Journal of Econometrics 187 (2015) 593–605

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

COMFORT: A common market factor non-Gaussian returns model[☆]

Marc S. Paolella[a,c,*], Paweł Polak[a,b]

[a] Department of Banking and Finance, University of Zurich, Switzerland

[b] Department of Statistics, Columbia University, NY, United States

[c] Swiss Finance Institute, Switzerland

ARTICLE INFO

Article history:
Available online 12 March 2015

JEL classification:
C51
C53
G11
G17

Keywords:
CCC
Common jumps
Density forecasting
EM-algorithm
Fat tails
GARCH
Multivariate asymmetric variance gamma distribution
Multivariate generalized hyperbolic distribution
Multivariate option pricing
Stochastic volatility

ABSTRACT

A new multivariate time series model with various attractive properties is motivated and studied. By extending the CCC model in several ways, it allows for all the primary stylized facts of financial asset returns, including volatility clustering, non-normality (excess kurtosis and asymmetry), and also dynamics in the dependency between assets over time. A fast EM-algorithm is developed for estimation. Each element of the vector return at time t is endowed with a common univariate shock, interpretable as a common market factor. This leads to the new model being a hybrid of GARCH and stochastic volatility, but without the estimation problems associated with the latter, and being applicable in the multivariate setting for potentially large numbers of assets. A feasible technique which allows for multivariate option pricing is presented, along with an empirical illustration.

1. Introduction

We consider modeling a set of asset returns via a conditional multivariate distribution with dynamics governed by a process which has features of both GARCH and stochastic volatility (SV). These two essentially disparate paradigms for capturing volatility clustering in asset returns have their individual advantages, and also limitations. In univariate models, both can capture changes in volatility, the leverage effect (at least for more recent SV models), and the leptokurtosis and possible asymmetry of the innovations distribution. The main disadvantage of SV models is the lack of an explicit form of the likelihood function and the necessity to use moment-based methods or simulation for estimation. The former are often simple, but inefficient, while the latter attempt to achieve a close approximation of the likelihood function through computationally intensive methods, and become problematic for more than a small number of assets (see, e.g., Asai et al. 2006; Asai and McAleer 2009; and Bos 2012; and the reference therein). Alternatively, in the univariate case, a GARCH model is trivial to estimate via (conditional) maximum likelihood, but it assumes that the volatility process is predictable. Both SV and many kinds of GARCH models share the same estimation infeasibility in the large-scale multivariate case, though models such as CCC, DCC, and some of their extensions, are feasible.

The recent literature has emphasized the importance of including a stochastic jump component in the volatility dynamics—a feature which can be easily incorporated into the SV structure (see, among others, Chernov et al. 2003; Eraker et al. 2003; Eraker 2004; and Todorov and Tauchen 2011) but is absent in GARCH models. (One exception is the model proposed by Chan and Maheu 2002, and Maheu and McCurdy 2004, see also Section 2.) Another

☆ We wish to thank Tim Bollerslev, Cathy Chen, Richard Davis, Dick van Dijk, Paul Embrechts, Matthias Fengler, Eric Jondeau, Loriano Mancini, Michael McAleer, Antonio Mele, Stefan Mittnik, Olivier Scaillet, Enrique Sentana, Mike Ka Pui So, Fabio Trojani, Michael Wolf, Philip Yu, and two anonymous referees, for many useful comments and suggestions. Financial support by the Swiss National Science Foundation (SNSF)– through project #150277 for Paolella and Early Postdoc.Mobility grant #151631 for Polak – is gratefully acknowledged.

* Corresponding author at: Department of Banking and Finance, University of Zurich, Switzerland.

E-mail address: marc.paolella@bf.uzh.ch (M.S. Paolella).

http://dx.doi.org/10.1016/j.jeconom.2015.02.041
0304-4076/

approach is to use high-frequency returns in order to construct realized variation measures which separate the volatility jump component from the smooth continuous movement of the underlying volatility (see, e.g., Andersen et al. 2007).

