a contagion-like spreading across different assets, see among others Ding et al. (2009); Aït-Sahalia et al. (2015); Fičura (2015); Azizpour et al. (2018); Calcagnile et al. (2018). As a consequence, widespread models describing jumps as processes with independent occurrences are not suited to accurately describe the empirical properties of asset returns.

In this paper we propose a new discrete-time model with observable volatility and jumps whose intensity is time-varying and persistent. Precisely, the local intensity follows an autoregressive process of order one where the realized number of jumps detected each day plays the role of an idiosyncratic shock. Formally, we extend the LHARG-RV class of models in Majewski et al. (2015) by adding to log-returns a jump component. The inclusion of jumps with persistent intensity is adequate to reproduce the clustering of extreme events in the physical process and to contemporary increase the short time-to-maturity implied volatility extracted from option quotes. These features improve both the likelihood of the process under the physical measure and the pricing performances under the risk-neutral measure.

The model we present widens the literature combining realized volatility measures with option pricing initiated by Stentoft (2008) and later extended by Corsi et al. (2013), Christoffersen et al. (2014) and Majewski et al. (2015). Specifically, we assume that the log-return dynamics is determined by the sum of two independent random processes accounting for the diffusive and discontinuous components of the price. The former corresponds to a sequence of random shocks which are normally distributed, with observable conditional variance given by the realized measure of the diffusive quadratic variation. The Realized Variance term follows an autoregressive gamma process (see Gourieroux and Jasiak (2006)) whose conditional mean is a linear function of the past realized variances and leverage terms aggregated over different time scales (daily, weekly, and monthly). The discontinuous component corresponds to a compound non homogeneous Poisson process whose intensity linearly depends on its first-order lagged value. The number of jump events detected in the preceding day provides the idiosyncratic shock to the intensity dynamics. Besides, the jump size is sampled from a normal distribution. The new model is dubbed Heterogeneous Auto-Regressive Gamma model for Realized Volatility with Leverage and Auto-Regressive Jumps, LHARG-ARJ for short. This model belongs to the class of exponentially affine processes, for which the recursive expression of the log-return moment generating function (MGF) under the physical measure is available in closed-form.

To price options, we compute recursively the MGF under the risk neutral measure. The

change of measure is performed adopting the same approach of Gerber and Shiu (1994); Christoffersen et al. (2009); Gagliardini et al. (2011); Corsi et al. (2013); Christoffersen et al. (2012, 2014, 2015) based on the specification of a discrete-time exponential affine stochastic discount factor (SDF) incorporating multiple risk premia. The SDF depends on four different risk premia in order to account for the different sources of risk entering our model: Two for the directional continuous and discontinuous movements of returns, and two for the non directional risk, i.e. the continuous and jump components of the Realized Volatility. Our specification of the multi-dimensional SDF improves the flexibility of the option pricing model under the risk neutral measure while preserving the analytical tractability. Indeed, we are able to derive the risk-neutral MGF and to show that the risk-neutral dynamics still belongs to the LHARG-ARJ model class. The latter result is achieved proving the existence of a one-to-one relation which maps each parameter describing the physical dynamics to the corresponding parameter for the risk-neutral dynamics.

Belonging to the class of models with observable variance components, LHARG-ARJ inherits the advantage of ease of estimation. In fact, the RV process is directly built from high-frequency log-returns, without the need of any filtering procedure. We compute the RV time series from tick-by-tick returns for the Standard and S&P500 Index Futures, from July 3, 1990 to June 28, 2011. In order to separate the continuous from the jump contributions of the log-return time series, we exploit the methodology introduced in Corsi et al. (2010). The test is based on the Threshold Bipower Variation and allows to detect the presence of at least one jump in a given trading day. In our model, this information is not sufficient since we need to recover the exact number of jumps occurring each day. Following the iterative procedure proposed by Andersen et al. (2010), we identify every jump occurring at intra-day level. With the time series for the continuous and discontinuous RV components, the number of jumps, and their size, we finally estimate the parameters of the LHARG-ARJ process by means of the Maximum Likelihood Estimator (MLE). The introduction of the jump intensity dynamics improves the likelihood of the model with respect to the benchmark LHARG model by Majewski et al. (2015).

Applying our analytically tractable model to a large sample of index options, we demonstrate its superior ability to price options. As benchmark models, we choose the LHARG-RV model by Majewski et al. (2015) and the model with Realized Volatility and realized jump variation of Christoffersen et al. (2015). We perform our analysis on plain vanilla options written on S&P500 Index whose valuation is given each Wednesday from January 1, 1996 to June 28, 2011. We calibrate the premia on the whole implied volatility surfaces and compute the option prices using the effective numerical COS method introduced by Fang and Oosterlee (2008). The results show that the LHARG-ARJ model represents a valid competitor class to state-of-the-art discrete-time models.

Our approach has several aspects in common with Christoffersen et al. (2012, 2014, 2015). Christoffersen et al. (2012) introduce a model for log-returns with latent conditional variance and a jump component with dynamic intensity. The two dimensional affine specification of the pricing kernel compensates for the continuous and jump directional shocks and allows to obtain closed-form recursive relations for the MGF under both physical and risk-neutral measures. However, the paper does not exploit any information coming from high frequency data, and the joint estimation-calibration is performed by means of Quasi Maximum Likelihood Estimation (QMLE). The economic value of including high-frequency information from tick-by-tick log-returns is investigated in Christoffersen et al. (2014). The paper does not consider any discontinuous component, but the inference on the latent two-component variance process is improved by the introduction of an observation equation which exploits the observable Realized Variance. Recently, Christoffersen et al. (2015) combine the insights gained from the two previous works, and present a model with daily realized BiPower and Jump Variation measures which is flexible and analytically tractable, dubbed BPJVM. Among the three, the latter work is the most similar to our approach but a closer look reveals several important differences (commented in more detail in the next sections). The first relevant difference is the method employed to identify and separate the continuous and jump components of the integrated variance. Christoffersen et al. (2015) compute a proxy of the continuous component of volatility by

means of the Bipower Variation from 5-minute returns and the jump contribution corresponds to the difference, when positive, between the Realized Variance and the Bipower Variation. The methodology does not consider any statistical test in order to assess the significance of the jump contribution. The literature – see for example Barndorff-Nielsen and Shephard (2004), Barndorff-Nielsen and Shephard (2006), Andersen et al. (2007), Corsi et al. (2010) – warns about the bias in the estimation of the continuous component of the integrated variance in finite sample, especially in presence of successive jump events. A second major difference is that the approach by Christoffersen and co-authors may be viewed as an improved and extended version of the Realized GARCH approach of Hansen et al. (2012), while the LHARG-ARJ extends the class of RV gamma models Corsi et al. (2013); Majewski et al. (2015). The role played by the observable realized measures in the two classes is essentially different. In the former, the conditional variance is a latent process with idiosyncratic shocks given by the RV measure – in the same spirit of the Realized GARCH. The latter directly models the dynamics of the RV components. The impact of the two modeling choices is relevant not only on the estimation methodology – which is based on QMLE for the BPJVM and on MLE for the LHARG-ARJ – but also, and more importantly, on the level of persistence of the conditional variance in the two models. The persistence of the BPJVM latent variance is nearly one, then a miss-specification of the current level of the volatility may lead to miss-fit the term structure of at-the-money (ATM) implied volatility, especially at longer maturities.

To summarize, the contribution of this paper is threefold. First, we introduce a model to describe the dynamics of asset prices including a multi-component structure for volatility and leverage and a novel observable jump component with persistent intensity. Second, we derive analytical formulas for the MGF under both physical and risk-neutral measures, and the no-arbitrage condition. This result is achieved by means of a flexible specification of the pricing kernel which compensates for all directional and non directional random components appearing in the return dynamics. Finally, we discuss the ability of our approach to price a large sample of S&P500 index options and benchmark the result with the LHARG by Majewski et al. (2015)

and the state-of-the-art BPJVM model of Christoffersen et al. (2015). When compared with the LHARG model, the LHARG-ARJ improves the pricing performances especially for short time-to-maturity options and out-of-the-money regions. Concerning the BPJVM, the LHARG-ARJ model fares much better than the benchmark model at longer maturities and for deep out-of-the-money options.

The rest of the paper is organized as follows. In Section 2 we introduce the LHARG-ARJ model of asset price dynamics and we derive the log-return MGF. We propose a change of measure based on a four-dimensional pricing kernel which takes into account both equity, volatility, and jump risk premia. Consistently, we derive the no-arbitrage condition and we show that risk-neutral dynamics still belongs to the LHARG-ARJ class of models. Section 3 discusses estimation of the model parameters. In Section 4, we describe the results of a comparative assessment of option pricing performances using the LHARG-ARJ model along with concurrent benchmarks. Section 5 draws the relevant conclusions.

2 The model

2.1 The motivation

Let us assume that log-return dynamics is defined on some stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ and described by a jump-diffusion process

$$Y(t) = \alpha(t) + \int_0^t \sigma(s) dW(s) + \sum_{i=1}^{N(t)} X_i, \tag{2.1}$$

where $W(t)$ is a Brownian motion, $\sigma(t)$ is an $\mathcal{F}_t$-measurable stochastic process corresponding to the continuous part of price volatility, $N(t)$ is an $\mathcal{F}_t$-measurable stochastic process describing the total number of jumps in price till time t and X_i are i.i.d. random variables capturing the size and direction of jumps. At this point, we only assume that $X_i \sim \mathcal{N}(\Lambda, \delta^2)$ without specifying any particular distribution of $N(t)$ or $\sigma(t)$.

The drift term process $\alpha(t)$ represents the reward that investors demand for bearing risks related to the jump-diffusive nature of price. An investor faces two types of risks related to price variation: One due to continuous, normal price movements, $\sigma(s)dW(s)$ and another one due to jumps, i.e. extreme market events, X_i, where $i \in \{1, \ldots, N(t)\}$. Using Itô isometry and assuming that $\text{Var}[N(t)] = \mathbb{E}[N(t)]$,[1] the components of log-return variance can be represented as

$$\text{Var}\left[\int_0^t \sigma(s)dW(s)\right] = \mathbb{E}\left[\int_0^t \sigma^2(s)ds\right] \quad \text{and} \quad \text{Var}\left[\sum_{i=1}^{N(t)} X_i\right] = \mathbb{E}\left[\sum_{i=1}^{N(t)} |X_i|^2\right].$$

Since jump and diffusion processes are independent, we obtain that log-return variance conditioned on $\alpha(t)$ can be written as a sum of expected integrated variance (IV) and jump variation (JV)

$$\text{IV}(t) = \int_0^t \sigma^2(s)ds \quad \text{and} \quad \text{JV}(t) = \sum_{i=1}^{N(t)} |X_i|^2.$$

The expected jump variation is equal to

$$\mathbb{E}[\text{JV}(t)] = \left(\Lambda^2 + \delta^2\right)\mathbb{E}[N(t)].$$

Assuming that investors are risk-averse to both types of price variation, we can write the drift process as

$$\alpha(t) = rt + \left(\lambda_c - \frac{1}{2}\right)\text{IV}(t) + (\lambda_j - \eta)\left(\Lambda^2 + \delta^2\right)N(t), \tag{2.2}$$

where r is the risk-free rate in the economy, λ_c and λ_j are the continuous and jump, respectively, components of equity risk premium, and

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

The compensation of two equity premia λ_c and λ_j in (2.2) by $\frac{1}{2}$ and η, respectively, ensures

[1] This assumption is satisfied by Poisson process.

that the conditional expectation of price process $S(t) = S(0)\mathrm{e}^{Y(t)}$ reads

$$\mathbb{E}\left[S(t)|\mathrm{IV}(t), N(t)\right] = S(0)\exp\left(rt + \lambda_c \mathrm{IV}(t) + \lambda_j\left(\Lambda^2 + \delta^2\right)N(t)\right).$$

If the market is neutral to the risk associated to continuous volatility and jumps ($\lambda_c = \lambda_j = 0$), the discounted price process is a martingale.

Since diffusion part in equation (2.1) can be seen as time-changed Brownian motion (see Ané and Geman (2000)), we can describe the dynamics of log-returns as

$$Y(t) = rt + \left(\lambda_c - \frac{1}{2}\right)\mathrm{IV}(t) + (\lambda_j - \eta)\left(\Lambda^2 + \delta^2\right)N(t) + W\left(\mathrm{IV}(t)\right) + \sum_{i=1}^{N(t)} X_i.$$

The major motivation of our work is that, exploiting the whole information from tick-by-tick data, the risk factors driving the log-return dynamics are observable. In Section 3.1 we describe the econometric procedure that enables us to measure the continuous component of realized variance, CRV, on a daily scale. Furthermore, the proposed procedure enables us to observe the number of jumps $N(t)$ and their size and direction, X_i.

