Volatility Bursts: A discrete-time option model with multiple volatility components

Francesca Lilla*

I propose an affine discrete-time model, called Vector Autoregressive Gamma with volatility Bursts (VARG-B) in which volatility experiences, in addition to frequent and small changes, periods of sudden and extreme movements generated by a latent factor which evolves according to the Autoregressive Gamma Zero process. A key advantage of the discrete-time specification is the possibility of estimating the model via the Extended Kalman Filter. Moreover, the VARG-B model leads to a fully analytic conditional Laplace transform which leads to a closed option pricing formula. When estimated on S&P500 index options and returns the new model provides more accurate option pricing and modelling of the IV surface with respect to some alternative models.

---

*Bank of Italy, Directorate General for Economics, Statistics and Research, francesca.lilla@bancaditalia.it.

1 Introduction[1]

The price of options depends on "extreme" movements of the underlying asset price, in addition to idiosyncratic asset changes. Since asset prices must follow semimartingales to avoid arbitrage opportunities, the literature has mainly focused on continuous time models. In this framework, extreme movements (typically associated with unexpected macroeconomic news announcements) are commonly modelled as discontinuities in the price trajectories: jumps. Starting from the seminal paper of Merton (1976), methods to distinguish between volatility and jump risk were assessed to fit low frequency data. However, the results provided in Christensen, Oomen, and Podolskij (2014) (COP henceforth) weaken the consensus in the literature about the presence of jumps in asset price! Using intraday data, COP examine the role of the jump component by applying new econometric techniques to a set of individual order-level tick data. The authors find that jumps account for only about 1% of quadratic price variation, which is substantially smaller than what typically found by lower-frequency literature. Moreover, COP show that the price continuity is often preserved. They suggest that sharp movements of asset prices over short periods of time are generated by high volatility episodes instead of genuine price jumps. The much reduced role for jumps and, consequently, the elevated role for volatility process, calls for a stronger effort in modelling the volatility dynamics and carries important implications for asset pricing models.

One goal of this paper is to replace jumps with volatility bursts in the dynamics of the underlying and understand if this replacement carries over to option valuation! This challenge is addressed in discrete time by specifying an affine discrete-time model, called Vector Autoregressive Gamma with volatility Bursts (VARG-B), which is characterised by a multifactor volatility specification. In the VARG-B model volatility experiences, in addition to frequent and small changes, periods of sudden and extreme movements, i.e.

volatility bursts. From a modelling perspective, the introduction of volatility bursts in a discrete time setting requires an additional state variable in the volatility dynamics. In this work, the volatility is assumed to be latent, and at each point in time, it is the result of the sum of two independent random variables. The frequent and small changes are generated by the first state variable, called *continuous* component, while the *volatility burst* component generates the volatility changes due to extreme price movements. The first variable is modelled as an Autoregressive Gamma process (ARG), see Gouriéroux and Jasiak (2006), while the second follows an Autoregressive Gamma Zero (ARG-Zero) process. This last process is proposed in Monfort et al. (2014) to model the Zero Lower Bound in the term structure of interest rate. The ARG-Zero process, whose introduction into the volatility dynamics represents the main innovation of this paper, is a suitable statistical channel for describing sudden volatility changes since it is coherent with non-negative volatility bursts and it accommodates extended periods of zero or close-to-zero values.

A key advantage of the proposed specification is that the model can be estimated using the Extended Kalman Filter, which allows filtering the time series of both volatility components. Thanks to this flexible estimation strategy it is possible to understand the relative contribution of both the volatility factors to the total conditional variance of log-returns. Moreover, the VARG-B model leads to a fully analytic conditional Laplace transform, thanks to its exponential affine form, which is mainly attractive for option pricing purposes. Indeed, the change of measure is performed adopting an exponentially affine stochastic discount factor which preserves all the analytical results in order to obtain a closed-form option pricing formula. Finally, the affine discrete-time model presented in this paper is intuitive and easy to estimate.

Specifically speaking, the VARG-B model extends the LHARG-RV model by Majewski et al. (2015) in which the volatility does not display sudden, and large changes and is perfectly observed through a realized measure (RM). More recently, Alitab et al. (2015) introduce the jump variation in the volatility dynamics which is, also in this case, observed in all its components through high frequency data-based estimators. To the best of my

3

knowledge Caporin et al. (2015) and Caporin et al. (2017) represent the only attempt at modelling the occurrence and the probability of volatility bursts in a discrete time setting. Caporin et al. (2015) extend the HAR model of Corsi (2009) with a linear and additive volatility burst factor (HAR-V-J, see Andersen et al., 2007 and Corsi et al., 2010) finding a positive probability of volatility bursts which are more likely to happen during financial crises. Caporin et al. (2017) extend the MEM model of Engle and Gallo (2006) with a multiplicative volatility burst factor (MEM-J). The authors find that the MEM-J significantly increases the model fit on the right tail of the volatility distribution[2]. In all these papers the authors assume that the price process is characterised by jumps and except for Alitab et al. (2015) which explicitly provide the jump variation component, they remove the spikes in the RM time series using a local volatility estimator: the RM is assumed to be the proxy of the returns integrated variation. Clearly, assuming the presence of jumps in the price process, the ability to disentangle the two sources of risk, strictly depends on the methodology employed for the jump identification and the results can be driven by the method chosen.

This paper differs from this discrete time literature since the price process is assumed to be free of jumps (following the COP intuition) and the extreme price movements being, instead, a by-product of volatility bursts. Since volatility is assumed to be latent, information about the state variables is recovered through a measurement equation introduced in the estimation procedure.

The empirical results on a large sample of S&P500 Index options emphasise the superior ability of the VARG-B model in pricing options along moneyness and time to maturity dimensions. These findings indicate the benefit of a multi-factor volatility specification that allows for volatility bursts in addition to usual changes.

The paper is organised as follows. In Section 2, I develop the return process and

---

[2]Caporin et al. (2017) compare alternative MEM specification (no jumps, constant and time-varying jump intensity) with respect to their ability in fitting the dynamics of the series analysing the dynamic properties of the residuals. Moreover, the authors study the ability of the MEM specifications to correctly predict the probability of tail events providing a Volatility-at-Risk exercise.

the volatility dynamics under the historical measure. Section 3 estimates the historical process via pseudo maximum likelihood with Extended Kalman Filter. In Section 4, I propose an economic application where I analyse the performance of the VARG-B model in an option pricing exercise. In particular, I derive the change of measure for the new model. I calibrate the model on options, provide the pricing and analyses its performance. Finally, Section 5 concludes. The proofs of all propositions are provided in the appendix.

2   The VARG-B model

The goal of this section is to build a model for option valuation that allows for volatility bursts in the return dynamics. I develop a model in which volatility is latent in all its components.

2.1   The VARG-B physical dynamics

The VARG-B model explicitly accounts for the probability of having large and sudden movements in the volatility dynamics through an additive component modelled with a new process, i.e. ARG-Zero.

Under the VARG-B specification the discrete-time stochastic volatility model for daily log-returns is the following:

$$y_{t+1} := \log\left(\frac{S_{t+1}}{S_t}\right) = r_{t+1} + \lambda f_{t+1} + \sqrt{f_{t+1}}\epsilon_{t+1} \tag{1}$$

where r_{t+1} denotes the risk-free rate at time $t+1$, assumed to be exogenous, and where λ is the market price of risk. The innovation ϵ_{t+1} is i.i.d. $\mathcal{N}(0,1)$. Moreover, the latent factor f_{t+1} denotes the true volatility, and it is equal to the sum of two independent components:

$$f_{t+1} = f_{1,t+1} + f_{2,t+1}. \tag{2}$$

Given the information set at time t, denoted $\mathcal{F}_t$, the continuous volatility component follows an Autoregressive Gamma (ARG) process:

$$f_{1,t+1}|\mathcal{F}_t \sim \gamma_v(\beta_1 f_{1,t}, \mu_1) \quad for \quad v > 0, \beta_1 > 0, \mu_1 > 0 \tag{3}$$

The process in (3) is defined by a shape parameter v, a noncentrality parameter $\beta_1 f_{1,t}$ and a scale parameter μ_1. The history of the process determines the entire noncentrality coefficient $\beta_1 f_{1,t}$, which is written as a linear function of the lagged value of the process.

Since the ARG process is a discretized version of the Cox, Ingersoll Jr, and Ross (1985) (CIR) model, it is sufficiently flexible to represent the volatility of financial asset.

Moreover, given $\mathcal{F}_t$ and $f_{1,t+1}$, the volatility burst component follows an Autoregressive Gamma Zero (ARG-Zero) process:

$$f_{2,t+1}|\mathcal{F}_t \sim \gamma_0(d_2 + \beta_2 f_{2,t}, \mu_2) \quad for \quad d_2 \geq 0, \beta_2 > 0, \mu_2 > 0 \tag{4}$$

denoted as $ARG_0(d_2 + \beta_2 f_{2,t}, \mu_2)$.

In (2) $f_{1,t+1}$ allows for frequent and small changes that characterize the volatility dynamics and it is called *continuous* volatility component. $f_{2,t+1}$ represents an "exceptional" volatility component that let volatility to experience periods of big and sudden changes. The latter is the focus of this model and is called *volatility burst* component.

On the contrary of the ARG by Gouriéroux and Jasiak (2006), the process in (4) is characterised by a zero lower bound: $f_{2,t+1}$ can take zero value with a strictly positive probability, stay at zero for a more or less extended period of time and become positive again. In order to understand the behaviour of an ARG-Zero process, I define its main characteristics.

The conditional probability density function $p(f_{2,t+1}|\mathcal{F}_t)$ of the $ARG_0(d_2 + \beta_2 f_{2,t}, \mu_2)$

is given by:

$$p(f_{2,t+1}|\mathcal{F}_t;\phi) = \sum_{z=1}^{+\infty} \left[\frac{\exp(-f_{2,t+1}/\mu_2)f_{2,t+1}^{z-1}}{(z-1)!\mu_2^z} \times \frac{\exp[-(d_2+\beta_2 f_{2,t})](d_2+\beta_2 f_{2,t})^z}{z!} \right] \mathbb{1}_{\{f_{2,t+1}>0\}}$$

$$+ \exp(-d_2 - \beta_2 f_{2,t})\mathbb{1}_{\{f_{2,t+1}=0\}}$$

(5)

where $\phi = d_2, \beta_2, \mu_2$.