In the multivariate case, jumps in individual assets can be observed to arrive simultaneously, forming what are called co-jumps. Following the arbitrage pricing theory, Bollerslev et al. (2008) distinguish the co-jumps in the idiosyncratic component, which are diversifiable, and the co-jumps in the common component, which are non-diversifiable, i.e., those which carry a risk premium. According to their statistical test, there are co-jumps which can be highly significant, but when considered on individual stocks, remain undetected. The importance of co-jumps structures for the consumption–portfolio selection problem is discussed by Aït-Sahalia et al. (2009). More recently, Gilder et al. (2014) provide empirical evidence for co-jumps and analyze the association between jumps in the market portfolio and co-jumps in the underlying stocks. Their results suggest that news events which have market-level influence are able to generate large co-jumps in individual stocks, while Bollerslev et al. (2013) show that extreme joint dependencies observed in daily data can be implied by the diffusive volatility *and* co-jumps observed in the high-frequency data.

In line with this high-frequency literature, we propose a new model which also splits the dynamics of volatility; however, it is a parametric approach which does not require high-frequency data and is applicable in the multivariate setting. We propose a solution which, in a conditional setup, utilizes a flexible, fat-tailed distribution, and combines univariate GARCH-type dynamics (most of the popular variations are possible) with a relatively simple, yet flexible, SV dynamic structure, based on the seminal work of Taylor (1982). By introducing a latent component similar to that used in the SV literature, the resulting hybrid GARCH-SV model is able to capture stochastic (co-)jumps in the volatility series and across assets. To the best of our knowledge, this is the first volatility model which combines GARCH and SV paradigms—and is applicable to large-dimensional multivariate settings, owing to the proposed EM-algorithm for maximum likelihood estimation (MLE) and the possibility for parallel computing.

The model contains a univariate, latent component which is common to all K assets. We term this a common market factor. It dictates the nonlinear dependency between margins, so that, conditional upon it (as realized via the EM-algorithm), what remains to be estimated are K univariate normal GARCH models. The (conditional) MLE of each of the latter is numerically fast and reliable to obtain, and the K estimations can be conducted in parallel.

Another important feature of our model is that it imposes a multiplicative structure on the volatility, and an infinitely divisible distribution which, in the iid case, generates an infinite activity jump process, as opposed to jump diffusion models, where finite activity jumps, modeled via a Poisson-distributed term, are added to Gaussian dynamics. This is in line with Aït-Sahalia and Jacod (2011), who propose two non-parametric statistical tests to discriminate between the two cases, and both tests point toward the presence of infinite-activity jumps in the data. Other examples of models which support infinite activity jumps, and are often used in the context of option pricing, are the variance gamma model of Madan and Seneta (1990) and Madan et al. (1998), the hyperbolic model of Eberlein and Keller (1995), and the CGMY model of Carr et al. (2002).

Motivated by these models from continuous time finance, we apply our model in option pricing. In order to maintain the applicability of the model for multivariate option pricing, we develop a new pricing algorithm which combines the equivalent martingale measure approach with Monte Carlo simulation. The GARCH version of the model is shown to be a good candidate for pricing options with maturity over one month. The hybrid GARCH-SV extension is flexible enough to price derivatives which are closer to expiration.

The remainder of the paper is as follows. The model and some of its properties are stated in Section 2. Section 3 discusses the proposed method of estimation. Section 4 provides the theoretical details on option pricing, while Section 5 demonstrates its empirical performance. Section 6 provides some concluding remarks, and an Appendix gathers various technical results.