2.2 The model dynamics

Daily log-returns, y_t, have the following dynamics

$$y_t = r + \left(\lambda_c - \frac{1}{2}\right)\mathrm{CRV}_t + (\lambda_j - \eta)\left(\Lambda^2 + \delta^2\right)n_t + \sqrt{\mathrm{CRV}_t}\epsilon_t + \sum_{i=1}^{n_t} X_{t,i}, \tag{2.3}$$

where ϵ_t are i.i.d. with standard normal random variables. CRV_t is the estimator of the continuous component of integrated variance on day t, n_t is the number of intra-day jumps at day t and $X_{t,i}$ are the jump sizes during day t, where $i \in \{1, \ldots, n_t\}$.

The continuous component of realized variance, CRV_t, follows the LHARG dynamics introduced by Majewski et al. (2015). CRV_{t+1} conditioned on information at day t is sampled from

a non-centred gamma distribution

$$\mathrm{CRV}_{t+1}|\mathcal{F}_t \sim \bar{\gamma}(\kappa, \Theta(\mathbf{CRV}_t, \mathbf{L}_t), \theta)\,, \tag{2.4}$$

with $\mathbf{CRV}_t = (\mathrm{CRV}_0, \ldots, \mathrm{CRV}_{t-1}, \mathrm{CRV}_t)$, $\mathbf{L}_t = \left(\ell_0^{(d)}, \ldots, \ell_{t-1}^{(d)}, \ell_t^{(d)}\right)$ and

$$\Theta(\mathbf{CRV}_t, \mathbf{L}_t) = \beta_d \mathrm{CRV}_t^{(d)} + \beta_w \mathrm{CRV}_t^{(w)} + \beta_m \mathrm{CRV}_t^{(m)} + \alpha_d \ell_t^{(d)} + \alpha_w \ell_t^{(w)} + \alpha_m \ell_t^{(m)}\,. \tag{2.5}$$

In the previous equation, the quantities

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

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.

We model the jump component of daily log-return as a compound Poisson process with i.i.d. normally distributed jumps

$$n_t|\omega_t \sim \mathrm{Poisson}\,(\omega_t) \quad \text{and} \quad X_{t,i} \sim \mathcal{N}(\Lambda, \delta^2), \tag{2.6}$$

where $\Lambda \in \mathbb{R}$ and $\delta \in \mathbb{R}^+$ are parameters specifying the mean and standard deviation of the jump size, respectively. The expected number of jumps depends on the time-varying intensity ω_t, whose dynamics is described by an auto-regressive process

$$\omega_{t+1} = \bar{\omega} + \xi\omega_t + \zeta n_t\,, \tag{2.7}$$

where $\bar{\omega}, \xi, \zeta$ are strictly positive parameters of the model. As follows from equation (2.7), the

jump intensity on day $t+1$ depends on the intensity ω_t and on the number of intra-day jumps (n_t) occurred the day before. This structure allows an extreme event to increase locally the intensity of the jump process in the following days. In the present setting, it is important to notice that the shock n_t is observable. This is similar in spirit to what has been proposed in Christoffersen et al. (2015), where the jump intensity follows an autoregressive process of order one but shock are represented by the jump component of the integrated daily volatility. In this way, the intensity process is affected not only by the realized number of intra-day jumps on a given day, but also by the square of their size. In our model, we disentangle the effect of realized jump frequency from the realized jump size and direction, and only the former determines the evolution of the intensity process. In this respect, our model mimics, in discrete time, the continuous-time self-exciting dynamics associated with non-marked exponential processes termed Hawkes processes, see Hawkes (1971); Daley and Vere-Jones (2003).

The model described by equations (2.3)-(2.7) has LHARG dynamics for the continuous component of realized variance and Auto-Regressive Jump intensity. Let us stress once again that all variables, apart from the innovation process ϵ_t, are observable. This property will significantly simplify the estimation procedure in comparison to alternative models with latent volatility and jump processes.

2.3 Statistical properties of the model

Firstly, let us observe that the conditionally expected daily return is given by

$$\mathbb{E}\left[\exp\left(y_{t+1}\right)|\mathrm{CRV}_{t+1}, n_{t+1}\right] = \exp(r + \lambda_c \mathrm{CRV}_{t+1} + \lambda_j \left(\Lambda^2 + \delta^2\right) n_{t+1}), \tag{2.8}$$

and if both equity risk premia are zero ($\lambda_c = \lambda_j = 0$), then, as for the continuous case, the price process corrected by the risk-free rate is a martingale.

As we have observed in the previous section, the structure of the dynamics in (2.7) includes the mechanism that increases the intensity of the jump process conditionally to the number of

extreme events happened the day before. This translates into positive correlation between the intensity and the number of jumps at the day before

$$\mathrm{Cov}_{t-1}\left(n_t, \omega_{t+1}\right) = \zeta\omega_t\,.$$

This mechanism introduces jump clustering in the model. Moreover, the covariance between log-return and the next day jump intensity is given by

$$\begin{aligned}\mathrm{Cov}_{t-1}\left(y_t, \omega_{t+1}\right) &= \left(\left(\lambda_j - \eta\right)\left(\Lambda^2 + \delta^2\right) + \Lambda\right)\zeta\omega_t, \\ &= \left(\lambda_j\left(\Lambda^2 + \delta^2\right) - 0.5\delta^2\right)\zeta\omega_t\,.\end{aligned} \tag{2.9}$$

By definition, the jump component of the market price of equity risk is non negative. If λ_j is zero or sufficiently small, then the above covariation is expected to be negative. This means that, on average, after a negative return the intensity is larger than after a positive return. Let us define the jump variation as

$$\mathrm{JRV}_t \doteq \sum_{i=1}^{n_t} |X_{t,i}|^2. \tag{2.10}$$

Then, as a consequence of (2.9) and for zero or small λ_j, the increase in jump variation is more likely after a negative shock in log-returns rather than after a positive one. This mechanism acts in a similar way to the leverage effect in the volatility modeling where the continuous component of the variance determines the negative correlation between past log-returns and future variances. Computing the formulae for the covariance between log-returns and variance components in the LHARG-ARJ model, we obtain

$$\begin{aligned}\mathrm{Cov}_{t-1}\left(y_t, \mathrm{JRV}_{t+1}\right) &= \left(\lambda_j\left(\Lambda^2 + \delta^2\right) - 0.5\delta^2\right)\left(\Lambda^2 + \delta^2\right)\zeta\omega_t, \\ \mathrm{Cov}_{t-1}\left(y_t, \mathrm{CRV}_{t+1}\right) &= -2\theta^2\alpha_d\gamma\left(\delta + \Theta(\mathbf{CRV}_{t-1}, \mathbf{L}_{t-1})\right)\,.\end{aligned} \tag{2.11}$$

By construction, the continuous and jump components of realized variance are uncorrelated.

A great advantage of the LHARG-ARJ model is that it inherits the affine property from

the LHARG class. Then, the conditional MGF takes an exponential-affine form, available in closed-form through a set of recursive relations.

Proposition 1 (Moment Generating Function)**.** *The MGF of the log-return* $y_{t,T} = \sum_{i=0}^{T-1} y_{t+i}$ *for the LHARG-ARJ model conditioned on the information available at time* t *is of the form*

$$\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{z y_{t,T}} | \mathcal{F}_t\right] = \exp\left(\mathrm{a}_t + \sum_{i=1}^{22} \mathrm{b}_{t,i}\mathrm{CRV}_{t+1-i} + \sum_{j=1}^{22} \mathrm{c}_{t,j}\ell_{t+1-j} + \mathrm{d}_{t+1}\left(\xi\omega_t + \zeta n_t\right)\right), \tag{2.12}$$

where the coefficients a_t*,* $\mathrm{b}_{t,i}$ *for* $i = 1, \ldots, 22$*,* $\mathrm{c}_{t,j}$ *for* $j = 1, \ldots, 22$*, and* d_{t+1} *satisfy the recursive relations given in (A.2).*

Proof: See Appendix A.

2.4 Risk-neutral dynamics

The standard problem of option pricing in incomplete markets is the specification of the stochastic discount factor. The shape of the pricing kernel determines the form of the risk-neutral measure. The latter then depends on the investor's attitude to the risk associated with uncertain future levels of returns, of variance, and possibility of large and sudden price variation. In this study we perform the risk-neutralization of the objective measure introducing a four-dimensional Esscher transform. This choice has three relevant advantages: It delivers a clear financial interpretation of the parameters, the resulting dynamics corresponds to a LHARG-ARJ process with risk-neutral parameters given by one-to-one mapping of the historical ones, and, finally, it provides a semi-closed expression for the log-return MGF under the risk-neutral measure.

We model the pricing kernel in terms of the Esscher transform

$$M_{t-1,t} = \frac{\mathrm{e}^{-\nu_c \mathrm{CRV}_t - \nu_j \mathrm{JRV}_t - \mu_c\sqrt{\mathrm{CRV}_t}\epsilon_t - \mu_j \sum_{i=1}^{n_t} X_{t,i}}}{\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{-\nu_c \mathrm{CRV}_t - \nu_j \mathrm{JRV}_t - \mu_c\sqrt{\mathrm{CRV}_t}\epsilon_t - \mu_j \sum_{i=1}^{n_t} X_{t,i}} \middle| \mathcal{F}_{t-1}\right]}, \tag{2.13}$$

with four parameters responsible for different risk premia. The value of ν_c and ν_j determines the

level of variance risk premium. The former is associated to CRV_t and, as it is also clear from the decomposition of the drift term in the return equation (2.3), represents the compensation that an investor requires from an investment with uncertain future level of the continuous component of variance. The latter compensates for JRV_t, i.e. the component of the integrated variance attributed to the discontinuous part of the log-return process. The remaining two parameters, μ_c and μ_j, determine the level of the equity risk premium. As before, the component $\mu_c\sqrt{\mathrm{CRV}_t}\epsilon_t$ remunerates the risk related to continuous directional changes in price, whereas the component $\mu_j \sum_{i=1}^{n_t} X_{t,i}$ captures the risk related to abrupt and large directional price changes. It is worth to comment that the pricing kernel is modelled in an apparently similar fashion in Christoffersen et al. (2015) (please refer to equation (18) in their paper) . However, there are several relevant differences which make non trivial a direct comparison between the two Esscher transforms. In the present framework, the number of risk premia is four, while in Christoffersen et al. (2015) it is equal to three. All the latter premia correspond to directional movements. Recalling the notation of the paper, $\nu_{1,t}$ and $\nu_{2,t}$ compensate for two idiosyncratic exogenous shocks $\epsilon_{1,t}$ and $\epsilon_{2,t}$ associated to the continuous component of the price process – the approach requires two components to reproduce the leverage effect in a similar fashion to stochastic volatility models – whereas ν_3 remunerates large directional price changes. Stated differently, the pricing of the uncertainty of future levels of volatility is not explicit but enters indirectly through the pricing of all directional components. Finally, the premia are time-varying and to ensure the affinity of the model under the $\mathbb{Q}$ measure the premium $\nu_{2,t}$ depends deterministically on $\nu_{1,t}$ and on the latent process of the conditional volatility. This reduces effectively the number of premia to $\nu_{1,t}$ and ν_3, and the number of free parameters to calibrate to two. In the LHARG-ARJ approach, the no-arbitrage restrictions fix the level of the directional premia, while the non directional ones, ν_c and ν_j, have to be calibrated on option quotes.

In order to play the role of the stochastic discount factor in our economy, the transform (2.13) has to guarantee absence of arbitrage opportunity. The following result holds.

Proposition 2 (No-arbitrage restriction)**.** *If the dynamics of the underlying price is described*

by (2.3)-(2.7), then the Esscher transform (2.13) is a stochastic discount factor if and only if the following conditions are satisfied

$$\mu_c = \lambda \quad \textit{and} \quad \mu_j = \frac{1}{2} + \frac{\Lambda + (\lambda_j - \eta)\left(\Lambda^2 + \delta^2\right)\left(1 + 2\nu_j\delta^2\right)}{\delta^2}. \tag{2.14}$$

Proof: See Appendix B.

From relation (2.14) 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. The estimation of the model parameters, from daily returns and realized measures of volatility, combined with the calibration of the premia parametrising the arbitrage-restricted pricing kernel allows the model to reconcile the time series properties of stock returns with the empirical properties of option panels.

Performing the change of measure by means of the Esscher transform (2.13) provides the risk-neutral dynamics which governs the return process. Remarkably, the risk-neutral log-returns follow a LHARG-ARJ process, whose parameters, denoted with a star, are readily obtained from the historical counterparts through a simple one-to-one mapping relation.