Moreover, f_{t+1} in (2) is stationary if and only if $\rho_j = \beta_j \mu_j < 1$ for $j = 1, 2$ [3].

The zero-point mass in equation (5) which is allowed by the zero shape parameter emphasizes the key feature of the ARG-Zero process: the zero lower bound. In fact, the probability of $f_{2,t+1}$ of reaching zero is equal to the second term on the right hand side of equation (5), i.e. $\exp(-d_2 - \beta_2 f_{2,t})$. Another important feature of the ARG-Zero process is represented by the positive intercept, i.e. d_2. The conditional probability of $f_{2,t+1}$ of remaining at the zero lower bound is $\exp(-d_2)$. When $d_2 = 0$, the zero lower bound becomes an absorbing state, since $\exp(-d_2) = 1$. On the contrary, the presence of a strictly positive intercept prevents $f_{2,t} = 0$ to be an absorbing state: if $d_2 > 0$, than $\exp(-d_2) < 1$. Indeed, the lower is d_2, the greater is the conditional probability of having more extended periods of zero values for $f_{2,t+1}$.

The VARG-B model in (1)-(4) has an important advantage from the option pricing perspective: it leads to a fully analytic conditional Laplace transform. Despite the complexity of the conditional density function in (5), the conditional Laplace transform is easy to manipulate, and it is equal to:

---

[3]The stationarity condition for the ARG-Zero process is illustrated in *Corollary 2.1.2* in Monfort et al. (2014). The stationarity condition for the ARG process is shown in *Proposition 2* in Gouriéroux and Jasiak (2006).

$$\varphi_t(u) := \mathbb{E}\left[\exp(u_1 f_{1,t+1} + u_2 f_{2,t+1})|\mathcal{F}_t\right]$$

$$= \exp\left[\frac{u_1\mu_1}{1 - u_1\mu_1}\beta_1 f_{1,t} + \frac{u_2\mu_2}{1 - u_2\mu_2}(d_2 + \beta_2 f_{2,t}) - \nu\log(1 - u_1\mu_1)\right] \quad (6)$$

$$for \quad u_1 < \frac{1}{\mu_1} \quad and \quad u_2 < \frac{1}{\mu_2}$$

I illustrate the relevance of the ARG-Zero process for volatility bursts modelling with a simple simulation exercise. The dynamics of the volatility burst factor is given by the ARG_0 in (4) with $d_2 = 0.1$ and $\mu_2 = 0.01$.

The chosen parameters are calibrated in such a way that the unconditional mean and the unconditional variance of $f_{2,t}$ are about 0.2 and 0.4, respectively. Given these parameters, I simulate 1000 periods for the process. The conditional probability for $f_{2,t}$ of remaining at zero is equal to $\exp(-d_2) = \exp(-0.1) = 0.9$.

From Figure 1, $f_{2,t}$ is characterised by extreme and sudden changes as well as by many episodes of periods of zero values. This behaviour of the ARG_0 process is particularly appealing to model the extraordinary nature of volatility bursts.

From an economic point of view, if the volatility burst component is modelled as an ARG-Zero process, d_2 identifies the average persistence of zero lower bound regimes. Given the exceptional characteristic of the bursts, d_2 should be greater than zero and small in magnitude. The evidence provided in Section 3.3 confirms this theoretical feature.

3   Performance on data

In the previous section, I have laid out the general framework for incorporating volatility bursts when modelling return dynamics. In this section, I develop a Kalman filter-based estimation method to estimate the physical ($\mathbb{P}$) parameters using daily observations on returns and information about volatility provided by the realized volatility measure (Andersen and Bollerslev, 1998). I also briefly describe two alternative models that I es-
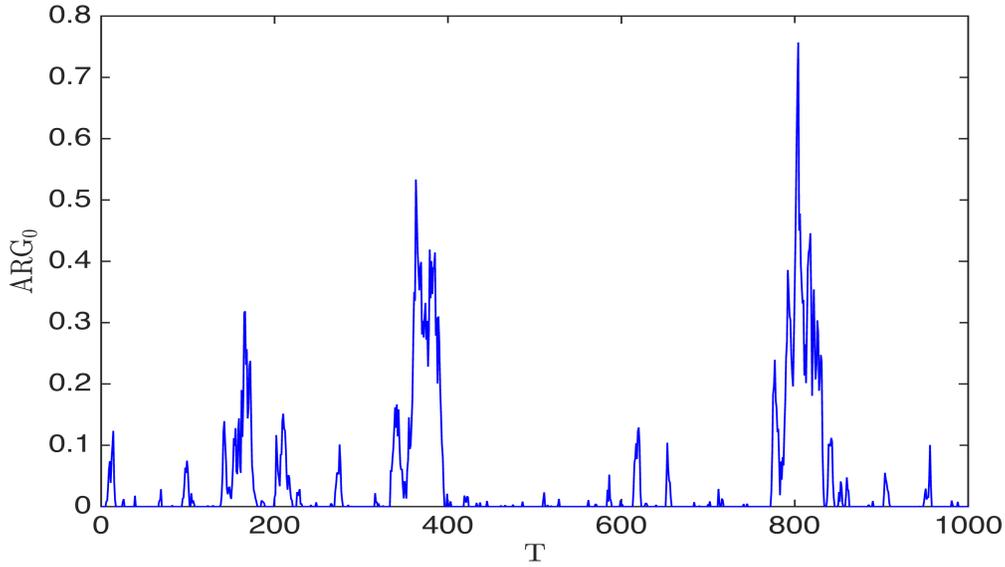
Figure 1: Simulated path of the volatility burst factor defined by the following conditional distribution: $f_{2,t+1}|\mathcal{F}_t \sim \gamma_0(d_2 + \beta_2 f_{2,t}, \nu_2)$. T is the total number of periods considered.

timate in order to have an idea of the relative VARG-B performance. The purpose of this section is dual: to illustrate how to estimate physical parameters on real data, and to propose empirical support to the VARG-B specification in which price jumps are replaced by volatility bursts.

3.1 Estimation strategy

The VARG-B model can be represented in a state-space form which can be estimated via pseudo-maximum likelihood with the Extended Kalman filter. The pseudo maximum likelihood is feasible since the first two conditional moments are available in closed-form. The Kalman filter strategy is natural in this framework since the VARG-B model is affine in the state variables.

The first measurement equation is directly obtained from the daily log-returns dynamics in (1). Assuming that returns of financial assets arise through discrete observations from an underlying continuous-time process, I augment the state space model with a sec-

9

ond measurement equation. The latter measurement equation relates the latent volatility to an ex-post estimator of daily quadratic variation of the log-price process over the period $t + 1$, i.e. Realized Volatility (RV) by Andersen and Bollerslev (1998).

The transition equations are given by the factor dynamics, specified according to the conditional mean and the conditional variance:

$$E(f_{jt+1}|\mathcal{F}_t) = \mu_j(d_j + \beta_j f_{jt}) + \mu_j v_j \qquad \text{for} \quad j = 1, 2 \qquad (7)$$

$$V(f_{jt+1}|\mathcal{F}_t) = 2\mu_j^2(d_j + \beta_j f_{jt}) + \mu_j^2 v_j \qquad \text{for} \quad j = 1, 2 \qquad (8)$$

where $v_1 = v$ as defined in (3) and $v_2 = 0$ as in (4).

In what follows, I formally lay down the state space model.

Using (7) and (8), the transition equations can be expressed as follows:

$$f_{1,t+1} = \mu_1 v + \mu_1 \beta_1 f_{1,t} + \sqrt{\mu_1^2(v + 2\beta_1 f_{1,t})} v_{t+1}^1 \qquad (9)$$

$$f_{2,t+1} = \mu_2 d_2 + \mu_2 \beta_2 f_{2,t} + \sqrt{\mu_2^2(2d_2 + 2\beta_2 f_{2,t})} v_{t+1}^2 \qquad (10)$$

where v_{t+1}^1 and v_{t+1}^2 are independent white noises with zero mean and unit variance.

The measurement equations describe the relationship between two types of observable variables and both the latent volatility factors:

$$y_t = r_t + \lambda(f_{1,t} + f_{2,t}) + \sqrt{f_{1,t} + f_{2,t}} \epsilon_t \qquad (11)$$

$$RV_t = \eta_0 + \eta_1(f_{1,t} + f_{2,t}) + \zeta_t \qquad (12)$$

$\zeta_t \sim IIDN(0, \sigma^2)$ and ϵ_t and ζ_t are independent.

The innovation term in (12) can be interpreted as a measurement error with zero mean and constant variance. Indeed, the estimator RV is characterised by an attenuation bias generated by the finite nature of the price sample (classic measurement problem) and by the absence of trading during the night (overnight effect).

I estimate parameters in (9)-(12) via pseudo-maximum likelihood with the Extended

Kalman filter. Since both y_t and f_t are conditionally heteroskedastic, the Extended Kalman filter is applied. Moreover, the true log-likelihoods derived from conditional non-central Gamma distributions are replaced by Gaussian distributions. This means that v_{t+1}^1 in (9) and v_{t+1}^2 in (10) are approximated by standard Gaussian white noises.

The application of this estimation strategy represents a strength of the VARG-B model. The availability of a state-space model makes the implementation of such a procedure straightforward, letting data dictate the relative contribution of both factors to the total volatility.

3.2 Alternative models

The VARG-B model is a discrete time specification in which the volatility dynamics is given by the sum of two independent factors.

First, by setting $f_{2,t+1} = 0$ in (2), I obtain the standard ARG model by Gouriéroux and Jasiak (2006).

Second, I can shut down the volatility burst component, and I can allow for long-memory in volatility, providing the HAR specification of Corsi (2009) for the non-central parameter in (3). In this case, I obtain the HARG model proposed in Corsi et al. (2013). Coherently with the basic idea of Corsi et al. (2013) and Majewski et al. (2015), in both the alternative specifications, the volatility (continuous component only) is assumed to be observed through the RV measure. In this way, the original estimation strategy is maintained and the RV-based option valuation framework is not distorted.