2. Model

We consider a set of K financial assets, with associated return vector at time t given by $\mathbf{Y}_t = (Y_{t,1}, Y_{t,2}, \ldots, Y_{t,K})'$, $t = 1, 2, \ldots, T$, whose conditional, time-varying distribution is taken to be multivariate generalized hyperbolic, hereafter MGHyp; see, e.g., McNeil et al. (2005). We observe a realization of $\mathbf{Y} = \left[\mathbf{Y}_1 \mid \mathbf{Y}_2 \mid \cdots \mid \mathbf{Y}_T\right]$, where the $\mathbf{Y}_t$ are equally spaced in time (ignoring the weekend effect for daily data) return vectors. The information set at time t is currently defined as the history of returns, $\boldsymbol{\Phi}_t = \{\mathbf{Y}_1, \ldots, \mathbf{Y}_t\}$, though extensions to the model which include the use of exogenous variables could be straightforwardly entertained. The dispersion matrix of $\mathbf{Y}_t \mid \boldsymbol{\Phi}_{t-1}$ is decomposed as the product of scale terms and a conditional dependency matrix (a correlation matrix when the MGHyp distribution approaches the multivariate normal). For each of the univariate scales, a GARCH-type structure is imposed, while the dependency matrix is specified as being constant over time. We will see in Section 2, Remark (ii) that, except for some special cases, the correlations are actually time-varying. Further methods of inducing time-variation in the dependency matrix are discussed in the conclusions.

Using the mixture representation of the MGHyp (see Section 3; Eberlein and Keller, 1995; and Eberlein et al., 1998), we can express the return vector as

$$\mathbf{Y}_t = \boldsymbol{\mu} + \boldsymbol{\gamma} G_t + \boldsymbol{\varepsilon}_t, \quad \text{with} \tag{1a}$$

$$\boldsymbol{\varepsilon}_t = \mathbf{H}_t^{1/2} \sqrt{G_t} \mathbf{Z}_t, \tag{1b}$$

where $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_K)'$ and $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_K)'$ are column vectors in $\mathbb{R}^K$; $\mathbf{H}_t$ is a positive definite, symmetric, conditional dispersion matrix of order K; $\mathbf{Z}_t \overset{\text{iid}}{\sim} \mathrm{N}(\mathbf{0}, \mathbf{I}_K)$ is a sequence of independent and identically distributed (iid) normal random vectors and $(G_t \mid \boldsymbol{\Phi}_{t-1}) \sim \mathrm{GIG}(\lambda_t, \chi_t, \psi_t)$ are mixing random variables, $t = 1, 2, \ldots, T$, independent of $\mathbf{Z}_t$, with typical GIG (generalized inverse Gaussian) density given by

$$f_G(x; \lambda, \chi, \psi) = \frac{\chi^{-\lambda}\left(\sqrt{\chi\psi}\right)^{\lambda}}{2\mathcal{K}_\lambda\left(\sqrt{\chi\psi}\right)} x^{\lambda-1} \exp\left(-\frac{1}{2}\left(\chi x^{-1} + \psi x\right)\right), \quad x > 0; \tag{2}$$

$\mathcal{K}_\lambda$ is the modified Bessel function of the third kind (and not to be confused with K, the number of assets), given by

$$\mathcal{K}_\lambda(x) = \frac{1}{2}\int_0^\infty t^{\lambda-1} \exp\left(-\frac{x}{2}\left(t + t^{-1}\right)\right) \mathrm{d}t, \quad x > 0; \tag{3}$$

and $\chi_t > 0$, $\psi_t \geq 0$ if $\lambda_t < 0$; $\chi_t > 0$, $\psi_t > 0$ if $\lambda_t = 0$; and $\chi_t \geq 0$, $\psi_t > 0$ if $\lambda_t > 0$.

There is a minor identification problem which needs to be addressed. The same MGHyp distribution arises from the parameter constellation $(\lambda_t, \chi_t/c, c\psi_t, \boldsymbol{\mu}, c\mathbf{H}_t, c\boldsymbol{\gamma})$ for any $c > 0$. Therefore, to achieve the identification one has to pin down the scale of either G_t (see, e.g., Protassov, 2004) or one of the elements of $\mathbf{H}_t$ (see, e.g., McNeil et al., 2005). We suggest to follow the former approach and fix χ or ψ (or both) parameters in the GIG distribution prior to the estimation.