Proposition 3. *Under the risk-neutral measure $\mathbb{Q}$ - corresponding to the SDF specification given by (2.13) - the log-return dynamics for the LHARG-ARJ model is governed by equations (2.3)-(2.7) with parameters λ^*, 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 C (equations (C.4) and (C.7)).*

Proof: See Appendix C.

Given the dynamics under $\mathbb{Q}$, the risk-neutral MGF readily follows as a straightforward consequence of Proposition 1.

Corollary 4. *Under $\mathbb{Q}$, the MGF for the LHARG-ARJ model has the same form as in (2.12) and (A.2) with parameters as in (C.4) and (C.7).*

As we detail in the next section, the existence of an analytic expression for the MGF under the risk-neutral measure allows to perform the calibration of the variance risk premia with very effective and reliable numerical methods based on the Fourier transform.

3 Model estimation

This section is dedicated to the estimation of parameters of the LHARG-ARJ model. Since the Realized Variance estimator is directly built from observed High-Frequency (HF) returns, it prevents the use of filtering procedure for determining the latent volatility process. Moreover, we introduce the state-of-the-art models by Majewski et al. (2015) and Christoffersen et al. (2015). They will be used as benchmark models for the LHARG-ARJ.

3.1 Return variation measurement and jump detection procedure

Our data set consists of HF returns of the S&P500 Futures Index from 1 July 1990 to 31 June 2011 provided by TickData. To build the series of Realized Variance we adopt the Two-Scale method proposed by Zhang et al. (2005) which has been proven to give an estimator (TSRV_t) of the quadratic variation of log-return process unbiased and robust to the presence of microstructure noise.

It is known that the RV series consistently accounts for the sum of the continuous and discontinuous components of the log-return variation. For our purposes, we need to disentangle the two contributions. We choose as a proxy for the continuous variation the Threshold Bipower Variation (TBPV_t) defined in Corsi et al. (2010). In order to detect if on a given day there has been a jump, we employ the Threshold-z (Tz) test statistics (see equation 3.5 in Corsi et al. (2010) for a corrected version of the test statistics). It is proved that under the null hypothesis of no-jumps, the Tz statistics is distributed according as a standard normal random variable. Hence, for a given significance level α, we can assess the statistically significance of the daily jump component looking at the deviation of the Tz variable from the quantile

of the standard normal, $\Phi_{1-\alpha}$. If $\text{Tz}_t > \Phi_{1-\alpha}$, then we reject the null hypothesis. If there is statistical evidence for a jump, we attribute the difference $\text{TSRV}_t - \text{TBPV}_t$ to the jump component, i.e. $\text{JRV}_t = \mathcal{I}_{\text{Tz}_t > \Phi_{1-\alpha}} (\text{TSRV}_t - \text{TBPV}_t)$. Coherently, the continuous component is obtained by setting $\text{CRV}_t = \text{TSRV}_t$ for those days in which we do not reject the null, and $\text{CRV}_t = \text{TBPV}_t$ for the remaining days in which the test detects a significant jump. The CRV_t series is eventually cleaned removing the most extreme observations, seemingly due to volatility jumps, employing a threshold-based jump detection method as suggested by Corsi et al. (2013). Finally, since our estimator for the volatility is computed by using the returns during the trading period (from opening to closing of the market), we rescale it to match the unconditional mean of squared daily returns (close-to-close), including the contribution coming from overnight returns.

The Threshold Bipower Variation allows to identify the days on which at least one jump occurs, but does not give information about how many intraday jumps have actually happened. In order to identify these events we follow the approach of Andersen et al. (2010), considering the series of intraday 5-min returns. Once we find a day with at least one jump, we remove the largest 5-min return from the daily sample, and substitute it with the average return for that day. Then, we repeat the Tz test for the adjusted intraday series. If the test does not reject the null, we conclude that only one jump has occurred. If the test rejects the null, the procedure is repeated. At every run, we identify a new intraday jump. Finally, when the null hypothesis is not rejected anymore, we are left with the series of intraday 5-min jump returns. By means of this procedure, the time series of number of jumps per day, n_t, and of jump sizes, $X_{t,i}$, are recovered in a non parametric way. For a given day, if the observable quantity $\sum_{i=1}^{n_t} |X_{t,i}|^2$ does not coincide with $\mathcal{I}_{\text{Tz}_t > \Phi_{1-\alpha}} (\text{TSRV}_t - \text{TBPV}_t)$, then $X_{t,i}$ is scaled to ensure the matching. Given the number and size of intra-day jumps, their contribution to the total daily return is readily obtained. The daily jump-adjusted return series can then be computed by simply subtracting $\sum_{i=1}^{n_t} X_{t,i}$ from the daily returns. A clear advantage of the non parametric procedure by Andersen et al. (2010) is that both n_t and $X_{i,t}$ are made observable

and can be directly used in the estimation of the model parameters via maximum likelihood. As it will be explained in the next section, this is especially useful in the construction of the likelihood function of JRV_t and for the filtering of the latent intensity process ω_t.

In Figure 1 we present the time series of log-returns of the S&P500 Futures Index from July 3 1990 to June 28 2011 with the realized variance computed by the Two-Scale method by Zhang et al. (2005). On the same figure we also report the decomposition of the realized variance into the continuous and jump components. On Figure 2 we plot the auto-correlation function computed from the data divided in two sub-periods, one from July 3 1990 to June 28 2007 and the second from July 2 2007 to June 28 2011. The first sample ends before the spreading of the news which has later led to the sub-prime crisis, and indeed the sample mean of the Realized Variance increases by a factor three from the first to the second period. Consistently with well-established stylized facts, the auto-correlation of returns is not significant, whereas the auto-correlations of realized variances are statistically significant for all considered lag orders (from 1 to 60).

3.2 Maximum Likelihood Estimation

After the separation of the continuous and the discontinuous components of the dynamics, we estimate the parameters under the physical measure of the LHARG-ARJ process using the Maximum Likelihood Estimator. According to the model, the log-likelihood function is given by the sum of one term related to the daily log-return process and two terms related to the continuous and discontinuous components of the Realized Variance process. As can be seen from equation (2.3) the jump-adjusted log-return $\tilde{y}_t = y_t - \sum_{i=1}^{n_t} X_{t,i}$ is distributed as $\mathcal{N}\left(r + \left(\lambda_c - \frac{1}{2}\right)\text{CRV}_t + (\lambda_j - \eta)\left(\Lambda^2 + \delta^2\right) n_t, \text{CRV}_t\right)$ conditionally to CRV_t and n_t. Its contribution to the log-likelihood is expressed by the following quantity

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

As concerns the continuous component of the Realized Variance, it is modelled as a random variable sampled from the non-central gamma distribution $\bar{\gamma}(\kappa, \Theta(\mathbf{CRV}_t, \mathbf{L}_t), \theta)$. The corresponding log-likelihood has the form

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

The last log-likelihood term takes into account the jump component of the Realized Variance. As defined in equation (2.10), conditionally to the number of observed intra-day jumps n_t, the jump variation is distributed as a non-central chi-square. Since the variable n_t is Poisson distributed, the log-likelihood is given by the following expression

$$L^{\mathrm{JRV}}\left(\bar{\omega}, \xi, \zeta, \Lambda, \delta\right) = \sum_{t=1}^{T}\log\left(e^{-\omega_t}\frac{\omega_t^{n_t}}{n_t!}\sum_{m=0}^{\infty}e^{-\frac{n_t\Lambda^2}{2\delta^2}}\frac{\left(\frac{n_t\Lambda^2}{2\delta^2}\right)^m}{m!}\frac{\left(\sum_{i=1}^{n_t}|X_{t,i}|^2\right)^{\frac{n_t+2m}{2}-1}e^{-\frac{1}{2\delta^2}\sum_{i=1}^{n_t}|X_{t,i}|^2}}{\Gamma\left(\frac{n_t+2m}{2}\right)\left(2\delta^2\right)^{\frac{n_t+2m}{2}}}\right). \tag{3.1}$$

It is important to notice that ω_t enters the expression (3.1) as a latent process. So the optimization of the likelihood function relies on the filtering of ω_t. Since ω_t obeys equation (2.7), at each step of the optimization the latent intensity is filtered in a recursive way as a function of the observed n_t and intensity parameters $\bar{\omega}$, ξ, and ζ.
Finally, the estimation of the parameters characterizing the LHARG-ARJ process is performed via maximization of the whole log-likelihood function $L = L^y + L^{\mathrm{CRV}} + L^{\mathrm{JRV}}$.

In order to slightly reduce the dimensionality of the parameter space, we restrict two of them by means of variance targeting. In this way we force the exact matching of the observed sample mean of the Realized Variance continuous and jump components. As target parameters, we consider κ and $\bar{\omega}$ and compute them using the following expressions for the unconditional

mean of CRV_t and JRV_t, respectively

$$\mathbb{E}\left[\mathrm{CRV}_t\right] = \frac{\theta\kappa}{1-\theta\left(\sum_{i=d,w,m}\beta_i\right)},$$
$$\mathbb{E}\left[\mathrm{JRV}_t\right] = \frac{\bar{\omega}}{1-(\xi+\zeta)}\left(\Lambda^2+\delta^2\right).$$

In Table 1 we present the estimated parameters of the LHARG-ARJ model. Since λ_j is not significantly different from zero neither for the first period nor for the second, according to equation (2.9) we have that log-returns are negatively correlated with the future jump intensity. Moreover, both covariances in (2.11) are negative. Then, for the LHARG-ARJ model the leverage effect is not only due to the continuous component of returns through the mechanism introduced by Heston and Nandi (2000), but it is also determined by the impact of the discontinuous part of the price process.

3.3 Benchmark models

To assess the performance of the LHARG-ARJ model, we use the LHARG model by Majewski et al. (2015) and the BPJVM model introduced by Christoffersen et al. (2015) as benchmarks. These represent the state-of-the-art for the class of models based on realized measure of volatilities. The former accounts only for the diffusive component of asset price dynamics. The latter is based on an approach incorporating a GARCH structure for the latent volatility and jump intensity where bipower and jump variations play the role of idiosyncratic components. Indeed, the BPJVM can be seen as an extension of the Realized GARCH model by Hansen et al. (2012); Huang et al. (2017) which includes jumps and provides a closed-form exact expression for option prices. On one side, the comparison with the LHARG model allows to evaluate the impact of the inclusion of a jump component to the heterogeneous structure of volatility and leverage of the gamma models. On the other side, benchmarking with the BPJVM, we consider a competitor model of comparable complexity, that, however, differs from ours in several

respects – a different construction of the observable quantities, their role within the dynamics, the impact of conditional variance filtering, and the estimation/calibration procedure.

For the discussion of the LHARG model, we refer to Majewski et al. (2015). We provide here a short review of the model by Christoffersen et al. (2015). The dynamics of log-returns in BPJVM is as follows

$$y_t = r + \left(\lambda_c - \frac{1}{2}\right) h_{z,t} + (\lambda_j - \eta) h_{y,t} + \sqrt{h_{z,t}}\epsilon_{1,t} + \sum_{i=0}^{n_t} X_{t,i}, \tag{3.2}$$

where $\epsilon_{1,t} \sim \mathcal{N}(0,1)$ are i.i.d innovations and jumps' sizes $X_{t,i}$ are pairwise independent with normal distribution $\mathcal{N}(\theta, \delta^2)$. The number of jumps n_t has Poisson distribution with intensity $h_{y,t-1}$. To obtain a relation similar to (2.8), η is set to $\exp(\theta + 0.5\delta^2) - 1$. The value of factor $h_{z,t}$ corresponds to the expected value of Realized Bipower Variation on the following day, while the value of factor $h_{y,t}$ corresponds to the expected value of realized jump variation on the following day multiplied by $\theta^2 + \sigma^2$. Both realized variations are observable. The dynamics of $h_{z,t}$ and $h_{y,t}$ is given by

$$\begin{aligned} h_{z,t+1} &= \omega_z - a_z\sigma + \left(b_z + a_z - a_z\sigma\gamma^2\right) h_{z,t} + a_z\sigma\left(\epsilon_{2,t+1} - \gamma\sqrt{h_{z,t}}\right)^2, \\ h_{y,t+1} &= \omega_y + a_y h_{y,t} + b_y \sum_{i=0}^{n_{t+1}} |X_{t+1,i}|^2, \end{aligned}$$

where $\epsilon_{2,t} \sim \mathcal{N}(0,1)$ are i.i.d. and have correlation ρ with the diffusive return shock, $\epsilon_{1,t}$ defined in equation (3.2).