3.3 Parameter estimates

The dependent variable in the first measurement equation is defined using daily returns of the S&P500 Index from January 5,1996 to December 30, 2005. The RV time series in the second measurement equation is obtained using returns on the S&P500 Index from January 5, 1996 to December 30, 2005 sampled at 5 minutes frequency which represents the trade-off between accuracy and microstructure noise (Madhavan, 2000, Biais et al.,

2005 and McAleer and Medeiros, 2008 for surveys on this topic).

Based on the two measurement equations (11)-(12) provided in the state space model, daily observations are used to filter both state variables, i.e. continuous and volatility burst. Table 1 shows the estimated parameters, and the relative standard errors for the VARG-B model and the alternative models presented in Section 3.2. According to the estimates, all VARG-B coefficients are statistically significant with the only exception of d_2, which is small in magnitude. Note that a small value for the intercept d_2 translates in an increasing ability of the ARG-Zero process to describe extraordinary and exceptional changes in the volatility, thus supporting the proposed specification. On the one hand, a small value for the intercept in (4) increases the probability of having extended periods of zero values after sudden movements. These unexpected changes have a specific persistence given by $\mu_2\beta_2$. On the other hand, a strictly positive value for d_2 prevents the zero lower bound of being an absorbing state. These are the two main features that most motivate a model specification for bursts in the volatility dynamics.

This result is in line with the updated[4] time series of both volatility factors reported in Figure 2. Indeed, $f_{1,t}$ in (3) describes the small and frequent volatility changes while $f_{2,t}$ in (4) allows volatility to experience extraordinary movements due to unexpected news. Figure 3 shows the RV and the updated values for $f_t = f_{1,t} + f_{2,t}$. The VARG-B model fits volatility changes, measured *ex-post* by RV. As explained before, the measurement error in (12) can be linked to the attenuation bias characterising RV. Indeed, augmenting the state space model with a relation between state variables and observed data deals with the usual errors-in-variables problem.

The ARG and HARG coefficients are all significant. The parameters of the HARG model show a decreasing impact of the past lags on the present value of the RV, which is in line with the literature. The model selection approach based on both the AIC (Akaike, 1998) and the BIC (Schwarz et al., 1978) select the VARG-B to model volatility. In particu-

---

[4]At each point in time, the current value of the state variable is updated on the basis of the observations of both returns and RV.

Table 1: Estimate of the parameters under the historical measure and standard errors (in parenthesis) for the VARG-B, ARG and HARG model. The parameters reported in the first column are estimated via pseudo maximum likelihood with Extended Kalman filter. The parameters of both ARG and HARG models are estimated using Maximum Likelihood. The historical data for all models are daily RV computed on 5-minutes data and daily returns of the S&P500 Index from January 5, 1996 to December 30, 2005.

Parameter	VARG-B	Parameter	ARG	HARG
λ	-0,0987 (0,0194)	λ	0,0085 (0,0115)	0,0085 (0,0115)
μ_1	0,2932 (0,0441)	μ	0,1691 (0,0001)	0,1403 (0,0026)
β_1	2,7360 (0,5102)	β_d	3,9891 (0,0088)	3,0614 (0,1533)
μ_2	2,2532 (0,7933)	β_w	-	1,7162 (0,2150)
ν	0,8085 (0,1358)	β_m	-	0,6910 (0,1863)
d_2	5,291e-08 (1,395e-07)	ν	1,5875 (0,0084)	1,3757 (0,0270)
β_2	0,4186 (0,1470)			
σ	0,0522 (0,0376)			
η_0	0,0705 (0,0114)			
η_1	0,4898 (0,0514)			
Log-likelihood	-4946	Log-likelihood	-14997	-14772

lar, for both information criteria, the VARG-B model registers the smallest value. Indeed among the models presented in this section, the VARG-B offers the best goodness of fit.
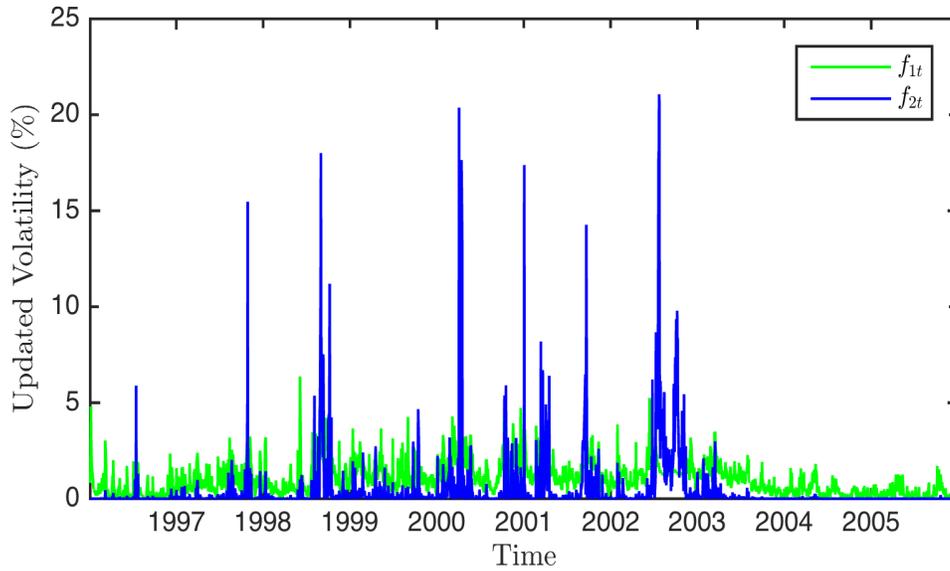
Figure 2: This figure shows the updated time series of the continuous ($f_{1,t}$, green line) and the burst ($f_{2,t}$, blue line) factors which are obtained applying the estimation procedure based on the Extended Kalman filter. The sample consists of S&P500 Index data from January 5, 1996 to December 30, 2005.

4   Economic application

In this section, I show how the VARG-B model developed under the physical measure can be used for option pricing. I first derive the moment generating function under the $\mathbb{P}$ measure for the VARG-B model and show that it is affine. I then define a stochastic discount factor which implies that the risk-neutral ($\mathbb{Q}$) moment generating function is of the same form as its physical counterpart. Then I approximate option prices using the COS efficient scheme[5] by Fang and Oosterlee (2008). All the propositions presented in this section are directly derived from the theoretical results presented in Majewski et al. (2015). Empirical results and VARG-B pricing performance follow.

---

[5]The COS method is an option pricing numerical method for European options based on the Fourier-cosine series. It takes advantage of the relation of the characteristic function with the series coefficients of the Fourier-cosine expansion of the density function. It is available once the characteristic function of the log-asset price is known and it has been proved to be fast and efficient.
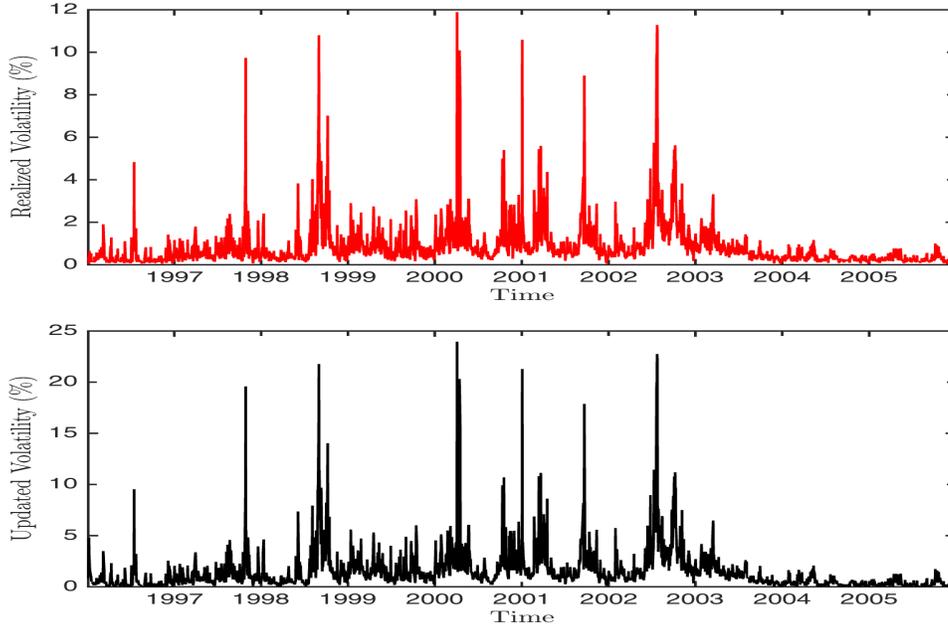
Figure 3: This figure shows the comparison between RV_t (top figure) and the updated time series of f_t (bottom figure). The sample consists of S&P500 Index data from January 5, 1996 to December 30, 2005.

4.1 The hystorical moment generating function

Equations (1) and (3)-(4) completely characterize the VARG-B model under the $\mathbb{P}$ measure. A great advantage of the VARG-B model is that its Moment Generating Function (MGF) satisfies the affine property.

Proposition 1. *The setup in Section 2.1 satisfies the Assumption 1 in Majewski et al. (2015):*

$$\mathbb{E}^{\mathbb{P}} \left[\exp(z y_{t+1} + \mathbf{b}' \mathbf{f}_{t+1}) | \mathcal{F}_t \right] \qquad (13)$$
$$= \exp \left[\mathcal{A}(z, \mathbf{b}) + \mathcal{B}(z, \mathbf{b})' \cdot \mathbf{f}_t \right]$$

for some functions $\mathcal{A} : \mathbb{R} \times \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}, \mathcal{B} : \mathbb{R} \times \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}^k$ given in (A.9) and (A.10).

Proof. See Appendix A.

Proposition 1 implies that the moment generating function of $\log(S_T/S_t)$ under the $\mathbb{P}$ measure is given by a recursive relation in terms of the functions $\mathcal{A}$ and $\mathcal{B}$.