We consider two specifications for the GIG parameters: (i) $G_t \mid \boldsymbol{\Phi}_{t-1}$ are iid with time-invariant parameters, i.e., $\lambda_t = \lambda$, $\chi_t = \chi$ and $\psi_t = \psi$; (ii) $G_t \mid \boldsymbol{\Phi}_{t-1}$ has time dependent, random, parameters with the dynamics described by a system of conditional moment equations

$$\mathbb{E}\left[G_t^r \mid \boldsymbol{\Phi}_{t-1}\right] = c_r + \rho_r \mathbb{E}\left[G_{t-1}^r \mid \boldsymbol{\Phi}_{t-2}\right] + \zeta_{r,t}, \tag{4}$$

for a set of positive integer values of r; $\zeta_{r,t} = \mathbb{E}\left[G_t^r \mid \boldsymbol{\Phi}_t\right] - \mathbb{E}\left[G_t^r \mid \boldsymbol{\Phi}_{t-1}\right]$; and c_r and ρ_r are parameters to be estimated. The dynamics in (4) are of the form of the first order autoregressive process and capture the possible persistence in the G_t. The error term $\zeta_{r,t}$ represents the unpredictable component affecting the rth moment of the mixing variable G_t. It contains all new information in forming expectations about G_t^r when moving from time $t-1$ to t. It can be used as a driver of the dynamics of $\mathbb{E}\left[G_t^r \mid \boldsymbol{\Phi}_{t-1}\right]$ in (4) because it is a martingale difference sequence (MDS) with respect to $\boldsymbol{\Phi}_{t-1}$, implying that $\mathbb{E}\left[\zeta_{r,t}\right] = 0$ and $\mathrm{Cov}\left(\zeta_{r,t}, \zeta_{r,t-s}\right) = 0$, $s = 1, 2, \ldots$. From (4), the dynamics of the parameters λ_t, χ_t, and ψ_t can be inferred by the expression for the moments of the GIG random variable given below in (43). However, for the special case of the MGHyp distribution we entertain, it turns out that the dynamics of only the λ parameter associated with G_t needs to be modeled. Therefore, we limit our consideration to the $r = 1$ case, for which we need to ensure that in the estimation, $\mathbb{E}\left[G_t \mid \boldsymbol{\Phi}_{t-1}\right]$ is positive. The dynamics in (4) can be rewritten as

$$\mathbb{E}\left[G_t^r \mid \boldsymbol{\Phi}_{t-1}\right] = \frac{c_r}{2} + \frac{\rho_r}{2}\mathbb{E}\left[G_{t-1}^r \mid \boldsymbol{\Phi}_{t-2}\right] + \frac{1}{2}\mathbb{E}\left[G_t^r \mid \boldsymbol{\Phi}_t\right], \tag{5}$$

so that the sufficient condition for $\mathbb{E}\left[G_t^r \mid \boldsymbol{\Phi}_{t-1}\right]$ to be positive for all $t > 0$ is $c_r > 0$ and $\rho_r \geq 0$.

When the parameters of G_t are allowed to be dynamic as in (4), we call model (1) a hybrid GARCH-SV extension because it can be linked to the seminal SV model of Taylor (1982); this link being detailed in Appendix B. Our model differs from it because ours is a multivariate model with GARCH dynamics in the individual scales and an SV component which describes the dynamics of the common market factor G_t (as detailed below). Moreover, the dynamics in (4) are in the same vein as Chan and Maheu (2002) and Maheu and McCurdy (2004), who model dynamics of the jump intensity of a Poisson process in individual stock returns. (One important difference is that our dynamics imply that G_t^r is *not* a deterministic function of the past returns.) In line with these works, we model the dynamics of the linear projections of G_t^r on past returns only, this being another difference between our approach and that of Taylor (1982).

Due to the MDS property of the $\zeta_{r,t}$ innovations, the conditional forecasts of the future conditional moments are given by

$$\mathbb{E}\left[G_{t+s}^r \mid \boldsymbol{\Phi}_t\right] = c_r \sum_{i=0}^{s-1} \rho_r^i + \rho_r^s\, \mathbb{E}\left[G_t^r \mid \boldsymbol{\Phi}_{t-1}\right], \quad s \geq 1, \tag{6}$$

where $\mathbb{E}\left[G_t^r \mid \boldsymbol{\Phi}_{t-1}\right]$ is measurable with respect to the information up to time $t-1$ and is given by (43). If $|\rho_r| < 1$, then the process in (4) is mean-reverting, and for $s \to \infty$, the forecast approaches the unconditional mean value $c_r/\left(1-\rho_r\right)$ of G_t^r.