For the reader's convenience, Table 2 summarizes the main properties of the LHARG-ARJ, LHARG, and BPJVM models.

As a major difference with the procedure described by Christoffersen and co-authors, we consider the estimation problem – based on QMLE, as in the original paper – separately from calibration. First, we perform parameter estimation from historical time series. Then, risk premia are calibrated minimizing the pricing errors computed by means of the analytic mapping

of the parameters from the objective to the risk-neutral dynamics. It is well-known that the joint estimation-calibration achieves a higher statistical efficiency. Indeed, it exploits in a single step all information from daily and tick-by-tick return time series and option quotes. However, we believe that keeping estimation separate from calibration represents an alternative valuable approach. It allows to disentangle the different role played by the historical measure from the forward-looking information implied by option quotes and to better understand the impact of risk premia in shaping the pricing kernel. Tables 3 and 4 present the parameter values estimated for the LHARG and BPJVM models for the 1990–2007 and 2007-2011 periods.

In Figure 3 we compare the persistence of the Realized Volatility components CRV_t and JRV_t and the latent intensity ω_t of the LHARG-ARJ model with that for the RBV_t and RJV_t time series and the filtered $h_{z,t}$ and $h_{y,t}$ processes from Christoffersen et al. (2015), before and during the financial crisis. We note that the auto-correlation of the observable continuous component is very high and significant for both models and periods, and that during the period before 2007 it is slightly more persistent for the LHARG-ARJ. However, in Christoffersen et al. (2015) the role of the conditional variance of the continuous component of returns is played by the filtered process $h_{z,t}$ while RBV_t enters the observation equation. So, the persistence of the CRV_t has to be compared with the persistence of $h_{z,t}$. The latter is extremely high and close to the unit value. We will comment more on the consequences of this level of persistence in the section dedicated to the option pricing exercise. Finally, we observe that the auto-correlation of the observable jump component of the Realized Variance is statistically significant and persistent for both models and periods. Comparing the jump intensity processes, we conclude that they are both strongly persistent, with a slightly larger value for the persistence of ω_t during the crisis period.

To check our modeling assumptions, we perform some model miss-specification tests. Comments refer to the entire period spanned by the available data, i.e. from 3 July, 1990 to 28 June, 2011. We compare the mean, variance, skewness and kurtosis of the idiosyncratic components ϵ_t in the LHARG-ARJ model, with that of the benchmark model LHARG. As far as the former

model is concerned, mean equals 0.09, variance 1.11, skewness 0.33, and excess kurtosis 0.81. For LHARG we estimate 0.12, 1.11, 0.40, and 1.07, respectively. In both cases the Jarque-Bera test rejects normality at 1% significance level. Thus, some mild inadequacies of both models survive. Nonetheless, when moving to the ARJ version of the LHARG class, there is a sizable decrease of skewness and excess kurtosis. To test for the presence of residual serial dependence in the idiosyncratic component, Figure 4 shows the autocorrelation of different powers of ϵ_t for lags ranging from zero to fifty. Moving clockwise, panels report the auto-correlation of ϵ_t, ϵ_t^2, ϵ_t^3, and ϵ_t^4, respectively, and the 95% confidence band. As it clearly emerges, there is no evidence of serial dependence. Finally, Figure 5 investigates the ability of the LHARG and LHARG-ARJ models to forecast one-day-ahead realized variance. The model predicted variance – CRV_t for LHARG and $CRV_t + JRV_t$ for LHARG-ARJ – is on the horizonal axis. The ex-post realized variance is on the vertical axis. The regression R^2 is 40% for LHARG and 41% for LHARG-ARJ, with coefficients equal to 1.58 and 1.56, respectively. The two models performs in a comparable way, with a slightly better performance of the LHARG-ARJ model.

4 Option pricing

Our dataset consists of Plain Vanilla 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$ (see Barone-Adesi et al. (2008), Corsi et al. (2013) and Majewski et al. (2015)). Defining the moneyness as K/S_t, we do not consider options with moneyness larger than 1.2 and with moneyness smaller than 0.8. We term options as deep out-of-the-money if the moneyness is between $0.8 \leq m \leq 0.9$ or $1.1 < m \leq 1.2$, as out-of-the-money if $0.9 < m \leq 0.98$ or $1.02 < m \leq 1.1$, and as at-the-money 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).

4.1 Model calibration and pricing method

In order to calibrate the model under the risk-neutral measure, the values for risk premium parameters $(\mu_c, \mu_j, \nu_c, \nu_j)$ need to be specified. As a consequence of Proposition 2, μ_c and μ_j are fixed by the no-arbitrage condition, while ν_c and ν_j remain undetermined parameters to be calibrated on the option data.

For the calibration procedure, we adopt a method 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 previously described – obtaining a 5×4 moneyness-maturity grid obtaining a 20-point discrete representation of the implied volatility surface. For each subset, we compute the unconditional average of the market implied volatilities. Then, we calculate the corresponding model implied volatility and obtain the optimal values of (ν_c, ν_j) as

$$\underset{(\nu_c,\nu_j)}{\arg\min} \left\{ f_{\text{obj}}(\nu_c, \nu_j) \right\} .$$

The objective function $f_{\text{obj}}(\nu_c, \nu_j)$ reads

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

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 adopt the COS numerical approach by Fang and Oosterlee (2008). This method, based on Fourier-cosine expansions, effectively evaluates the price of Plain Vanilla options from the characteristic function of log-returns.

To summarize the whole numerical procedure, we proceed in four steps. First, we estimate the model parameters under the physical measure via MLE and obtain the values given in Table 1 for LHARG-ARJ model. Second, the premia ν_c and ν_j appearing in the pricing kernel are

calibrated on the option dataset. Since neither λ_c nor λ_j are statistically significant, calibration is performed by fixing their values to zero. The results are reported in Table 1, too. In the third step, we switch from the physical to the martingale measure using Proposition 3. Finally, we compute option prices for each Wednesday in our dataset employing the COS method. Due to significant differences in the statistics of the time series of Realized Volatilities arising across the financial crisis which started in the summer of 2007 and reached its peak in the fall of 2008, we divide our option dataset in two periods too: A pre-crisis period going from July, 1990 to June, 2007 and a crisis/post-crisis period from July, 2007 to June, 2011. We repeat the procedure described above for the two periods, separately.

Concerning the BPJVM model, risk-neutralization is achieved following two distinct approaches. In a first case, we apply a three-dimensional Esscher transform with (χ, ν_3) as free parameters to be calibrated on option data. This is the approach adopted in Christoffersen et al. (2015) and we refer to the paper for further details. A second different strategy, ensuring a fair comparison between the BPJVM and our model, employs the pricing kernel (2.13). In this case, the SDF depends on four risk premia but the no-arbitrage constraints, detailed in appendix D, fix two of them. As a result, only (ν_c, ν_j) survive as independent risk premia to be calibrated on market prices.

The risk premia panels in Tables 3 and 4 report the values of premia calibrated on options for the periods 1996–2007 and 2007-2011. As for the LHARG-ARJ model, estimated values for the coefficients λ_c, λ_z, and λ_y are not statistically significant. Then, calibration of the premia is performed by fixing the value of all λ's equal to zero.

4.2 Pricing performances

The pricing performance is evaluated with the percentage Implied Volatility Root Mean Square Error ($RMSE_{IV}$) introduced by Renault (1997) and computed as

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

where N is the number of options, IV^{mkt} and IV^{mod} represent the market and model implied volatilities, respectively. Table 5 shows that LHARG-ARJ improves the pricing performance with respect to LHARG and BPJVM for both region of moneyness, before and after financial crisis of 2008. The improvement ranges from $8\% - 9\%$ of relative $RMSE$ with respect to LHARG for the pre-crisis period, while it reduces to 2% for the crisis/post-crisis period. This difference suggests that in the 2007-2008 period of financial distress option prices are more influenced by the high level of the diffusive component – a so-called volatility burst – than by the discontinuous component of price dynamics. From a refined comparison, we observe that for options belonging to the per-crisis period, the best performance of the LHARG-ARJ model concentrates on pricing contracts with maturity less than 50 days. The RMSE is up to 15% smaller than the error obtained for the LHARG (see Table 6). These results confirm the well established fact that the inclusion of a jump component is essential for the correct description of the volatility surface implied by short-term options. Moving towards longer maturities, in the ATM region performances are in favor of the LHARG dynamics, with the exception of the longest maturities. When pricing deep-out-of-the-money and out-of-the-money options, on average the LHARG-ARJ performs better. Again, these results conform to the intuition that extreme regions of the moneyness, especially deep-out-of-the-money put, are more sensitive to extreme price events. The table referring to the crisis/post-crisis period (Table 7) shows results in line with the previous period. In this case, the performances are more balanced between the two models. Again, this is a plausible consequence of the unprecedented high level of volatility. This effect, evident from Figure 1 (left panel on the second row), tends to decrease the modeling

advantage of the ARJ model.

As far as the comparison between the LHARG-ARJ and the BPJVM model is concerned, we observe that the former globally performs better than the latter in both periods (see Table 5, left column numbers). The best performance is for contracts with moneyness between 0.9 and 1.1 (around 24%) in the pre-crisis period. LHARG-ARJ implied volatilities are more accurate by a factor of 0.81 considering a wider range of moneyness ($0.8 < m < 1.2$), and the gain in relative error is of order 16% for the 2007 – 2011 period. The result for the central region of the volatility surface confirms that the heterogeneous structure of the LHARG-ARJ is a parsimonious and effective way to provide a good description of the ATM implied volatility dynamics. These results are rooted on the pricing kernel employed by Christoffersen et al. (2015). To ensure a fair comparison with the BPJVM and to decipher whether the best performance derives from a better specification of the historical dynamics or has to be attributed to the different pricing kernel, the right column in Table 5 (rows referring to the BPJVM model) reports the result from a second comparison. In this case, we employ the pricing kernel (2.13). It can be readily recognized that performances of the BPJVM improve. For the pre-crisis period, the relative error diminished from 24% to 9% ATM, and from 19% to 6% when including deep-out-of-the-money options. The improvement is also evident for the crisis/post-crisis period, although in this case the relative gain is smaller. Nonetheless, in both cases the global performance is still clearly in favor of the LHARG-ARJ model.

By commenting the details of Tables 8 and 9, we only refer to Panels C and D. These correspond to pricing results based on the SDF (2.13). The qualitative picture agrees with that of Panels A and B, but numerical results decrease the gap between the LHARG-ARJ and BPJVM models. In the pre-crisis period, for short to moderate maturities the RMSE are balanced between the two models. They more evidently favorite the LHARG-ARJ model when moving to longer horizons. As clear from Tables 1 and 4, the persistence of the latent intensity is comparable for both LHARG-ARJ and BPJVM models. It reads 0.994 for the LHARG-ARJ and 0.986 for the BPJVM in the 1990 – 2007 period. For the 2007 – 2011 period, it rises to

0.999 for the LHARG-ARJ and 0.998 for the BPJVM. Then, such mild difference cannot be responsible for the sizable difference in the pricing performance between the two models in the long horizon. Moreover, the contribution of the jump variation to the total quadratic variation, measured as the ratio between the historical unconditional mean of JRV and the average sum of JRV and CRV, amounts to 18% in the 1990 – 2007 period and to 9% in the 2007 – 2011 period. The unconditional level of CRV increases by nearly a factor three – passing from 6.62e-05 to 1.91e-04, while the unconditional level of JRV increases from 1.21e-05 to 1.74e-05. Then, the difference in performance can reasonably be attributed to the substantial difference in persistence of the continuous component of the conditional volatility between the two models. While for the LHARG-ARJ model it ranges between 0.812 – 0.831, it saturates to 0.999 for the BPJVM model. Even though the latter model tracks adequately well the short term of the implied volatility, it systematically over-prices the long term of the volatility surface. The LHARG-ARJ seems to be more flexible than the BPJVM in reproducing the term structure of implied volatilities, especially at longer maturities. Moving to the crisis/post-crisis period, the previous features are confirmed by the relative errors. However, the performance of the BPJVM deteriorates in the short maturity region and for all deep-out-of-the-money options, irrespectively of the maturity.

5 Conclusions

In this paper, we present a heterogeneous autoregressive model for the asset log-return dynamics with observable volatility and jump component with dynamic intensity. We devise our proposal with the purpose to describe the empirical properties of financial returns and volatilities, as long memory, leverage effect, jumps, and jump clustering. We present the analytical characterisation of the log-return moment generating function under both the physical and the risk-neutral measure. For the change of measure, we adopt a flexible exponential affine pricing kernel. It is designed to differentiate among directional and non-directional sources of financial risk, and to separately compensate for them, introducing separate premia for the continuous and

discontinuous equity risk, and the continuous and discontinuous variance components. Finally, we detail an application to option pricing, and show the improvements of the novel approach. The model is able to reproduce successfully the different features of the implied volatility surface and to provide better performances when compared with state-of-the-art discrete time pricing models based on Realized Volatility measures.