15

Proposition 2. *Let $y_{t,T} = \log(S_T/S_t)$, under $\mathbb{P}$, the MGF for the VARG-B model has the following form*

$$\varphi^{\mathbb{P}}_{t,T,z} = \mathbb{E}^{\mathbb{P}}[e^{zy_{t,T}}|\mathcal{F}_t] = exp\left(a_t + \mathbf{b}'_t\mathbf{f}_t\right) \tag{14}$$

with $z \in \mathbb{R}$, $\mathbf{b} \in \mathbb{R}^2$ and where

$$
\begin{aligned}
a_s &= a_{s+1} + zr - \nu W_{1,s+1} + d_2 V_{2,s+1} \tag{15} \\
\mathbf{b}'_s &= \mathbf{b}'_{s+1} + (V_{1,s+1}, V_{2,s+1})\boldsymbol{\beta} \tag{16}
\end{aligned}
$$

with

$$x_{h,s+1} = x_h(z, \mathbf{b}_{s+1}) = \mathbf{b}'_{s+1} + z\lambda + \frac{z^2}{2}, \quad h = 1,2$$

subject to the initial conditions:

$$a_T = 0, \quad \mathbf{b}'_T = 0$$

The functions V and W are defined as follows:

$$
\begin{aligned}
V_{h,s+1} &= V_h(x_{h,s+1}, \mu_h) = \frac{x_{h,s+1}\mu_h}{1 - x_{h,s+1}\mu_h}, \quad h = 1,2 \\
W_{1,s+1} &= W_1(x_{1,s+1}, \mu_1) = \log\left[1 - x_{1,s+1}\mu_1\right]
\end{aligned}
$$

Proof. *See Appendix D*

The parameters under $\mathbb{P}$ are given by

$$\boldsymbol{\psi} = [\lambda, \nu, \mu_1, \mu_2, d_2, \beta_1, \beta_2] \tag{17}$$

Apart from λ, all of them are assumed to be nonnegative.

4.2 The VARG-B risk-neutral dynamics

I introduce an assumption on the Stochastic Discount Factor (SDF) that allows to obtain the risk neutral distribution and, therefore, to compute option prices. In specifying the SDF, I follow Majewski et al. (2015)[6]:

$$M_{t,t+1} = \frac{\exp(-\delta_2 y_{t+1} - \delta_{11} f_{1,t+1} - \delta_{12} f_{2,t+1})}{\mathbb{E}^{\mathbb{P}}[\exp(-\delta_2 y_{t+1} - \delta_{11} f_{1,t+1} - \delta_{12} f_{2,t+1})|\mathcal{F}_t]} \tag{18}$$

This SDF is a very flexible specification since it identifies two risk premia, i.e. δ_{11} and δ_{12} in addition to the usual equity premium, i.e. δ_2.

More precisely, δ_{11} compensates for the continuous volatility while δ_{12} compensates for the burst source of risk[7].

Proposition 3. *Under the model specification in (1) and (3)-(4) with the SDF specified as in (18), the VARG-B satisfies the no-arbitrage condition if and only if*

$$\delta_2 = \lambda + \frac{1}{2} \tag{19}$$

Proof. *See Appendix B*

Given the result in Proposition 3 and the market incompleteness, δ_{11} and δ_{12} are free parameters to be calibrated while δ_2 is considered as fixed. So, the no-arbitrage condition fixes the level of the equity risk premium, while both the continuous and burst variance risk premia are free parameters to be calibrated on the option prices sample.

The SDF in (18) belongs to the family of the exponential-affine factors. Indeed, it is possible to compute a recursion under the $\mathbb{Q}$ measure analogous to that given under $\mathbb{P}$.

Corollary 4. *Under $\mathbb{Q}$, the MGF for the VARG-B model has the following form:*

---

[6]Corsi et al. (2013) introduce a SDF involving both the log-returns and Realized Volatility, applying a modified version of the standar discrete-time exponential affine SDF applied in Gourieroux and Monfort (2007). Majewski et al. (2015) present a more general and flexible version.

[7]Many authors (see Gagliardini et al., 2011,Christoffersen et al., 2013, Corsi et al., 2013 and Majewski et al., 2015) recognized the importance of variance-dependent risk premia in SDF in reconciling the time series properties of asset returns with the cross-section of option prices.

$$\varphi^Q_{\delta_2, \delta_1}(t, T, z) = \mathbb{E}^Q\left(e^{zy_{t,T}} | \mathcal{F}_t\right) = \exp\left(a_t^* + \mathbf{b}_t^{*\prime}\mathbf{f}_t\right) \tag{20}$$

where:

$$a_s^* = a_{s+1}^* + \mathcal{A}(z - \delta_2, \mathbf{b}_{s+1}^* - \boldsymbol{\delta_1}) - \mathcal{A}(-\delta_2, -\boldsymbol{\delta_1}) \tag{21}$$

$$\mathbf{b}_s^* = \mathbf{b}_{s+1}^* + \mathcal{B}(z - \delta_2, \mathbf{b}_{s+1}^* - \boldsymbol{\delta_1}) - \mathcal{B}(-\delta_2, -\boldsymbol{\delta_1}) \tag{22}$$

with terminal conditions: $a_T^* = 0, \mathbf{b}_T^* = \mathbf{0}$ and $\mathcal{A}(\cdot, \cdot)$, $\mathcal{B}(\cdot, \cdot)$ as in (A.9) and (A.10), respectively.

Proof. *See Appendix C*

The comparison between the physical and the risk-neutral MGFs provides a one-to-one mapping between the set of parameters under $\mathbb{P}$ and the set of parameters under $\mathbb{Q}$.

Proposition 5. *Under the risk-neutral measure, $\mathbb{Q}$ the latent volatility still follows a VARG-B process with parameters*

$$
\begin{aligned}
d_2^Q &= \frac{d_2}{1 - y_2^* \mu_2} & \beta_1^Q &= \frac{\beta_1}{1 - y_1^* \mu_1} & \mu_1^Q &= \frac{\mu_1}{1 - y_1^* \mu_1} \\
\beta_2^Q &= \frac{\beta_2}{1 - y_2^* \mu_2} & \mu_2^Q &= \frac{\mu_2}{1 - y_2^* \mu_2} & \nu^Q &= \nu
\end{aligned}
\tag{23}
$$

where $y_h^* = -\delta_{1h} - \delta_2 \lambda + \frac{\delta_2^2}{2}$ for $h = 1, 2$.

   Proof. *See Appendix E*

The result proved in Proposition 5 and the analytical tractability of the VARG-B process simplify the computation of the risk-neutral MGF. In fact, the latter is obtained starting from the MGF under $\mathbb{P}$ and substituting the parameters under $\mathbb{P}$ with those under $\mathbb{Q}$.

Corollary 6. *Under $\mathbb{Q}$ the MGF for the VARG-B model has the same form as in (14) with equity risk premium*
$\lambda^Q = -\frac{1}{2}$ and $d_2^Q, \boldsymbol{\beta}^Q, \mu_1^Q, \mu_2^Q, \nu^Q$ as in (23).

Therefore, f_t is still a VARG-B process under the $\mathbb{Q}$ measure and the two risk premia δ_{11} and δ_{12} are the only parameters to be calibrated on option prices, as explained in Section 4.4. Once the values of δ_{11} and δ_{12} are calibrated, all the parameters in (23) can be computed in closed-form following Corollary 6.

4.3 Data and stylized facts

The data used in this exercise consist of European call option prices written on the S&P500 Index for the same time period of data used in Section 3.3. The observations for the option prices range from January 5, 1996 to December 30, 2005. I only use Wednesday options data[8] yielding a total of 30061 observations. As it is customary in the literature (see Barone-Adesi et al., 2008, Corsi et al., 2013, Majewski et al., 2015), I filter out options with time to maturity less than 10 days or more than 360 days, implied volatility larger than 70% or prices less than 0.05. To perform the analysis, I split options into different categories according to time to maturity and moneyness. Moneyness (m) is defined as the underlying index level divided by the option strike price. A call option is defined as DOTM (deep out-of-the-money) if m≤ 0.94, OTM if 0.94 <m≤ 0.97, ATM if 0.97 <m≤ 1.03, DITM if 1.03 <m≤ 1.06 and ITM if m> 1.06. Based instead on the time to maturity (T), options are classified in four categories: short maturity if T≤50, medium-short maturity if 50 <T≤ 90, medium-long if 90 < T≤ 160 and long maturity if T> 160. Table 2 reports some descriptive statistics for the options classified by the moneyness and maturity definitions given above. From Panel A, the DOTM call options are heavily traded especially at longer maturity. According to the summary statistics in Panel B and Panel C, the observed implied volatility increases as option intrinsic value increases. Hence, ITM (DITM) calls are more expensive compared to OTM (DOTM) calls.

---

[8]The first motivation for using Wednesday data is that Wednesday is the day of the week least likely to be a holiday. Therefore, it is less likely than other days to be affected day-of-the-week effects (see Bakshi et al., 1997, Christoffersen et al., 2008).

Table 2: Summary statistics for the S&P500 Index option data. The observations refer to each Wednesday during January 5, 1996 to December 30, 2005. Panel A shows the number of option contracts sorted by moneyness and maturity. Panel B shows the average option prices classified by moneyness and maturity. Panel C shows the average implied volatilities sorted by moneyness and maturity. Implied volatilities are calculated using the Black & Scholes formula. T refers to the number of days to maturity while m represents the moneyness defined as the underlying index level divided by the option strike price.

Moneyness	$T\leq 50$	$50<T\leq 90$	$90<T\leq 160$	$T>160$	All
Panel A: Number of Contracts					
$m\leq 0.94$	2391	2294	1496	2650	8831
$0.94<m\leq 0.97$	2246	1393	595	656	4890
$0.97<m\leq 1$	2671	1732	637	809	5849
$1<m\leq 1.03$	2169	1160	475	633	4437
$1.03<m\leq 1.06$	1240	546	274	285	2345
$m>1.06$	1631	853	602	623	3709
All	12348	7978	4079	5656	30061
Panel B: Average Option Prices					
$m\leq 0.94$	0.9188	3.2835	6.9051	17.044	7.3858
$0.94<m\leq 0.97$	3.8515	11.450	23.346	46.767	14.145
$0.97<m\leq 1$	13.009	24.920	39.597	65.467	26.687
$1<m\leq 1.03$	29.895	42.033	55.952	81.544	43.226
$1.03<m\leq 1.06$	54.066	65.985	78.103	103.23	65.625
$m>1.06$	131.48	148.11	169.64	195.13	152.19
All	31.740	34.817	48.920	58.595	39.940
Panel C: Average Implied Volatility					
$m\leq 0.94$	0.1726	0.1555	0.1498	0.1474	0.1568
$0.94<m\leq 0.97$	0.1378	0.1410	0.1490	0.1556	0.1425
$0.97<m\leq 1$	0.1454	0.1503	0.1593	0.1611	0.1505
$1<m\leq 1.03$	0.1594	0.1619	0.1644	0.1664	0.1616
$1.03<m\leq 1.06$	0.1800	0.1785	0.1758	0.1735	0.1784
$m>1.06$	0.2762	0.2312	0.2199	0.2034	0.2445
All	0.1725	0.1624	0.1650	0.1599	0.1664

4.4 The calibration of risk premia

Given the estimates of the parameters under the $\mathbb{P}$ measure obtained via the procedure described in Section 3.1, the risk premium parameters in (18) need to be calibrated to derive the risk-neutral dynamics.