The conditional dispersion matrix $\mathbf{H}_t$ is decomposed as

$$\mathbf{H}_t \equiv \mathbf{S}_t \boldsymbol{\Gamma} \mathbf{S}_t, \tag{7}$$

where $\mathbf{S}_t$ is a diagonal matrix composed of the strictly positive conditional scale terms $s_{k,t}$, $k = 1, 2, \ldots, K$, and $\boldsymbol{\Gamma}$ is a time-invariant, symmetric, with ones on the main diagonal, conditional dependency matrix, such that $\mathbf{H}_t$ is positive definite. The univariate scale terms $s_{k,t}$ are modeled by a GARCH-type process. In particular, the simplest realistic choice is the GARCH(1, 1) model

$$s_{k,t}^2 = \omega_k + \alpha_k \varepsilon_{k,t-1}^2 + \beta_k s_{k,t-1}^2, \tag{8}$$

where $\varepsilon_{k,t} = y_{k,t} - \mu_k - \gamma_k G_t$ is the kth element of the $\boldsymbol{\varepsilon}_t$ vector in (1), and $\omega_k > 0, \alpha_k \geq 0, \beta_k \geq 0$, for $k = 1, 2, \ldots, K$.

Remarks. (i) In order to maintain the news effect in future volatilities, the innovation process used in the GARCH recursions in (8) is $\varepsilon_{k,t} = y_{k,t} - \mu_k - \gamma_k G_t$. If instead, we were to use $\varepsilon_{k,t}/\sqrt{G_t}$, then the next period volatility would not be influenced by the current spike in G_t. Hence, there would be no volatility persistence after news; the model would not capture the stylized fact of volatility clustering, and use of the GARCH-type dynamics for the scale term would be inadequate.

(ii) In model (1), $\boldsymbol{\mu}$ is the location vector and $\mathbf{H}_t$ is the dispersion matrix of the conditional distribution of $\mathbf{Y}_t$, while the mean and the covariance matrix are given by

$$\mathbb{E}\left[\mathbf{Y}_t \mid \boldsymbol{\Phi}_{t-1}\right] = \boldsymbol{\mu} + \mathbb{E}\left[G_t \mid \boldsymbol{\Phi}_{t-1}\right]\boldsymbol{\gamma} \tag{9}$$

and

$$\mathrm{Cov}\left(\mathbf{Y}_t \mid \boldsymbol{\Phi}_{t-1}\right) = \mathbb{E}\left[G_t \mid \boldsymbol{\Phi}_{t-1}\right]\mathbf{H}_t + \mathbb{V}\left(G_t \mid \boldsymbol{\Phi}_{t-1}\right)\boldsymbol{\gamma}\boldsymbol{\gamma}', \tag{10}$$

respectively, where $\mathbb{V}\left(G_t \mid \boldsymbol{\Phi}_{t-1}\right) = \mathbb{E}\left[G_t^2 \mid \boldsymbol{\Phi}_{t-1}\right] - \left(\mathbb{E}\left[G_t \mid \boldsymbol{\Phi}_{t-1}\right]\right)^2$. Analogously, $\boldsymbol{\Gamma}$ is a correlation matrix only conditionally on the realization of the mixing process. For this reason, we call $\boldsymbol{\Gamma}$ the dependency matrix.