References

Aït-Sahalia, Y., 2004. Disentangling diffusion from jumps. Journal of Financial Economics 74 (3), 487–528.

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., 2009. Testing for jumps in a discretely observed process. The Annals of Statistics 37 (1), 184–222.

Andersen, T. G., Bollerslev, T., Diebold, F., 2007. 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., Frederiksen, P., Ørregaard Nielsen, M., 2010. Continuous-time models, realized volatilities, and testable distributional implications for daily stock returns. Journal of Applied Econometrics 25 (2), 233–261.

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

Azizpour, S., Giesecke, K., Schwenkler, G., 2018. Exploring the sources of default clustering. Journal of Financial Economics 129 (1), 154–183.

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., 2006. Econometrics of testing for jumps in financial economics using bipower variation. Journal of Financial Econometrics 4 (1), 1–30.

Barndorff-Nielsen, O. E., Shephard, N., 2007. Variation, Jumps, and High-Frequency Data in Financial Econometrics. In: Blundell, R., Newey, W., Persson, T. (Eds.), Advances in Economics and Econometrics. Vol. 3. Cambridge University Press, pp. 328–372.

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

Bollerslev, T., Kretschmer, U., Pigorsch, C., Tauchen, G., 2009. A discrete-time model for daily S & P500 returns and realized variations: Jumps and leverage effects. Journal of Econometrics 150 (2), 151–166.

Bollerslev, T., Law, T. H., Tauchen, G., 2008. Risk, jumps, and diversification. Journal of Econometrics 144 (1), 234–256.

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

Calcagnile, L. M., Bormetti, G., Treccani, M., Marmi, S., Lillo, F., 2018. Collective synchronization and high frequency systemic instabilities in financial markets. Quantitative Finance 18 (2), 237–247.

Chen, K., Poon, S.-H., 2013. Variance swap premium under stochastic volatility and self-exciting jumps. Available at ssrn.com/abstract=2200172.

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., Elkamhi, R., Feunou, B., Jacobs, K., 2009. Option valuation with conditional heteroskedasticity and nonnormality. The Review of Financial Studies 23 (5), 2139–2183.

Christoffersen, P., Fenou, B., Jeon, Y., 2015. Option valuation with observable volatility and jump dynamics. Journal of Banking & Finance 61 (S2), S101–S120.

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., 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.

Cont, R., Tankov, P., 2004. Financial Modelling with jump processes. Chapman & Hall/CRC.

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.

Daley, D. J., Vere-Jones, D., 2003. An Introduction to the Theory of Point Processes Volume I: Elementary Theory and Methods. Springer, Heidelberg.

Ding, X., Giesecke, K., Tomecek, P. I., 2009. Time-changed birth processes and multiname credit derivatives. Operations Research 57 (4), 990–1005.

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.

Fičura, M., 2015. Modelling jump clustering in the four major foreign exchange rates using high-frequency returns and cross-exciting jump processes. Procedia Economics and Finance 25, 208 – 219.

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

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

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

Hansen, P. R., Huang, Z., Shek, H. H., 2012. Realized garch: a joint model for returns and realized measures of volatility. Journal of Applied Econometrics 27 (6), 877–906.

Hawkes, A. G., 1971. Spectra of some self-exciting and mutually exciting point processes. Biometrika 58, 83–90.

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

Huang, Z., Wang, T., Hansen, P. R., 2017. Option pricing with the realized garch model: An analytical approximation approach. Journal of Futures Markets 37 (4), 328–358.

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.

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.

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

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

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.

A Computation of the Moment Generating Function

As starting point, we recall that a LHARG process with zero mean leverage can be mapped in a LHARG process with parabolic leverage as shown in Majewski et al. (2015). We rewrite $\Theta(\mathbf{CRV}_t, \mathbf{L}_t)$ in (2.5), as

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

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 LHARG with zero mean leverage in (2.5), and $\ell_t = \left(\epsilon_t - \gamma\sqrt{\mathrm{CRV}_t}\right)^2$ is the parabolic leverage. We further define

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

Then we compute the one-step-forward MGF by applying the 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 $Z\sim\mathcal{N}(0,1)$,

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

we obtain

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

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

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

and

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

we can rewrite (A.1) 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, the previous expression factorizes as

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

To shorten the notation, in the remaining part of the paper we write x for $x(z,b,c)$ and v for $v(z)$. Using equations (8)-(9) from Gourieroux and Jasiak (2006) we obtain

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

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

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

Collecting all previous results, an exponentially affine form for the physical MGF follows

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

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

The computation of the MGF for the log-return $y_{t,T} = \sum_{i=t}^{T} y_i$ between t and T involves

the iterated 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 the coefficients are defined as follows

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

Moving a time step backward, we isolate all the random variables at time $T-1$ and compute the expectation conditioning to information up to $T-2$,

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

Iterating the reasoning, the final expression of the conditional moment generating function reads

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

where the coefficients satisfy the following recursive relations

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

with

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

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

B No-arbitrage condition

The no-arbitrage constraint

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

for transform (2.13) reads

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

Firstly, we compute the denominator of the expression above

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

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

$$=\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{r+\left(\frac{\nu_2^2}{2}-\nu_c\right)\mathrm{CRV}_t}\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{-\nu_j\sum_{i=1}^{n_t}X_{t,i}^2-\mu_j\sum_{i=1}^{n_t}X_{t,i}}\middle|n_t\right]\middle|\mathcal{F}_{t-1}\right]$$

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

where

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

Similar computations for the numerator give

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

$$=\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{-\nu_c \mathrm{CRV}_t-\nu_j \mathrm{JRV}_t-\mu_c\sqrt{\mathrm{CRV}_t}\epsilon_t-\mu_j\sum_{i=1}^{n_t}X_{t,i}+r+\left(\lambda_c-\frac{1}{2}\right)\mathrm{CRV}_t+(\lambda_j-\eta)\left(\Lambda^2+\delta^2\right)n_t+\sqrt{\mathrm{CRV}_t}\epsilon_t+\sum_{i=1}^{n_t}X_{t,i}}\middle|\mathcal{F}_{t-1}\right]$$

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

where Putting together numerator and denominator, the no-arbitrage condition (B.1) becomes

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

In order to be satisfied, the above equation implies the 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

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

C Mapping of parameters from $\mathbb{P}-$ to $\mathbb{Q}-$measure

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

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

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

Proceeding as in Appendix A we obtain

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

where $\bar{x}$, $\bar{v}$ are given in (B.2) and

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

A comparison of (C.1) with the equivalent expression under physical measure $\mathbb{P}$ yields the following relations

$$\begin{aligned}
\kappa\left(\mathcal{W}\left(\hat{x},\theta\right)-\mathcal{W}\left(\bar{x},\theta\right)\right) &= \kappa^*\mathcal{W}\left(x^*,\theta^*\right)\\
d\left(\mathcal{V}\left(\hat{x},\theta\right)-\mathcal{V}\left(\bar{x},\theta\right)\right) &= d^*\mathcal{V}\left(x^*,\theta^*\right)\\
\alpha_j\left(\mathcal{V}\left(\hat{x},\theta\right)-\mathcal{V}\left(\bar{x},\theta\right)\right) &= \alpha_j^*\mathcal{V}\left(x^*,\theta^*\right)\\
\beta_i\left(\mathcal{V}\left(\hat{x},\theta\right)-\mathcal{V}\left(\bar{x},\theta\right)\right) &= \beta_i^*\mathcal{V}\left(x^*,\theta^*\right)\\
\omega_t e^{\bar{v}} &= \omega_t^*\\
\hat{v}-\bar{v} &= v^*,
\end{aligned} \tag{C.2}$$

where

$$\begin{aligned}
x^* &= z\left(\lambda^* - \frac{1}{2}\right) + b^* + \frac{\frac{z^2}{2} + (\gamma^*)^2 c^* - 2c^*\gamma^* z}{1-2c^*}\\
v^* &= \frac{z^2(\delta^*)^2}{2} + z\Lambda^* + z\left(\lambda_j^*-\eta^*\right)\left((\Lambda^*)^2+(\delta^*)^2\right).
\end{aligned} \tag{C.3}$$

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

The first four relations in (C.2) together with no-arbitrage condition imply the following

mapping of the non-centered gamma process parameters

$$\begin{aligned}
&\beta_d^* = \tfrac{1}{1-\theta \mathrm{y}^*}\beta_d\,, \quad \beta_w^* = \tfrac{1}{1-\theta \mathrm{y}^*}\beta_w\,, \quad \beta_m^* = \tfrac{1}{1-\theta \mathrm{y}^*}\beta_m\,,\\
&\alpha_d^* = \tfrac{1}{1-\theta \mathrm{y}^*}\alpha_d\,, \quad \alpha_w^* = \tfrac{1}{1-\theta \mathrm{y}^*}\alpha_w\,, \quad \alpha_m^* = \tfrac{1}{1-\theta \mathrm{y}^*}\alpha_m\,,\\
&\theta^* = \tfrac{1}{1-\theta \mathrm{y}^*}\theta\,, \quad \delta^* = \delta\,, \quad \gamma^* = \gamma + \lambda_c\,,\\
&d^* = \tfrac{1}{1-\theta \mathrm{y}^*}d\,, \quad \lambda_c^* = 0,
\end{aligned} \tag{C.4}$$

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

The last two relations in (C.2) are related to the mapping of the jump component. However, to get the complete characterisation of jumps under the risk-neutral measure we need to derive the MGF of sum of jumps over two days

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

where v^* is given in (C.3). Assuming that the dynamics of ω^* under $\mathbb{Q}$ is given by equation (2.7), we can write

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

Finally, recalling that n_{t+1}^* is Poisson and $X_{t+1,j}^*$ are i.i.d. normally distributed random vari-

ables, we obtain

$$\begin{aligned}&\mathbb{E}^{\mathbb{Q}}\left[\mathrm{e}^{z\left(\lambda_j^*-\eta^*\right)\left((\Lambda^*)^2+(\delta^*)^2\right)(n_{t+1}+n_{t+2})+z\sum_{i=t+1}^{t+2}\sum_{k=1}^{n_i^*}X_{i,k}^*}\middle|\mathcal{F}_t\right]\\&=\exp\left(\left(\bar{\omega}^*+\xi^*\omega_{t+1}^*\right)\left(\mathrm{e}^{v^*}-1\right)+\omega_{t+1}^*\left(\mathrm{e}^{u^*}-1\right)\right),\end{aligned}\tag{C.5}$$

where

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

We compare (C.5) with

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

where

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

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

Plugging in the dynamics of the intensity ω provided by equation (2.7), we get

$$\mathbb{E}^{\mathbb{P}}\left[\prod_{i=t}^{t+1} M_{i,i+1}\mathrm{e}^{z(\lambda_j-\eta)\left(\Lambda^2+\delta^2\right)n_{i+1}+z\sum_{k=1}^{n_{i+1}}X_{i+1,k}}|\mathcal{F}_t\right]=\exp\left(\mathrm{e}^{\bar{v}}\left(\bar{\omega}+\xi\omega_{t+1}\right)\left(\mathrm{e}^{v-\bar{v}}-1\right)+\omega_{t+1}\mathrm{e}^{\bar{v}}\left(\mathrm{e}^{u-\bar{v}}-1\right)\right) \tag{C.6}$$

where

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

Comparing (C.5) with (C.6) 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 jump intensity

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

and the following mapping of the parameters

$$\begin{aligned}\Lambda^*&=\Lambda-\left(\mu_j+2\nu_j\Lambda\right)\frac{\delta^2}{(1+2\nu_j\delta^2)},\\ (\delta^*)^2&=\frac{\delta^2}{(1+2\nu_j\delta^2)},\\ \lambda^*_j&=\frac{\left(\lambda_j-\eta\right)\left(\Lambda^2+\delta^2\right)}{\left(\Lambda^*\right)^2+\left(\delta^*\right)^2}+\eta^*,\\ \bar{\omega}^*&=e^{\bar{v}}\bar{\omega},\\ \xi^*&=\xi,\\ \zeta^*&=e^{\bar{v}}\zeta,\end{aligned} \tag{C.7}$$