Specifically, δ_2 is determined by the no-arbitrage condition in the Proposition 3, and δ_{11} and δ_{12} are calibrated on observed option prices. The purpose of the calibration is the selection of risk premia such that the model implied unconditional volatility under the risk-neutral measure matches the unconditional risk-neutral volatility. Notice that it is not possible to directly observe the latter, then I follow the same strategy used in Corsi et al. (2013). The market-observed implied volatility (IV) is used as an instrument to be matched with the model-generated IV since both depend on the volatility under $\mathbb{Q}$ measure. The two risk premia, δ_{11} and δ_{12}, are calibrated by minimising the loss function which measures the distance between the model generated IV for two options and the market IV corresponding to the same two options in the sample:

$$f(\delta_{11}, \delta_{12}) = \sqrt{\frac{1}{2} \left[(IV_1^{mkt} - IV_1^{mod})^2 + (IV_2^{mkt} - IV_2^{mod})^2 \right]} \times 100 \qquad (24)$$

where $IV_{i,t}^{mkt}$ is the market IV of the option i, $IV_{i,t}^{mod}$ is the IV computed from the model for the same option, with $i = 1, 2$. In order to deal with the risk premia calibration for the VARG-B model, I randomly select two ATM options from the most liquid ones, observed in two different days [9].

Then, I proceed in pricing options: first I map the parameters of the model estimated under $\mathbb{P}$ into the parameters under $\mathbb{Q}$ according to Proposition 5; second I approximate option prices by the COS method introduced by Fang and Oosterlee (2008), using the

---

[9]For the alternative models, ARG and HARG, the motivation behind the calibration is the same but, since the risk premium to be calibrated is only δ_{11}, I select just one option from the sample. To avoid the problem of a possible unfair comparison among models, I randomly choose one of the two options used to minimise the objective function in (24). I also calibrate both ARG and HARG models on the latter option (the option that is not randomly selected at the beginning), and the results are comparable in terms of pricing performance.

MGF formula in Proposition 2 with the parameters in Proposition 5. Finally, I compute the IVs for each model. As expected[10], both δ_{11} and δ_{12} are negative and equal to -0,8197 and -0,2386, respectively.

4.5 Option pricing performance

As it is customary in the literature, I analyse the option pricing performance of each model in terms of the Root Mean Square Error on the percentage IV:

$$RMSE_{IV} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(IV_i^{mkt} - IV_i^{mod})^2} \times 100 \qquad (25)$$

where $IV_{i,t}^{mkt}$ is the market IV of option i, $IV_{i,t}^{mod}$ is the IV computed from the model for the same option, with $i = 1, 2, \ldots N$ and N is the total number of options in the sample. For completeness I report the performance results also for the Root Mean Square Error on option prices:

$$RMSE_P = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_i^{mkt} - P_i^{mod})^2} \qquad (26)$$

where $P_{i,t}^{mkt}$ is the market price of the option i, $P_{i,t}^{mod}$ is the price computed from the model of the option i, for $i = 1, 2, \ldots N$. The former metric represents an intuitive weighting of options across strikes and maturities. The latter gives more weight to options with high intrinsic value (DITM) and time value (longer maturity) but has the advantage of interpreting RMSE as $ errors.

Table 3 reports the global option pricing performance on the S&P500 call options from January 5, 1996 to December 30, 2005. The first row shows the absolute $RMSE_{IV}$ and the $RMSE_P$ for the VARG-B model, computed over the entire sample. The remaining rows display an indicator for the VARG-B relative performance with respect to the alternative

---

[10]Options are volatility-sensitive investments. They typically pay off in adverse states of nature, i.e. when the marginal utility of wealth is high. This means that such investment is negative-beta and, in turn, are characterised by negative risk premia.

models which is computed as the ratio between the RMSE$_{IV}$ (RMSE$_P$) of the VARG-B and that of the ARG and HARG, respectively. Note that this indicator is a ratio between two loss functions, indeed a value less than one indicates an outperformance of the model set as the numerator, i.e. VARG-B model.

A comparison between the ARG and the VARG-B model illustrates the importance of a multifactor specification of volatility in pricing options. By comparing the VARG-B with the HARG model, I can shed some light on the importance of volatility bursts in fitting medium/long-term part of the implied volatility surface, especially for at-the-money options, with respect to the long-memory feature.

Table 3: Global option pricing performance. The first row shows the implied volatility root mean square error and the price root mean square error. RMSE$_{IV}$ and RMSE$_P$ are both expressed in percentage. The second and the third rows show the RMSE$_{IV}$ and RMSE$_P$ of the competitor models relative to the VARG-B. A ratio smaller than 1 indicates an outperformance of the VARG-B model. I use the parameter estimates from Table 1 and S&P500 call options from January 5, 1996 to December 30, 2005.

Model	RMSE$_{IV}$	RMSE$_P$
VARG-B	6.2924	0.6648
VARG-B/ARG	0.8661	0.5992
VARG-B/HARG	0.9649	0.6389

At first sight, the VARG-B model outperforms all the alternatives, both via RMSE$_{IV}$ and RMSE$_P$. Specifically, looking at the RMSE$_{IV}$ (RMSE$_P$), the VARG-B model improvement is about 14% (40%) over ARG and around 4% (36%) over HARG model.

In order to get a deeper understanding of the VARG-B pricing performance, Table 4 (Table 5) reports the results in terms of RMSE$_{IV}$ (RMSE$_P$) disaggregated for different maturities and moneyness[11]. The results in Table 4 confirm that modelling the occurrence

---

[11]The results recorded in Table 5 are in line with those reported in Table 4. For this reason, I will discuss only the pricing performance in terms of RMSE$_{IV}$.

and the probability of having bursts in the volatility carries advantages over option evaluation. The VARG-B model offers a flexible volatility specification that translates in an increasing ability to capture the volatility smile.

Panel B of Table 4 compares the performance of VARG-B and ARG. It shows the advantage of providing an extraordinary burst factor in the volatility dynamics. Improvement for long maturities $T > 90$ and moneyness $0.97 < m \leq 1.03$ reaches more than 40%. For the remaining maturities and moneyness, VARG-B still improves over ARG, recording 10% - 44% smaller error.

The relative performance between VARG-B and HARG is displayed in Panel C of Table 4. In this comparison, I focus on the long-term part of the IV surface, where the persistence of the volatility process plays a fundamental role. Note that the heterogeneous structure for the volatility in the HARG model is introduced to mimic the long-memory characterising the volatility process. The advantage of volatility burst component is strong for long maturities $T > 90$ and all moneyness. For all options with the maturity greater than 90 days, the $RMSE_{IV}$ for VARG-B is smaller than that registered for the HARG. Improvement for long maturities $T > 90$ and at-the-money options (moneyness $0.97 < m \leq 1.03$) is about 40%. The improvements are systematic along both dimensions considered.

Indeed, the ability of the VARG-B model to reproduce higher level of persistence permits this more flexible model to outperform the HARG model and to improve pricing over the long-term part of the IV surface.

To summarise, the proposed VARG-B model is better able to reproduce the IV level for options along both the moneyness and the time to maturity dimensions. It improves upon both the alternative models, especially for ATM and longer maturities. The volatility burst component appears to be an important and necessary ingredient for a more accurate option pricing and modelling of the IV surface.

Table 4: Option pricing performance via the percentage implied volatility root mean square error (RMSE$_{IV}$). The Panel A shows the RMSE$_{IV}$ of the VARG-B model sorted by moneyness and maturity. Panel B shows the RMSE$_{IV}$ of the ARG model relative to the VARG-B classified by moneyness and maturity. Panel C shows the RMSE$_{IV}$ of the HARG model relative to the VARG-B sorted by moneyness and maturity. A ratio smaller than 1 indicates an outperformance of the VARG-B model. I use the parameter estimates from Table 1 and S&P500 Call options from January 5, 1996 to December 30, 2005. T refers to the number of days to maturity while m represents the moneyness defined as the underlying index level divided by the option strike price.

Moneyness	T$\leq$50	50<T$\leq$90	90<T$\leq$160	T>160
Panel A: VARG-B Implied Volatility RMSE				
m$\leq$0.94	7.1135	5.8068	4.8609	4.4883
0.94<m$\leq$0.97	5.6552	4.9842	4.2245	4.1220
0.97<m$\leq$1	4.6264	4.3465	3.7574	3.7756
1<m$\leq$1.03	4.0896	3.9320	3.6435	3.7939
1.03<m$\leq$1.06	4.6424	4.1310	3.7545	3.6637
m>1.06	15.570	9.9653	7.5613	7.5643
Panel B: VARG-B/ARG Implied Volatility RMSE				
m$\leq$0.94	1.1848	0.8495	0.6589	0.5767
0.94<m$\leq$0.97	0.8528	0.6492	0.5758	0.5651
0.97<m$\leq$1	0.7676	0.6302	0.5763	0.5557
1<m$\leq$1.03	0.7394	0.6446	0.5896	0.5826
1.03<m$\leq$1.06	0.9011	0.7679	0.6664	0.6050
m>1.06	1.1488	1.1255	0.9567	1.0451
Panel C: VARG-B/HARG Implied Volatility RMSE				
m$\leq$0.94	1.1663	1.0004	0.6908	0.5540
0.94<m$\leq$0.97	1.2042	0.8638	0.6573	0.5652
0.97<m$\leq$1	1.1133	0.8702	0.6697	0.5534
1<m$\leq$1.03	1.0999	0.9222	0.7010	0.5857
1.03<m$\leq$1.06	1.1961	1.1177	0.7995	0.6057
m>1.06	1.1392	1.2127	0.9863	1.0953

Table 5: Option pricing performance via the percentage price root mean square error (RMSE$_P$). The Panel A shows the RMSE$_P$ of the VARG-B model sorted by moneyness and maturity. Panel B shows the RMSE$_P$ of the ARG model relative to the VARG-B sorted by moneyness and maturity. Panel C shows the RMSE$_P$ of the HARG model relative to the VARG-B classified by moneyness and maturity. A ratio smaller than 1 indicates an outperformance of the VARG-B model. I use the parameter estimates from Table 1 and S&P500 Call options from January 5, 1996 to December 30, 2005. T refers to the number of days to maturity while m represents the moneyness defined as the underlying index level divided by the option strike price.