While $\boldsymbol{\Gamma}$ in (7) is not dynamic, the conditional correlation matrix of $\mathbf{Y}_t \mid \boldsymbol{\Phi}_{t-1}$ is time-varying when $\boldsymbol{\gamma} \neq \mathbf{0}$ and $\mathbb{E}\left[G_t \mid \boldsymbol{\Phi}_{t-1}\right] \neq \mathbb{V}\left(G_t \mid \boldsymbol{\Phi}_{t-1}\right)$. If $\boldsymbol{\gamma} = \mathbf{0}$ or $\mathbb{E}\left[G_t \mid \boldsymbol{\Phi}_{t-1}\right] = \mathbb{V}\left(G_t \mid \boldsymbol{\Phi}_{t-1}\right)$ (e.g., in the MAVG distribution below), then $\mathrm{Corr}\left(\mathbf{Y}_t \mid \boldsymbol{\Phi}_{t-1}\right) = \boldsymbol{\Gamma}$, or $\mathrm{Corr}\left(\mathbf{Y}_t \mid \boldsymbol{\Phi}_{t-1}\right) = \boldsymbol{\Gamma} + \boldsymbol{\gamma}\boldsymbol{\gamma}'$, respectively, and the dynamics in the parameters of $G_t \mid \boldsymbol{\Phi}_{t-1}$ in (4) influence only the variances.

All the conditional moments implied by the model (including the limiting cases of the mixing law) are available in Scott et al. (2011). The co-skewness and co-kurtosis matrices are also tractable; see the expressions given in Mencía and Sentana (2009).

Finally, the unconditional mean and covariance of $\mathbf{Y}_t$ can be expressed in terms of the unconditional mean of G_t and unconditional covariance function of $\mathbf{Y}_t \mid G_t$, respectively as $\mathbb{E}\left[\mathbf{Y}_t\right] = \boldsymbol{\mu} + \mathbb{E}\left[G_t\right]\boldsymbol{\gamma}$ and $\mathrm{Cov}\left(\mathbf{Y}_t\right) = \mathbb{E}\left[\mathrm{Cov}\left(\mathbf{Y}_t \mid G_t\right)\right] + \mathbb{V}\left(G_t\right)\boldsymbol{\gamma}\boldsymbol{\gamma}'$.

(iii) From (10) it follows that the vector of conditional volatilities, defined as the square root of the conditional variances, and denoted by $\mathrm{vol}_{t|t-1}$, is given by

$$\mathrm{vol}_{t|t-1} = \sqrt{\left(\mathbb{E}\left[G_t \mid \boldsymbol{\Phi}_{t-1}\right]\mathbf{S}_t^2 + \mathbb{V}\left(G_t \mid \boldsymbol{\Phi}_{t-1}\right)\boldsymbol{\gamma}^2\right)}. \tag{11}$$

Note that, when the G_t are iid, the volatility persistence is captured by $\boldsymbol{\beta} = \left(\beta_1, \ldots, \beta_K\right)'$ from the scale term dynamics, while in the hybrid GARCH(1, 1)-SV extension, there are two sources of volatility persistence. Consider a simple example, a univariate ($K = 1$), symmetric ($\gamma = 0$), hybrid GARCH(1, 1)-SV model with $r = 1$. From (11), the next period volatility is given by

$$\begin{aligned}
\mathrm{vol}_{t+1|t}^2 &= \mathbb{E}\left[G_{t+1} \mid \boldsymbol{\Phi}_t\right] s_{t+1}^2 \\
&= \left(c_1 + \rho_1 \mathbb{E}\left[G_t \mid \boldsymbol{\Phi}_{t-1}\right] + \zeta_{1,t+1}\right)\left(\omega + \alpha\varepsilon_t^2 + \beta s_t^2\right) \\
&= c_1\omega + \left(\frac{\omega}{s_t^2}\rho_1 + \frac{c_1}{\mathbb{E}\left[G_t \mid \boldsymbol{\Phi}_{t-1}\right]}\beta + \beta\rho_1\right)\mathrm{vol}_{t|t-1}^2 + \cdots.
\end{aligned}$$

Hence, the volatility persistence is a sum of three terms, and even if the conditional moment of G_t is not persistent at all ($\rho_1 = 0$), the volatility persistence is non-zero. It is equal to $\beta c_1/\left(c_1 + \zeta_{1,t}\right)$, where $\mathbb{E}\left[\zeta_{1,t} \mid \boldsymbol{\Phi}_{t-1}\right] = 0$ because $\zeta_{1,t}$ is a martingale difference sequence.