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

D BPJVM risk-neutralization with a four-dimensional SDF

D.1 No-arbitrage condition

We detail here the formal derivation of the no-arbitrage condition for the BPJVM model, when the SDF (2.13) is employed. In order to ensure absence of arbitrage, the following condition has to be satisfied

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

First, we compute the denominator of the SDF (2.13)

$$\mathbb{E}^{\mathbb{P}}\left[e^{-\nu_c \mathrm{RBV}_t - \nu_j \mathrm{RJV}_t - \mu_c\sqrt{h_{z,t-1}}\epsilon_{1,t} - \mu_j \sum_{i=0}^{n_t} X_{t,i}}|\mathcal{F}_{t-1}\right],$$

by separating the continuous contribution given by

$$\mathbb{E}^{\mathbb{P}}\left[e^{-\nu_c\sigma\left(\epsilon_{2,t}-\gamma\sqrt{h_{z,t-1}}\right)^2-\mu_c\sqrt{h_{z,t-1}}\epsilon_{1,t}}|\mathcal{F}_{t-1}\right] = e^{-\frac{1}{2}\log(1+2\nu_c\sigma)+\left(\frac{(\mu_c\rho-2\nu_c\sigma\gamma)^2}{2(1+2\nu_c\sigma)}+\frac{1}{2}\left(1-\rho^2\right)\mu_c^2-\nu_c\sigma\gamma^2\right)h_{z,t-1}}$$

from the jump contribution corresponding to

$$\mathbb{E}^{\mathbb{P}}\left[e^{-\nu_j\sum_{i=0}^{n_t}X_{t,i}^2-\mu_j\sum_{i=0}^{n_t}X_{t,i}}|\mathcal{F}_{t-1}\right] =$$

$$\exp\left(\left(e^{-\frac{1}{2}\log(1+2\nu_j\delta^2)-\mu_j\theta-\nu_j\theta^2+\frac{\left(\mu_j+2\nu_j\theta\right)^2\delta^2}{2\left(1+2\nu_j\delta^2\right)}}-1\right)h_{y,t-1}\right).$$

Collecting all terms, the following result holds

$$\mathbb{E}^{\mathbb{P}}\left[e^{-\nu_c \mathrm{RBV}_t - \nu_j \mathrm{RJV}_t - \mu_c\sqrt{h_{z,t-1}}\epsilon_{1,t} - \mu_j \sum_{i=0}^{n_t} X_{t,i}} | \mathcal{F}_t\right] =$$

$$\exp\left(\begin{array}{l} -\frac{1}{2}\log(1+2\nu_c\sigma) + \nu_c\sigma + \left(\frac{(\mu_c\rho - 2\nu_c\sigma\gamma)^2}{2(1+2\nu_c\sigma)} + \frac{1}{2}\left(1-\rho^2\right)\mu_c^2 - \nu_c\right) h_{z,t-1} \\ \\ + \left(e^{-\frac{1}{2}\log(1+2\nu_j\delta^2) - \mu_j\theta - \nu_j\theta^2 + \frac{\left(\mu_j + 2\nu_j\theta\right)^2\delta^2}{2\left(1+2\nu_j\delta^2\right)}} - 1\right) h_{y,t-1} \end{array}\right).$$

The numerator of the no-arbitrage condition can be expressed as

$$\mathbb{E}^{\mathbb{P}}\left[e^{-\nu_c \mathrm{RBV}_t - \nu_j \mathrm{RJV}_t + r + \left(\lambda_c - \frac{1}{2}\right)h_{z,t-1} + (\lambda_j - e^{\theta+\frac{1}{2}\delta^2}+1)h_{y,t-1} - (\mu_c - 1)\sqrt{h_{z,t-1}}\epsilon_{1,t} - (\mu_j - 1)\sum_{i=0}^{n_t} X_{t,i}} | \mathcal{F}_{t-1}\right],$$

which, neglecting for a moment the constant term $e^{r+\left(\lambda_c-\frac{1}{2}\right)h_{z,t-1}+(\lambda_j - e^{\theta+\frac{1}{2}\delta^2}+1)h_{y,t-1}}$, can be computed by simply substituting $\mu_c \to \mu_c - 1$ and $\mu_j \to \mu_j - 1$ in the expression for the denominator. After simple algebra, the no-arbitrage condition implies the following equality

$$\begin{aligned} &\left(\lambda_c - \mu_c\left(1 - \frac{2\nu_c\sigma\rho^2}{1+2\nu_c\sigma}\right) - \frac{\nu_c\sigma\rho\left(\rho - 2\gamma\right)}{1+2\nu_c\sigma}\right) h_{z,t-1} \\ &+ \left(\lambda_j - e^{\theta+\frac{1}{2}\delta^2} + 1 + e^{v(\mu_j - 1, \nu_j)} - e^{v(\mu_j,\nu_j)}\right) h_{y,t-1} = 0 \end{aligned}$$

where

$$v(\mu, \nu_j) = -\frac{1}{2}\log(1+2\nu_j\delta^2) - \mu_j\theta - \nu_j\theta^2 + \frac{\left(\mu_j + 2\nu_j\theta\right)^2\delta^2}{2\left(1+2\nu_j\delta^2\right)}.$$

This equation is equivalent to the two relations which follow

$$\begin{aligned} &\lambda_c - \mu_c\left(1 - \frac{2\nu_c\sigma\rho^2}{1+2\nu_c\sigma}\right) - \frac{\nu_c\sigma\rho\left(\rho - 2\gamma\right)}{1+2\nu_c\sigma} = 0 \\ &\lambda_j - e^{\theta+\frac{1}{2}\delta^2} + 1 + e^{v(\mu_j - 1, \nu_j)} - e^{v(\mu_j,\nu_j)} = 0. \end{aligned}$$

The two relations determine the equity risk premia μ_c and μ_j as functions of ν_c and ν_j. The latter two premia remain free parameters of the model to be calibrated on market option prices.

D.2 Risk-neutral MGF

For pricing purposes, we need to compute the MGF of the BPJVM

$$\begin{aligned}
&\mathbb{E}^{\mathbb{Q}}\left[e^{zy_{t+1}}|\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{P}}\left[M_{t,t+1}e^{zy_{t+1}}|\mathcal{F}_t\right] = \\
&= \mathbb{E}^{\mathbb{P}}\left[\frac{e^{-\nu_c \mathrm{RBV}_{t+1} - \nu_j \mathrm{RJV}_{t+1} - \mu_c\sqrt{h_{z,t}}\epsilon_{1,t+1} - \mu_j \sum_{i=0}^{n_{t+1}} X_{t+1,i} + zy_{t+1}}}{\mathbb{E}^{\mathbb{P}}\left[e^{-\nu_c \mathrm{RBV}_{t+1} - \nu_j \mathrm{RJV}_{t+1} - \mu_c\sqrt{h_{z,t}}\epsilon_{1,t+1} - \mu_j \sum_{i=0}^{n_{t+1}} X_{t+1,i}}|\mathcal{F}_t\right]}|\mathcal{F}_t\right] = \\
&= \exp\left(\begin{array}{l}\mathcal{C}\left(z,-\nu_c\right) - \mathcal{C}\left(0,-\nu_c\right) + \left(\mathcal{D}\left(z,-\mu_c,-\nu_c\right) - \mathcal{D}\left(0,-\mu_c,-\nu_c\right)\right)h_{z,t} \\ \\ + \left(\mathcal{F}\left(z,-\mu_j,-\nu_j\right) - \mathcal{F}\left(0,-\mu_j,-\nu_j\right)\right)h_{y,t}\end{array}\right) \\
&= \exp\left(c_t + d_t h_{z,t} + f_t h_{y,t}\right),
\end{aligned}$$

where the following quantities have been introduced

$$\begin{aligned}
\mathcal{C}\left(x,y\right) &= rx - \sigma y - \frac{1}{2}\log\left(1-2\sigma y\right) \\
\mathcal{D}\left(x,y,z\right) &= \left(\lambda_c - \frac{1}{2}\right)x + z + \frac{1}{2}\left(1-\rho^2\right)\left(x-y\right)^2 + \frac{\left((x-y)\rho - 2\sigma\gamma z\right)^2}{2(1-2\sigma z)} \\
\mathcal{F}\left(x,y,z\right) &= (\lambda_j - e^{\theta+\frac{1}{2}\delta^2} + 1)x + e^{-\frac{1}{2}\log(1-2z\delta^2) + (x-y)\theta + z\theta^2 + \frac{((x-y)+2z\theta)^2\delta^2}{2(1-2z\delta^2)}} - 1\,.
\end{aligned}$$

Computing one step backward, we obtain

$$\mathbb{E}^{\mathbb{P}}\left[M_{t-1,t}e^{zy_t+c_t+d_th_{z,t}+f_th_{y,t}}|\mathcal{F}_{t-1}\right]=$$

$$=\frac{1}{\mathbb{E}^{\mathbb{P}}\left[e^{-\nu_c\mathrm{RBV}_t-\nu_j\mathrm{RJV}_t-\mu y_t}|\mathcal{F}_{t-1}\right]}\mathbb{E}^{\mathbb{P}}\left[\begin{array}{l}e^{-\nu_c\mathrm{RBV}_t-\nu_j\mathrm{RJV}_t-\mu_c\sqrt{h_{z,t}}\epsilon_{1,t+1}-\mu_j\sum_{i=0}^{n_t}X_{t,i}+zy_t}\times\\ \\ e^{c_t+d_t(\omega_z+b_zh_{z,t-1}+a_z\mathrm{RBV}_t)+f_t(\omega_y+b_yh_{y,t-1}+a_z\mathrm{RVJ}_t)}|\mathcal{F}_{t-1}\end{array}\right]=$$

$$=\exp\left(\begin{array}{l}c_t+\omega_zd_t+\omega_yf_t+\mathcal{C}\left(z,a_zd_t-\nu_c\right)-\mathcal{C}\left(0,-\nu_c\right)\\ \\ +\left(d_tb_z+\mathcal{D}\left(z,-\mu_c,a_zd_t-\nu_c\right)-\mathcal{D}\left(0,-\mu_c,-\nu_c\right)\right)h_{z,t-1}\\ \\ +\left(f_tb_y+\mathcal{F}\left(z,-\mu_j,a_yf_t-\nu_j\right)-\mathcal{F}\left(0,-\mu_j,-\nu_j\right)\right)h_{y,t-1}\end{array}\right)$$

$$=\exp\left(c_{t-1}+d_{t-1}h_{z,t-1}+f_{t-1}h_{y,t-1}\right).$$

The last equality allows to define the following backward recursive formulas for the coefficients of the exponential affine risk-neutral MGF:

$$\begin{aligned}c_{t-1}&=c_t+\omega_zd_t+\omega_yf_t+\mathcal{C}\left(z,a_zd_t-\nu_c\right)-\mathcal{C}\left(0,-\nu_c\right)\\ d_{t-1}&=d_tb_z+\mathcal{D}\left(z,-\mu_c,a_zd_t-\nu_c\right)-\mathcal{D}\left(0,-\mu_c,-\nu_c\right)\\ f_{t-1}&=f_tb_y+\mathcal{F}\left(z,-\mu_j,a_yf_t-\nu_j\right)-\mathcal{F}\left(0,-\mu_j,-\nu_j\right).\end{aligned}$$

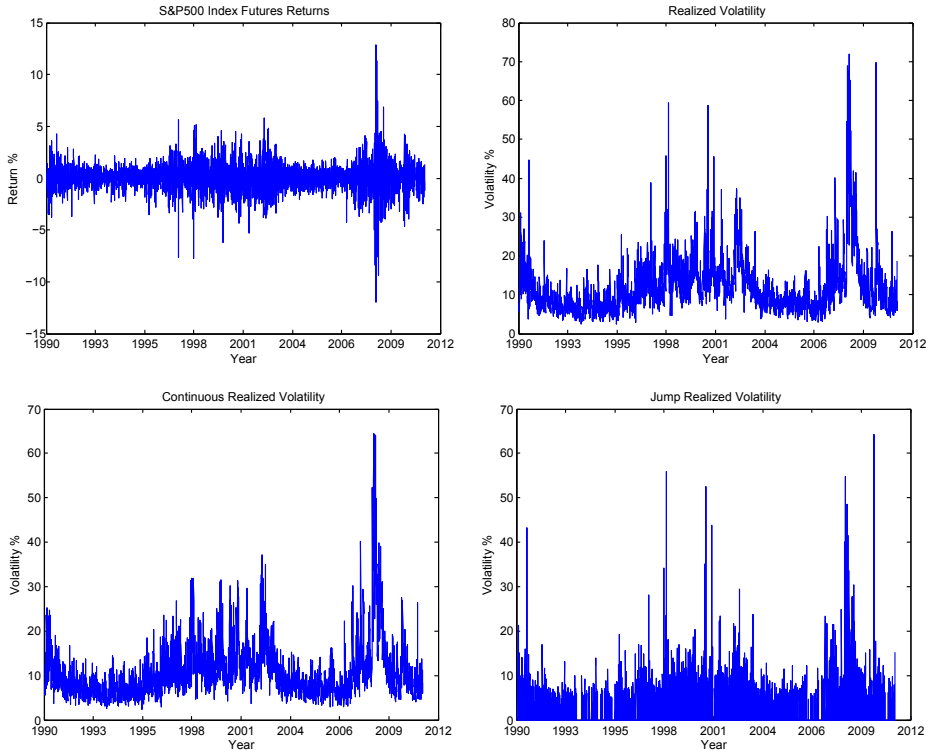

Figure 1: Daily returns and realized volatility time series for SP&500 Index Futures. Realized Volatility is computed using the Two-Scale method by Zhang et al. (2005). The continuous and the jump components are constructed following the procedure in Section 3. The sample starts on 3 July, 1990 and ends on 28 June, 2011.