Moneyness	T≤50	50<T≤90	90<T≤160	T>160
Panel A: VARG-B Price RMSE				
m≤0.94	0.2801	0.4441	0.6308	0.9909
0.94<m≤0.97	0.4015	0.7520	0.9128	1.2694
0.97<m≤1	0.4781	0.8102	0.9014	1.2068
1<m≤1.03	0.4184	0.6788	0.8588	1.2224
1.03<m≤1.06	0.3549	0.5680	0.8026	1.0551
m>1.06	0.2423	0.3772	0.5922	0.8743
Panel B: VARG-B/ARG Price RMSE				
m≤0.94	1.3588	0.6709	0.5202	0.4926
0.94<m≤0.97	0.7974	0.6156	0.5503	0.5604
0.97<m≤1	0.7213	0.6216	0.5784	0.5790
1<m≤1.03	0.7113	0.6252	0.6185	0.6383
1.03<m≤1.06	0.8950	0.7580	0.7215	0.6523
m>1.06	0.9925	0.9943	0.9173	0.8830
Panel C: VARG-B/HARG Price RMSE				
m≤0.94	1.5242	0.8285	0.5458	0.4632
0.94<m≤0.97	1.1435	0.8287	0.6151	0.5479
0.97<m≤1	1.0532	0.8536	0.6581	0.5625
1<m≤1.03	1.0616	0.8861	0.7199	0.6295
1.03<m≤1.06	1.1979	1.1149	0.8479	0.6382
m>1.06	1.1036	1.2833	1.0844	0.8735

5 Conclusion

In this paper, I propose an affine discrete-time model, labelled VARG-B, in which volatility experiences, in addition to frequent and small changes, periods of sudden and extreme movements, i.e. volatility bursts. The former changes are generated by a state variable, called continuous component while the volatility burst component generates the volatility changes due to extreme price movements. The total volatility is equal to the sum of these two independent factors which are both assumed as latent. The continuous component is modelled as an Autoregressive Gamma (ARG) while the volatility burst factor follows an Autoregressive Gamma Zero (ARG-Zero) process.

A great advantage of VARG-B is represented by the estimation strategy which allows to filter the time series of both the volatility components and to understand the relative contribution of both the factors to the total conditional variance of log-returns. The state space is augmented with a measurement equation that relates the latent volatility to an ex-post estimator of daily quadratic variation of the log-price process, i.e. Realized Volatility.

The VARG-B model leads to a fully analytic conditional Laplace transform (that is, exponential affine), which is particularly attractive for option pricing purposes. Indeed, the change of measure is performed adopting an exponentially affine stochastic discount factor which preserves all the analytical results in order to obtain closed-form option pricing formula.

The proposed VARG-B model is better able to reproduce the IV level for options along both the moneyness and the time to maturity dimensions with respect to some alternatives. The greatest improvement is registered for D-OTM and ITM options at medium and long maturity. For these option categories, the VARG-B model outperforms both alternative models and provides an improvement also for ATM options.

The more flexible volatility specification allows the VARG-B model to reproduce higher level of persistence, which improves the pricing along the long-term part of the IV surface.

From the evidence reported in this paper, the volatility burst component is an impor-

27

tant and necessary ingredient for a more accurate option pricing and modelling of the IV surface.

The VARG-B model can be extended to include leverage effect in the volatility dynamics as well as a dependence between volatility components. I leave the possibility to study these features to future research.

A Proof of Proposition 1

The *Assumption 1* in Majewski et al. (2015) is the following:

$$\mathbb{E}^{\mathbb{P}}\left[\exp(zy_{t+1} + \mathbf{b}'\mathbf{f}_{t+1} + \mathbf{c}'\mathbf{l}_{t+1})|\mathcal{F}_t, \mathcal{L}_t\right] \tag{A.1}$$

$$= \exp\left[\mathcal{A}(z, \mathbf{b}, \mathbf{c}) + \sum_{i=1}^{p}\mathcal{B}_i(z, \mathbf{b}, \mathbf{c})' \cdot \mathbf{f}_{t+1-i} + \sum_{j=1}^{q}\mathcal{C}_j(z, \mathbf{b}, \mathbf{c})' \cdot \mathbf{l}_{t+1-j}\right]$$

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

$$\mathbb{E}^{\mathbb{P}}\left[\exp(zy_{t+1} + \mathbf{b}'\mathbf{f}_{t+1})|\mathcal{F}_t\right] \tag{A.2}$$

$$= \exp\left[\mathcal{A}(z, \mathbf{b}) + \mathcal{B}(z, \mathbf{b})' \cdot \mathbf{f}_t\right]$$

For the setup in Section 2.1, assumption (A.2) is satisfied with $\mathbf{l}_t = 0$ for $t = 1, \ldots, T$ and $p = 1$. Without loss of generality I assume $r_{t+1} = r$ for $t > 0$, since r_{t+1} is predetermined (that is, known at t).

To derive the expressions for $\mathcal{A}_t(z, \mathbf{b}, \mathbf{c})$ and $\mathcal{B}_i(z, \mathbf{b}, \mathbf{c})$, I write:

$$\mathbb{E}^{\mathbb{P}}\left[\exp(zy_{t+1} + \mathbf{b}'\mathbf{f}_{t+1})|\mathcal{F}_t\right]$$

$$= \mathbb{E}^{\mathbb{P}}\left[\exp(zr + z\lambda f_{t+1} + z\sqrt{f_{t+1}}\epsilon_{t+1} + b_1 f_{1,t+1} + b_2 f_{2,t+1})|\mathcal{F}_t\right]$$

$$= e^{zr}\mathbb{E}^{\mathbb{P}}\left[\exp[(b_1 + z\lambda)f_{1,t+1} + (b_2 + z\lambda)f_{2,t+1} + z\sqrt{f_{t+1}}\epsilon_{t+1}]|\mathcal{F}_t\right]$$

$$= e^{zr}\mathbb{E}^{\mathbb{P}}\left\{\exp\left[(b_1 + z\lambda)f_{1,t+1} + (b_2 + z\lambda)f_{2,t+1}\right]\right.$$

$$\left. \times \mathbb{E}^{\mathbb{P}}\left[\exp\left[z\sqrt{f_{t+1}}\epsilon_{t+1}\right]\Big| f_{1,t+1}, f_{2,t+1}, \mathcal{F}_t\right]\Big|\mathcal{F}_t\right\}$$

where $\mathbf{f}_{t+1} = (f_{1,t+1}, f_{2,t+1})'$.

To compute the inner expectation I now use the following property: if $Z \sim \mathcal{N}(0,1)$ and $Y = aZ$, then

$$\mathbb{E}\{\exp[xY]\} = \exp\left[\frac{1}{2}(xa)^2\right]$$

Hence:

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[\exp(zy_{t+1}+\mathbf{b}'\mathbf{f}_{t+1})|\mathcal{F}_t\right] \\
&= e^{zr}\mathbb{E}^{\mathbb{P}}\left\{\exp\left[(b_1+z\lambda)f_{1,t+1}+(b_2+z\lambda)f_{2,t+1}+\frac{z^2}{2}f_{t+1}\right]\middle|\mathcal{F}_t\right\} \\
&= e^{zr}\mathbb{E}^{\mathbb{P}}\left\{\exp\left[(b_1+z\lambda)f_{1,t+1}+(b_2+z\lambda)f_{2,t+1}+\frac{z^2}{2}f_{1,t+1}+\frac{z^2}{2}f_{2,t+1}\right]\middle|\mathcal{F}_t\right\} \\
&= e^{zr}\mathbb{E}^{\mathbb{P}}\left\{\exp\left[x_1(z,\mathbf{b})f_{1,t+1}+x_2(z,\mathbf{b})f_{2,t+1}\right]|\mathcal{F}_t\right\}
\end{aligned}
$$

where:

$$
x_1(z,\mathbf{b}) = b_1+z\lambda+\frac{z^2}{2} \tag{A.3}
$$

$$
x_2(z,\mathbf{b}) = b_2+z\lambda+\frac{z^2}{2} \tag{A.4}
$$

In what follows I will sometimes simplify the notation using x_1 (resp. x_2) instead of $x_1(z,\mathbf{b})$ (resp. $x_2(x,\mathbf{b})$). I now use the following property of the noncentral Gamma-Zero distribution: if $Z \sim \gamma_0(\theta,\mu)$, then

$$
\mathbb{E}[\exp(xZ)] = \exp\left[\frac{x\mu}{1-x\mu}\theta\right].
$$

Since $f_{2,t+1}|\mathcal{F}_t \sim \gamma_0(d_2+\beta_2 f_{2,t},\mu_2)$, defining $\theta_{2t}=d_2+\beta_2 f_{2,t}$, I get:

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[\exp(zy_{t+1}+\mathbf{b}'\mathbf{f}_{t+1})|\mathcal{F}_t\right] \\
&= e^{zr}\mathbb{E}^{\mathbb{P}}\left\{\exp[x_1(z,\mathbf{b})f_{1,t+1}]\mathbb{E}^{\mathbb{P}}\left[\exp\left(x_2(z,\mathbf{b})f_{2,t+1}\right)|f_{1,t+1},\mathcal{F}_t\right]\middle|\mathcal{F}_t\right\} \\
&= e^{zr\frac{x_2\mu_2}{1-x_2\mu_2}\theta_{2t}}\mathbb{E}^{\mathbb{P}}\left\{\exp[x_1 f_{1,t+1}]|\mathcal{F}_t\right\} \\
&= e^{zr+V_2(x_2,\mu_2)\theta_{2t}}\mathbb{E}^{\mathbb{P}}\left\{\exp[x_1 f_{1,t+1}]|\mathcal{F}_t\right\}
\end{aligned}
$$

where:

$$
V_2[x_2,\mu_2] = \frac{x_2(z,\mathbf{b})\mu_2}{1-x_2(z,\mathbf{b})\mu_2} \tag{A.5}
$$

I now use the following property of the noncentral Gamma distribution: if $Z \sim \gamma_\nu(\theta, \mu)$, then

$$\mathbb{E}[\exp(xZ)] = \exp\left[\frac{x\mu}{1-x\mu}\theta - \nu\log(1-x\mu)\right].$$

Since $f_{1,t+1}|\mathcal{F}_t \sim \gamma_\nu(\beta_1 f_{1,t}, \mu_1)$, defining $\theta_{1t} = \beta_1 f_{1,t}$, I get:

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[\exp(zy_{t+1} + \mathbf{b}'\mathbf{f}_{t+1})|\mathcal{F}_t\right] \\
&= \exp\left\{zr + V_2(x_2, \mu_2)\theta_{2t} - \nu\log(1-x_1\mu_1) + \frac{x_1\mu_1}{1-x_1\mu_1}\theta_{1t}\right\} \\
&= \exp\left\{zr - \nu W_1(x_1, \mu_1) + V_1(x_1, \mu_1)\theta_{1t} + V_2(x_2, \mu_2)\theta_{2t}\right\} \quad\quad\quad \text{(A.6)}
\end{aligned}
$$

where

$$
\begin{aligned}
W_1[x_1, \mu_1] &= \log[1 - x_1(z, \mathbf{b})\mu_1] &\text{(A.7)}\\
V_1[x_1, \mu_1] &= \frac{x_1(z, \mathbf{b})\mu_1}{1 - x_1(z, \mathbf{b})\mu_1} &\text{(A.8)}
\end{aligned}
$$

Substituting the expressions for non-centrality parameters in (A.6) and collecting terms, it is easy to check that Assumption A.2 is satisfied, with:

$$
\begin{aligned}
\mathcal{A}(z, \mathbf{b}) &= zr - \nu W_1(x_1, \mu_1) + V_2(x_2, \mu_2)d_2 &\text{(A.9)}\\
\mathcal{B}(z, \mathbf{b})' &= [V_1(x_1, \mu_1), \ V_2(x_2, \mu_2)]\beta &\text{(A.10)}
\end{aligned}
$$

where $\beta = (\beta_1, \beta_2)'$.

B   Proof of Proposition 3

The assumed SDF is

$$M_{t,t+1} = \frac{\exp(-\delta_2 y_{t+1} - \delta_{11} f_{1,t+1} - \delta_{12} f_{2,t+1})}{\mathbb{E}^{\mathbb{P}}[\exp(-\delta_2 y_{t+1} - \delta_{11} f_{1,t+1} - \delta_{12} f_{2,t+1})|\mathcal{F}_t]} \quad\quad\quad \text{(B.11)}$$

The no-arbitrage conditions are

$$\mathbb{E}^{\mathbb{P}}[M_{s,s+1}|\mathcal{F}_s] = 1 \qquad \text{for} \quad s \in \mathbb{N} \qquad \text{(B.12)}$$

$$\mathbb{E}^{\mathbb{P}}[M_{s,s+1}e^{y_{s+1}}|\mathcal{F}_s] = e^r \qquad \text{for} \quad s \in \mathbb{N} \qquad \text{(B.13)}$$

The first condition is satisfied by definition of $M_{t,t+1}$.

Let $\delta_1 = (\delta_{11}, \delta_{12})'$. To enforce no arbitrage, I use *Proposition 2* in Majewski et al. (2015), which shows that the second condition is equivalent to:

$$\mathcal{A}(1 - \delta_2, -\delta_1) = r + \mathcal{A}(-\delta_2, -\delta_1)$$

$$\mathcal{B}(1 - \delta_2, -\delta_1) = \mathcal{B}(-\delta_2, -\delta_1)$$

These equalities are implied by

$$x_1(1 - \delta_2, -\delta_1) = x_1(-\delta_2, -\delta_1)$$

$$x_2(1 - \delta_2, -\delta_1) = x_2(-\delta_2, -\delta_1).$$

For this to hold, it is easy to check that it is sufficient to impose

$$\delta_2 = \lambda + \frac{1}{2} \qquad \text{(B.14)}$$

C   Proof of Corollary 4

Let $y_{t,T} = \log(S_T/S_t)$ I have to show that:

$$\varphi^{\mathbb{Q}}_{\delta_2,\delta_1}(t, T, z) = \mathbb{E}^{\mathbb{Q}}\left(e^{z y_{t,T}}|\mathcal{F}_t\right) = \exp\left(a_t^* + \mathbf{b}_t^{*\prime}\mathbf{f}_t\right) \qquad \text{(C.15)}$$

where:

$$a_s^* = a_{s+1}^* + \mathcal{A}(z - \delta_2, \mathbf{b}_{s+1}^* - \boldsymbol{\delta_1}) - \mathcal{A}(-\delta_2, -\boldsymbol{\delta_1}) \tag{C.16}$$

$$\mathbf{b}_s^* = \mathbf{b}_{s+1}^* + \mathcal{B}(z - \delta_2, \mathbf{b}_{s+1}^* - \boldsymbol{\delta_1}) - \mathcal{B}(-\delta_2, -\boldsymbol{\delta_1}) \tag{C.17}$$

subject to the terminal conditions:

$$a_T^* = 0, \quad \mathbf{b}_T^* = \mathbf{0}$$

The above relation is derived using the expression for the SDF in (18) repeatedly and using the tower law of conditional expectation:

$$\varphi_{\delta_2,\delta_1}^{\mathbb{Q}}(t, T, z)$$

$$= \mathbb{E}^{\mathbb{Q}}\left[e^{zy_{t,T}} \middle| \mathcal{F}_t\right]$$

$$= \mathbb{E}^{\mathbb{P}}\left[M_{t,t+1} \ldots M_{T-1,T} e^{zy_{t,T}} \middle| \mathcal{F}_t\right]$$

$$= \mathbb{E}^{\mathbb{P}}\left[M_{t,t+1} \ldots M_{T-2,T-1} e^{zy_{t,T-1}} \mathbb{E}^{\mathbb{P}}\left[M_{T-1,T} e^{y_T} \middle| \mathcal{F}_{T-1}\right] \middle| \mathcal{F}_t\right]$$

$$= \mathbb{E}^{\mathbb{P}}\left[M_{t,t+1} \ldots M_{T-2,T-1} e^{zy_{t,T-1}} \mathbb{E}^{\mathbb{P}}\left[\frac{e^{-\delta_2 y_T - \delta_{11} f_{1,T} - \delta_{12} f_{2,T} + zy_T}}{\mathbb{E}^{\mathbb{P}}\left[e^{-\delta_2 y_T - \delta_{11} f_{1,T} - \delta_{12} f_{2,T}} \middle| \mathcal{F}_{T-1}\right]} \middle| \mathcal{F}_{T-1}\right] \middle| \mathcal{F}_t\right]$$

$$= \mathbb{E}^{\mathbb{P}}\left[M_{t,t+1} \ldots M_{T-2,T-1} e^{zy_{t,T-1} - \mathcal{A}(-\delta_2, -\delta_1) - \mathcal{B}(-\delta_2, -\delta_1) f_{T-1}} \mathbb{E}^{\mathbb{P}}\left[e^{(z-\delta_2)y_T - \delta_1 f_T} \middle| \mathcal{F}_{T-1}\right] \middle| \mathcal{F}_t\right]$$

$$= \mathbb{E}^{\mathbb{P}}\left[M_{t,t+1} \ldots M_{T-2,T-1} e^{zy_{t,T-1} - \mathcal{A}(-\delta_2, -\delta_1) - \mathcal{B}(-\delta_2, -\delta_1) f_{T-1} + \mathcal{A}(z-\delta_2, -\delta_1) + \mathcal{B}(z-\delta_2, -\delta_1) f_{T-1}} \middle| \mathcal{F}_t\right]$$

$$= \mathbb{E}^{\mathbb{P}}\left[M_{t,t+1} \ldots M_{T-2,T-1} e^{zy_{t,T-1} + a_{T-1}^* + b_{T-1}^* f_{T-1}} \middle| \mathcal{F}_t\right]$$

$$= \mathbb{E}^{\mathbb{P}}\left[M_{t,t+1} \ldots M_{T-3,T-2} e^{zy_{t,T-2} + a_{T-1}^*} \mathbb{E}^{\mathbb{P}}\left[M_{T-2,T-1} e^{zy_{T-1} + b_{T-1}^* f_{T-1}} \middle| \mathcal{F}_{T-2}\right] \middle| \mathcal{F}_t\right]$$

$$= \mathbb{E}^{\mathbb{P}}\left[M_{t,t+1} \ldots M_{T-3,T-2} e^{zy_{t,T-2} + a_{T-1}^* - \mathcal{A}(-\delta_2, -\delta_1) - \mathcal{B}(-\delta_2, -\delta_1) f_{T-2} + \mathcal{A}(z-\delta_2, b_{T-1}^* - \delta_1) + \mathcal{B}(z-\delta_2, b_{T-1}^* - \delta_1) f_{T-2}} \middle| \mathcal{F}_t\right]$$

$$= \ldots$$

$$= e^{(a_t^* + \mathbf{b}_t^{*\prime} \mathbf{f}_t)}$$

I now specialize these expression for the setup outlined in Section 2.1. Consider equation (C.16). Using (A.9), I get

$$a_s^* = a_{s+1}^* + zr - \nu(W_{1,s+1}^* - W_1^y) + d_2(V_{2,s+1}^* - V_2^y) \tag{C.18}$$

where:

$$
\begin{aligned}
x_{h,s+1}^* &= x_h(z - \delta_2, b_{s+1}^* - \delta_1), \quad h = 1,2 \\
y_h^* &= x_h(-\delta_2, -\delta_1) = -\delta_{1h} - \delta_2\lambda + \frac{\delta_2^2}{2}, \quad h = 1,2 \\
V_{h,s+1}^* &= V_h(x_{h,s+1}^*, \mu_h), \quad h = 1,2 \\
V_h^y &= V_h(y_h^*, \mu_h), \quad h = 1,2 \\
W_{1,s+1}^* &= W_1(x_{1,s+1}^*, \mu_1) \\
W_1^y &= W_1(y_1^*, \mu_1)
\end{aligned}
$$