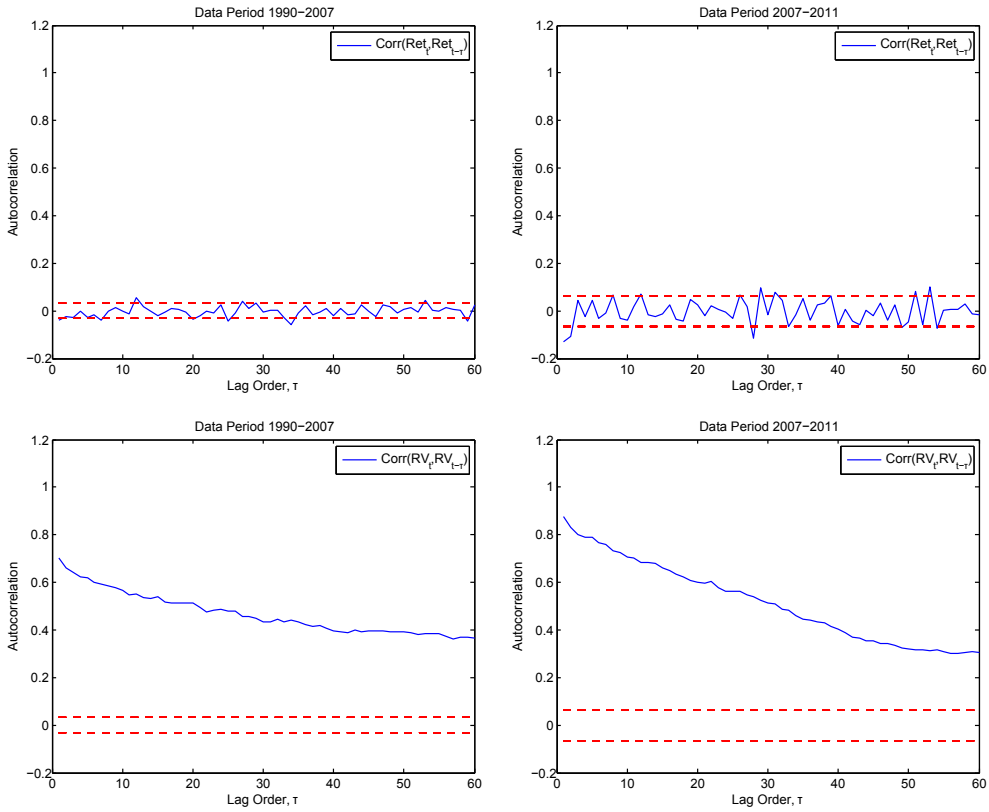

Figure 2: Auto-correlations of daily returns and Realized Volatilities computed from two periods. The first sample (left column) starts on 3 July, 1990 and ends on 28 June, 2007. The second sample (right column) starts on July 2 2007 and ends on June 28 2011.

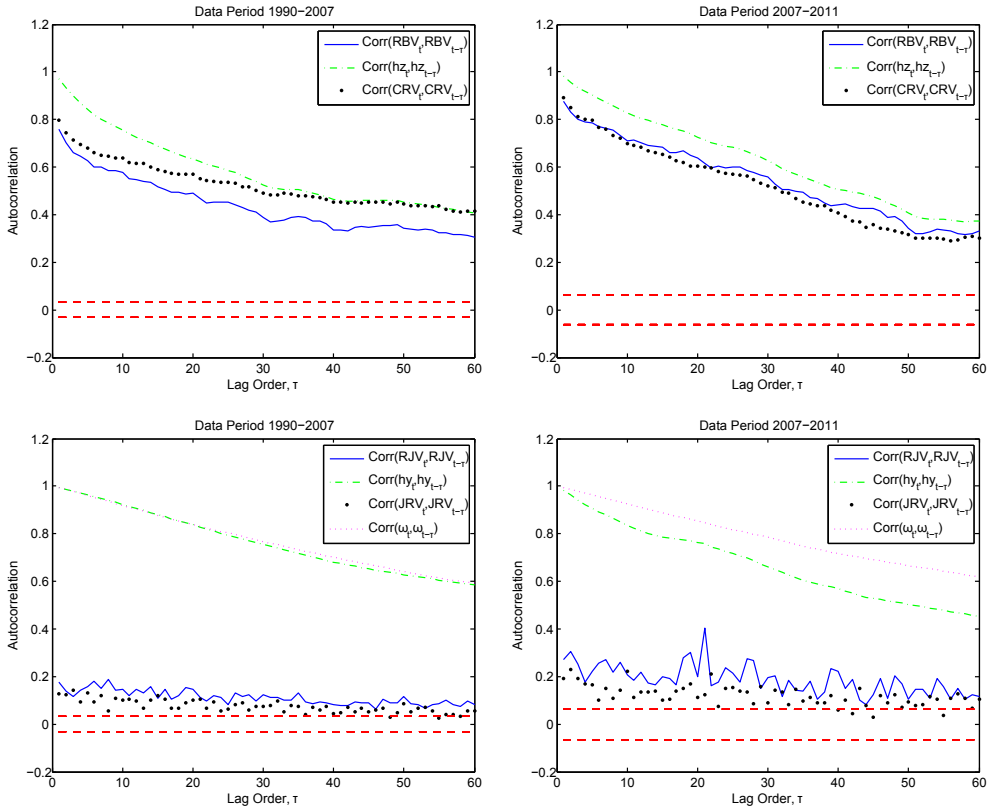

Figure 3: First row: Comparison among the auto-correlations of the Realized Volatility continuous components CRV_t and RBV_t (see Christoffersen et al. (2015)), and of the latent conditional volatility process $h_{z,t}$. Second row: Comparison among the auto-correlations of the Realized Volatility jump components JRV_t and RJV_t (see Christoffersen et al. (2015)), and of the latent intensity processes ω_t and $h_{y,t}$.

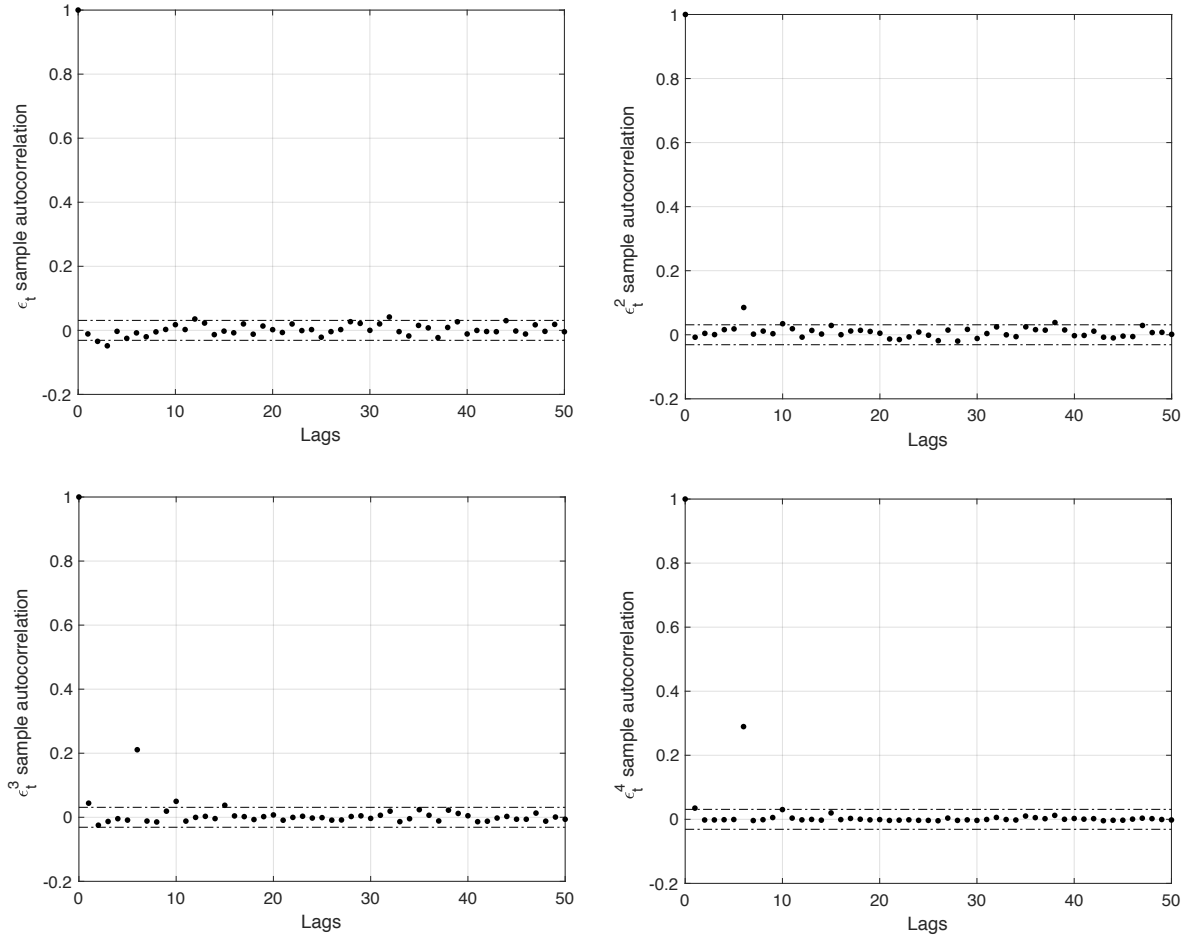

Figure 4: (Clockwise) Empirical autocorrelation of ϵ_t, ϵ_t^2, ϵ_t^3, and ϵ_t^4 for the LHARG-ARJ model and 95% confidence band for the period 3 July, 1990 – 28 June, 2011.

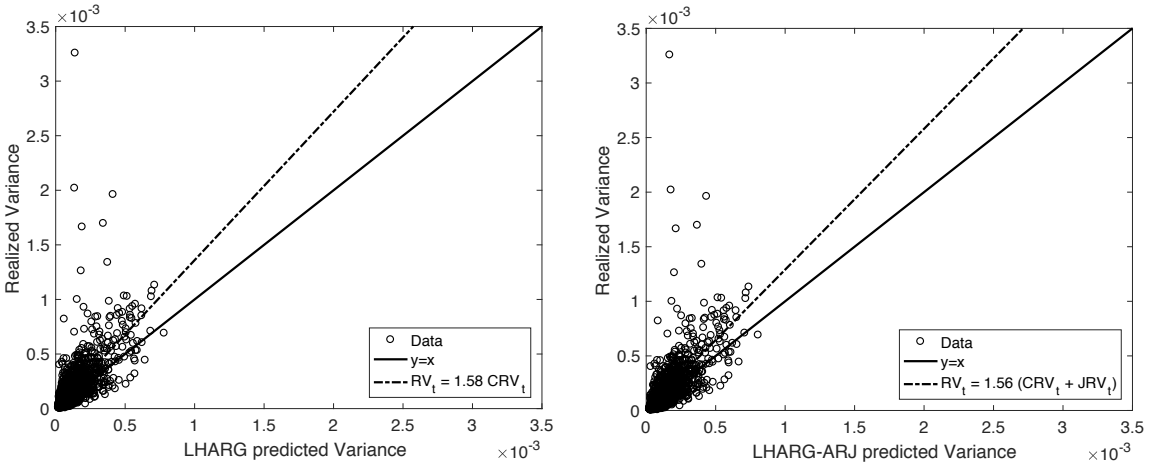

Figure 5: Realized variance and predicted variance from LHARG model (left panel) and LHARG-ARJ model (right panel). The sample starts on July 3, 1990 and ends on June 28, 2011.

	LHARG-ARJ	
Parameter	Data period 1990-2007	Data period 2007-2011
λ_c	1.3 (1.9)	-2 (2)
λ_j	0.5 (1.6)	3 (10)
θ	8.7e-06 (5e-07)	2.6e-05 (2e-06)
κ	1.43	1.18
β_d	4.2e+04 (4e+03)	1.9e+04 (3e+03)
β_w	3.4e+04 (4e+03)	1.2e+04 (3e+03)
β_m	1.7e+04 (3e+03)	9e+02 (1.5e+03)
α_d	1.3e-01 (3e-02)	5.7e-02 (1.5e-02)
α_w	1.1e-01 (2e-02)	6.7e-01 (1.8e-02)
α_m	1.1e-01 (4e-02)	5e-03 (4e-02)
γ	3.5e+02 (6e+01)	4.2e+02 (9e+01)
$\bar{\omega}$	2.9e-03	5.0e-04
ξ	9.7e-01 (1.0e-02)	9.80e-01 (1.9e-02)
ζ	2.4e-02 (7e-03)	2.0e-02 (1.5e-02)
Λ	-4.4e-04 (1.6e-04)	-1.5e-04 (4e-04)
δ	5.0e-03 (2e-04)	5.9e-03 (3e-04)
	Risk premia	
ν_c	-8.11e+03	-1.29e+03
ν_j	-8.28e+03	-2.39e+03
	Log-likelihood	
L^y	13908	2731
L^{CRV}	-22884	-6156
L^{JRV}	5735	1637
Persistence CRV_t	0.812	0.831
Persistence ω_t	0.994	0.999

Table 1: Maximum likelihood estimates (standard errors in parenthesis) for LHARG-ARJ on S&P500 Index for the two periods July 1990 – June 2007 and July 2007 – June 2011.

	LHARG Majewski et al. (2015)	LHARG-ARJ	BPJVM Christoffersen et al. (2015)
Observable conditional variance	•	•	
Heterogeneous variance structure	•	•	
Heterogeneous leverage structure	•	•	
Jumps		•	•
Latent jump intensity with AR(1) dynamics		•	•

Table 2: Review of the main features of the LHARG-ARJ, LHARG, and BPJVM models.

	LHARG	
Parameter	Data period 1990-2007	Data period 2007-2011
λ_c	3.8 (2.5)	-1.6 (3.3)
θ	8.5e-06 (4e-07)	2.6e-05 (2e-06)
κ	1.37	1.16
β_d	4.3e+04 (4e+03)	1.9e+04 (3e+03)
β_w	3.6e+04 (4e+03)	1.2e+04 (3e+03)
β_m	1.7e+04 (3e+03)	1.0e+03 (1.7e+03
α_d	1.6e-01 (3e-02)	4e-02 (1e-02)
α_w	1.0e-01 (2e-02)	5.3e-02 (1.9e-02)
α_m	8e-03 (1.3e-02)	5e-03 (1.3e-02)
γ	3.0e+02 (5e+01)	4.7e+02 (1.1e+02)
	Risk premium	
ν_c	-9.27e+03	-1.58e+03
	Log-likelihood	
L^y	13401	2634
L^{CRV}	-22866	-6160
Persistence CRV_t	0.824	0.834

Table 3: Maximum likelihood estimates (standard errors in parenthesis) for LHARG on S&P500 Index for the two periods July 1990 – June 2007 and July 2007 – June 2011.

	BPJVM	
Parameter	Data period 1990-2007	Data period 2007-2011
λ_z	1 (4)	-6 (5)
λ_y	4e-05 (8e-05)	2.7e-04 (1.9e-04)
γ	1.44e+04 (1.7e+02)	1.45e+04 (3e+02)
ω_z	2.48e-08	1.79e-09
ω_y	4.23e-02	8.53e-02
σ	1.86e-07 (2e-09)	2.43e-07 (5e-09)
θ	1e-05 (5e-05)	1.6e-04 (1.2e-04)
δ	1.278e-03 (1.5e-05)	2.01e-03 (6e-05)
ρ	3.3e-01 (2e-02)	3e-01 (4e-02)
b_z	6.50e-01 (3e-02)	6.1e-01 (5e-02)
b_y	9.51e-01 (1.2e-02)	9.09e-01 (1.7e-02)
a_z	3.5e-01 (3e-02)	3.9e-01 (4e-02)
a_y	2.2e+04 (5e+03)	2.2e+04 (5e+03)
Risk premia		
χ -5.16 ν_c -779	χ -2.01 ν_c -249	
ν_3 2.59 ν_j -6.25e+04	ν_3 1.70 ν_j -2.02e+03	
Quasi-log-likelihood	96343	21181
Persistence h_z	0.999	0.999
Persistence h_y	0.986	0.998

Table 4: Quasi maximum likelihood estimates (standard errors in parenthesis) for BPJVM on S&P500 Index for the two periods July 1990 – June 2007 and July 2007 – June 2011.

	$RMSE_{IV}$			
	1996-2007		2007-2011	
Model \ Moneyness	$0.9 < m < 1.1$	$0.8 < m < 1.2$	$0.9 < m < 1.1$	$0.8 < m < 1.2$
LHARG	5.61	7.00	6.24	7.84
LHARG-ARJ/LHARG	0.91	0.92	0.98	0.98
BPJVM	6.77 | 5.61	7.96 | 6.90	7.34 | 6.92	9.17 | 8.87
LHARG-ARJ/BPJVM	0.76 | 0.91	0.81 | 0.94	0.84 | 0.89	0.84 | 0.87

Table 5: Comparison of global option pricing performance on S&P500 Index options for models LHARG-ARJ, LHARG and BPJVM during two periods: from 1996 to 2007 and from 2007 to 2011. Rows referring to the BPJVM model: (left column) pricing kernel as in Christoffersen et al. (2015); (right column) pricing kernel as in equation (2.13).

	Maturity			
Moneyness	$\tau \leq 50$	$50 < \tau \leq 90$	$90 < \tau \leq 160$	$160 < \tau$
Panel A	LHARG Implied Volatility RMSE			
$0.8 \leq m \leq 0.9$	15.60	8.81	6.05	5.03
$0.9 < m \leq 0.98$	7.51	4.24	3.83	6.21
$0.98 < m \leq 1.02$	4.16	3.56	4.11	7.17
$1.02 < m \leq 1.1$	6.76	3.51	4.03	7.14
$1.1 < m \leq 1.2$	22.76	6.18	3.24	6.01
Panel B	LHARG-ARJ/LHARG Implied Volatility RMSE			
$0.8 \leq m \leq 0.9$	0.94	0.90	0.92	0.92
$0.9 < m \leq 0.98$	0.85	0.88	0.98	0.92
$0.98 < m \leq 1.02$	0.85	1.09	1.08	0.94
$1.02 < m \leq 1.1$	0.86	1.10	1.11	0.96
$1.1 < m \leq 1.2$	0.93	0.86	1.05	0.99

Table 6: Detailed comparison of option pricing performance on S&P500 Index options for models LHARG and LHARG-ARJ before the financial crisis from July 1996 to June 2007.

	Maturity			
Moneyness	$\tau \leq 50$	$50 < \tau \leq 90$	$90 < \tau \leq 160$	$160 < \tau$
Panel A	LHARG Implied Volatility RMSE			
$0.8 \leq m \leq 0.9$	14.56	7.85	6.03	5.50
$0.9 < m \leq 0.98$	7.72	5.44	5.56	5.74
$0.98 < m \leq 1.02$	5.32	5.58	6.09	6.53
$1.02 < m \leq 1.1$	5.96	6.26	6.78	7.05
$1.1 < m \leq 1.2$	10.84	5.92	6.59	7.60
Panel B	LHARG-ARJ/LHARG Implied Volatility RMSE			
$0.8 \leq m \leq 0.9$	0.97	0.99	1.00	0.99
$0.9 < m \leq 0.98$	0.93	1.00	0.99	0.96
$0.98 < m \leq 1.02$	0.98	1.05	1.00	0.96
$1.02 < m \leq 1.1$	1.02	1.08	1.03	0.99
$1.1 < m \leq 1.2$	0.97	1.08	1.05	0.99

Table 7: Detailed comparison of option pricing performance on S&P500 Index options for models 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			
$0.8 \leq m \leq 0.9$	12.99	5.51	4.71	11.61
$0.9 < m \leq 0.98$	6.33	4.01	4.83	11.18
$0.98 < m \leq 1.02$	3.58	3.95	5.25	11.61
$1.02 < m \leq 1.1$	5.69	4.94	6.78	14.21
$1.1 < m \leq 1.2$	18.49	8.05	10.41	21.19
Panel B	LHARG-ARJ/BPJVM Implied Volatility RMSE			
$0.8 \leq m \leq 0.9$	1.01	1.15	0.95	0.41
$0.9 < m \leq 0.98$	0.91	0.91	0.87	0.56
$0.98 < m \leq 1.02$	1.00	1.09	0.99	0.62
$1.02 < m \leq 1.1$	0.93	0.86	0.77	0.52
$1.1 < m \leq 1.2$	1.00	0.45	0.42	0.30
Panel C	BPJVM RMSE Implied Volatility RMSE			
$0.8 \leq m \leq 0.9$	12.96	5.63	4.25	7.81
$0.9 < m \leq 0.98$	6.22	3.82	4.09	7.75
$0.98 < m \leq 1.02$	3.39	3.52	4.35	8.21
$1.02 < m \leq 1.1$	5.55	4.52	5.74	10.23
$1.1 < m \leq 1.2$	18.30	7.34	8.64	14.83
Panel D	LHARG-ARJ/BPJVM Implied Volatility RMSE			
$0.8 \leq m \leq 0.9$	1.02	1.13	1.05	0.61
$0.9 < m \leq 0.98$	0.93	0.96	1.03	0.80
$0.98 < m \leq 1.02$	1.05	1.22	1.19	0.88
$1.02 < m \leq 1.1$	0.95	0.94	0.92	0.72
$1.1 < m \leq 1.2$	1.00	0.50	0.51	0.43

Table 8: Detailed comparison of option pricing performance on S&P500 Index options for models BPJVM and LHARG-ARJ before the financial crisis from July 1996 to June 2007. Panels A and B: pricing kernel as in Christoffersen et al. (2015); Panels C and D: pricing kernel as in equation (2.13).

	Maturity			
Moneyness	$\tau \leq 50$	$50 < \tau \leq 90$	$90 < \tau \leq 160$	$160 < \tau$
Panel A	BPJVM Implied Volatility RMSE			
$0.8 \leq m \leq 0.9$	14.41	8.15	7.58	9.06
$0.9 < m \leq 0.98$	8.06	7.05	7.53	9.20
$0.98 < m \leq 1.02$	5.40	6.36	7.20	9.37
$1.02 < m \leq 1.1$	5.33	6.64	7.89	12.46
$1.1 < m \leq 1.2$	9.50	9.40	10.85	14.27
Panel B	LHARG-ARJ/BPJVM Implied Volatility RMSE			
$0.8 \leq m \leq 0.9$	0.87	0.79	0.79	0.56
$0.9 < m \leq 0.98$	0.84	0.76	0.73	0.60
$0.98 < m \leq 1.02$	1.00	1.07	0.87	0.67
$1.02 < m \leq 1.1$	1.18	1.05	0.90	0.55
$1.1 < m \leq 1.2$	0.97	0.72	0.69	0.53
Panel C	BPJVM Implied Volatility RMSE			
$0.8 \leq m \leq 0.9$	14.41	8.07	7.24	7.64
$0.9 < m \leq 0.98$	8.00	6.75	6.87	7.62
$0.98 < m \leq 1.02$	5.26	5.93	6.46	7.85
$1.02 < m \leq 1.1$	5.26	6.34	7.31	10.89
$1.1 < m \leq 1.2$	9.42	8.99	10.16	12.76
Panel D	LHARG-ARJ/BPJVM Implied Volatility RMSE			
$0.8 \leq m \leq 0.9$	0.88	0.80	0.83	0.66
$0.9 < m \leq 0.98$	0.84	0.79	0.79	0.72
$0.98 < m \leq 1.02$	1.03	1.02	0.97	0.81
$1.02 < m \leq 1.1$	1.19	1.10	0.97	0.63
$1.1 < m \leq 1.2$	0.98	0.75	0.73	0.60

Table 9: Detailed comparison of option pricing performance on S&P500 Index options for models BPJVM and LHARG-ARJ during and after the financial crisis from July 2007 to June 2011. Panels A and B: pricing kernel as in Christoffersen et al. (2015); Panels C and D: pricing kernel as in equation (2.13).