Using (A.10), equation (C.17) becomes:

$$\mathbf{b}_s^{*\prime} = \mathbf{b}_{s+1}^{*\prime} + (V_{1,s+1}^* - V_1^y, \ V_{2,s+1}^* - V_2^y)\beta \tag{C.19}$$

D Proof of Proposition 2

To compute the MGF of $y_{t,T}$ under $\mathbb{P}$, I simply need to plug $\delta_2 = 0$ and $\delta_1 = \mathbf{0}$ in the expression of $\varphi_{\delta_2,\delta_1}^{\mathbb{Q}}(t, T, z)$ in (C.15):

$$\varphi_{0,0}^{\mathbb{Q}}(t, T, z) = \mathbb{E}^{\mathbb{P}}[e^{zy_{t,T}}|\mathcal{F}_t] = exp\left(a_t + \mathbf{b}_t'\mathbf{f}_t\right) \tag{D.20}$$

where

$$
\begin{aligned}
a_s &= a_{s+1} + zr - \nu W_{1,s+1} + d_2 V_{2,s+1} \tag{D.21} \\
\mathbf{b}_s' &= \mathbf{b}_{s+1}' + (V_{1,s+1}, V_{2,s+1})\beta \tag{D.22}
\end{aligned}
$$

with

$$x_{h,s+1} = x_h(z, \mathbf{b}_{s+1}) = \mathbf{b}_{s+1} + z\lambda + \frac{z^2}{2}, \quad h = 1,2$$

and

$$
\begin{aligned}
V_{h,s+1} &= V_h(x_{h,s+1}, \mu_h) = \frac{x_{h,s+1}\mu_h}{1 - x_{h,s+1}\mu_h}, \quad h = 1,2 \\
W_{1,s+1} &= W_1(x_{1,s+1}, \mu_1) = \log(1 - x_{1,s+1}\mu_1)
\end{aligned}
$$

subject to the initial conditions:

$$a_T = 0, \quad \mathbf{b}'_T = 0$$

E Proof of Proposition 5

The MGFs $\varphi^{\mathbb{Q}}_{\delta_2,\delta_1}(t, T, z)$ and $\varphi^{\mathbb{P}}_{0,0}(t, T, z)$ derived above depend on the parameters under $\mathbb{P}$, $\boldsymbol{\psi}$ defined in (17), and on the risk premium parameters $\delta = (\delta_2, \delta'_1)'$ introduced in the SDF (18). I now show that the MGF under $\mathbb{Q}$ can be rewritten as the MGF under $\mathbb{P}$ using a new set of parameters $\boldsymbol{\psi}^{\mathbb{Q}}$ i.e. the risk-neutral ones

$$\boldsymbol{\psi}^{\mathbb{Q}} = [\lambda^{\mathbb{Q}}, \nu^{\mathbb{Q}}, \mu_1^{\mathbb{Q}}, \mu_2^{\mathbb{Q}}, d_2^{\mathbb{Q}}, \beta^{\mathbb{Q}}] \tag{E.23}$$

To derive the expression of $\boldsymbol{\psi}^{\mathbb{Q}}$ as a function of $\boldsymbol{\psi}$ and δ, I match the parameters using the identity:

$$\varphi^{\mathbb{Q}}_{\delta_2,\delta_1}(t, T, z; \boldsymbol{\psi}, \delta) = \varphi^{\mathbb{P}}_{0,0}(t, T, z; \boldsymbol{\psi}^{\mathbb{Q}}) \tag{E.24}$$

It is useful to denote

$$
\begin{aligned}
x^{\mathbb{Q}}_{h,s+1} &= x_h(z, \boldsymbol{b}^*_{s+1}; \boldsymbol{\psi}^{\mathbb{Q}}), \quad h = 1,2 \\
V^{\mathbb{Q}}_{h,s+1} &= V_h(x^{\mathbb{Q}}_{h,s+1}, \mu_h^{\mathbb{Q}}), \quad h = 1,2 \\
W^{\mathbb{Q}}_{1,s+1} &= W_1(x^{\mathbb{Q}}_{1,s+1}, \mu_1^{\mathbb{Q}})
\end{aligned}
$$

For (E.24) to hold, (C.18) needs to be matched with (15) and (C.19) with (16) , where (15) and (16) are evaluated at $V^Q_{1,s+1}$, $V^Q_{2,s+1}$ and $W^Q_{1,s+1}$. Note that since I start from the same initial conditions (E.24) requires

$$\nu(W^*_{1,s+1} - W^y_1) = \nu^Q W^Q_{1,s+1} \tag{E.25}$$

$$d_2(V^*_{2,s+1} - V^y_2) = d^Q_2 V^Q_{2,s+1} \tag{E.26}$$

$$(V^*_{1,s+1} - V^y_1 , V^*_{2,s+1} - V^y_2)\beta = (V^Q_{1,s+1} , V^Q_{2,s+1})\beta^Q \tag{E.27}$$

for all s.

Consider (E.25). This requires

$$\nu[\log(1 - x^*_{1,s+1}\mu_1) - \log(1 - y^*_1\mu_1)] = \nu^Q \log(1 - x^Q_{1,s+1}\mu^Q_1)$$

Sufficient conditions for this equality to hold are

$$\nu^Q = \nu, \quad \mu^Q_1 = \frac{\mu_1}{1 - y^*_1\mu_1} \quad \text{and} \quad x^Q_{1,s+1} = x^*_{1,s+1} - y^*_1.$$

In turn, it can be checked that the latter equality is valid if I pose

$$\lambda^Q = -\frac{1}{2}.$$

Note that under these conditions I also have $x^Q_{2,s+1} = x^*_{2,s+1} - y^*_2$.

Now turn to (E.26):

$$d_2\left(\frac{x^*_{2,s+1}\mu_2}{1 - x^*_{2,s+1}\mu_2} - \frac{y^*_2\mu_2}{1 - y^*_2\mu_2}\right) = d^Q_2 \frac{x^Q_{2,s+1}\mu^Q_2}{1 - x^Q_{2,s+1}\mu^Q_2}.$$

If I substitute for $x^Q_{2,s+1}$ the expression obtained above, I get:

$$d^Q_2 = \frac{d_2}{1 - y^*_2\mu_2}$$

Its validity is guaranteed if μ_2^Q is:

$$\mu_2^Q = \frac{\mu_2}{1 - y_2^* \mu_2}$$

Note that these solutions also imply that $V_{2,s+1}^Q = (1 - y_2^* \mu_2)(V_{2,s+1}^* - V_2^y)$

Finally, I turn to (E.27) which implies:

$$\beta_h^Q = \frac{\beta_h}{1 - y_h^* \mu_h}, \quad h = 1, 2$$

Note that under this condition $V_{1,s+1}^Q = (1 - y_1^* \mu_1)(V_{1,s+1}^* - V_1^y)$.

References

Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In *Selected Papers of Hirotugu Akaike*, pp. 199–213. Springer.

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

Andersen, T. G. and T. Bollerslev (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International economic review*, 885–905.

Andersen, T. G., T. Bollerslev, and F. X. Diebold (2007). Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *The Review of Economics and Statistics 89*(4), 701–720.

Bakshi, G., C. Cao, and Z. Chen (1997). Empirical performance of alternative option pricing models. *The Journal of finance 52*(5), 2003–2049.

Barone-Adesi, G., R. F. Engle, and L. Mancini (2008). A GARCH option pricing model with filtered historical simulation. *Review of Financial Studies 21*(3), 1223–1258.

Biais, B., L. Glosten, and C. Spatt (2005). Market microstructure: A survey of microfoundations, empirical results, and policy implications. *Journal of Financial Markets 8*(2), 217–264.

Caporin, M., E. Rossi, and P. S. de Magistris (2017). Chasing volatility: A persistent multiplicative error model with jumps. *Journal of Econometrics 198*(1), 122–145.

Caporin, M., E. Rossi, P. S. de Magistris, et al. (2015). Volatility jumps and their economic determinants. *Journal of Financial Econometrics 14*(1), 29–80.

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

Christoffersen, P., S. Heston, and K. Jacobs (2013). Capturing option anomalies with a variance-dependent pricing kernel. *Review of Financial Studies 26*(8), 1963–2006.

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

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, nbp001.

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

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

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

Engle, R. F. and G. M. Gallo (2006). A multiple indicators model for volatility using intra-daily data. *Journal of Econometrics 131*(1), 3–27.

Fang, F. and C. W. Oosterlee (2008). A novel pricing method for european options based on fourier-cosine series expansions. *SIAM Journal on Scientific Computing 31*(2), 826–848.

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

Gouriéroux, C. and J. Jasiak (2006). Autoregressive gamma processes. *Journal of Forecasting 25*(2), 129–152.

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

Madhavan, A. (2000). Market microstructure: A survey. *Journal of financial markets 3*(3), 205–258.

Majewski, A. A., G. Bormetti, and F. Corsi (2015). Smile from the past: A general option pricing framework with multiple volatility and leverage components. *Journal of Econometrics*.

McAleer, M. and M. C. Medeiros (2008). Realized volatility: A review. *Econometric Reviews 27*(1-3), 10–45.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of financial economics 3*(1-2), 125–144.

Monfort, A., F. Pegoraro, J.-P. Renne, and G. Roussellet (2014). Recursive discrete-time affine processes and asset pricing. *Technical report, mimeo.*.

Schwarz, G. et al. (1978). Estimating the dimension of a model. *The annals of statistics 6*(2), 461–464.