(iv) As the volatility shock of the SV component is univariate, it influences (in a multiplicative way) each of the asset-specific conditional volatilities via (1b). Moreover, it drives the dynamics of higher conditional moments (e.g., skewness and kurtosis) and co-moments of the returns. One could argue that modeling volatility shocks with a univariate process is not sufficient because the reaction of the asset-specific volatilities to the common shock should vary across assets and through time. Nevertheless, because of the asset-specific conditional asymmetry coefficient γ_k, the impact of the SV component on each volatility is not equal across assets and its conditional expected value varies over time. (An empirical demonstration of this is given in the companion paper (Paolella and Polak, 2015b), which emphasizes out-of-sample density forecasting performance of the model, and applications to minimum expected shortfall portfolio optimization.)

(v) In order to estimate the dynamics in (4), the starting values of $\zeta_{r,t}$ and $\mathbb{E}\left[G_t^r \mid \boldsymbol{\Phi}_{t-1}\right]$ have to be set. In our empirical analysis, at each iteration in the estimation, we set them equal to the unconditional expected values, i.e., $\zeta_{r,0} = 0$ for all r and $\mathbb{E}\left[G_1^r \mid \boldsymbol{\Phi}_0\right] = c_r/(1-\rho_r)$, respectively. In addition, we assume that the roots of the polynomials $(1-\rho_r L)$, where L is the lag operator associated with (4), lie outside the unit circle (i.e., modulus $|\rho_r| < 1$), so that the unconditional expected value of G_t^r exists.

(vi) For notational convenience later, we collect the parameters of the model into three blocks (process, distribution, and correlation) as follows:

$$\boldsymbol{\theta}_P = (\boldsymbol{\mu}, \boldsymbol{\gamma}, \boldsymbol{\omega}, \boldsymbol{\alpha}, \boldsymbol{\beta}), \qquad \boldsymbol{\theta}_D = (\mathbf{c}, \boldsymbol{\rho}) \quad \text{and} \quad \boldsymbol{\theta}_C = \boldsymbol{\Gamma}, \tag{12}$$

where $\boldsymbol{\omega}, \boldsymbol{\alpha}, \boldsymbol{\beta}$ are K-dimensional vectors of GARCH(1, 1) parameters from (8); and $\mathbf{c}$ and $\boldsymbol{\rho}$ denote vectors of c_r and ρ_r parameters, respectively, though note that, in the case of constant G_t parameters, $\boldsymbol{\theta}_D$ reduces to just the set (λ, χ, ψ), of which in our special case, only λ needs to be estimated.

3. Estimation

The explicit form of the MGHyp density and the structure of (1) implies that, if $\boldsymbol{\gamma} = \mathbf{0}$ (necessary to make the scale shock $\boldsymbol{\varepsilon}_t$ independent of the unobserved realizations of G_t), then the estimation of the parameters in the model can be performed by direct maximization of the corresponding conditional (on $\boldsymbol{\Phi}_{t-1}$) likelihood function. In particular, with $\boldsymbol{\theta}_P^S = \boldsymbol{\theta}_P \setminus \boldsymbol{\gamma}$, vector $\mathbf{Y}_t$ is assumed to have an MGHyp distribution with density $f_{\mathbf{Y}_t}\left(\mathbf{y}; \boldsymbol{\theta}_P^S, \boldsymbol{\theta}_D, \boldsymbol{\theta}_C\right)$ given by

$$\int_0^\infty \frac{1}{(2\pi)^{K/2} |\mathbf{H}_t|^{1/2} g^{K/2}} \exp\left\{-\frac{(\mathbf{y}-\boldsymbol{\mu})' \mathbf{H}_t^{-1} (\mathbf{y}-\boldsymbol{\mu})}{2g}\right\} \times f_G(g; \boldsymbol{\theta}_D)\, dg,$$

where $f_G(\cdot)$ is the density of the unobserved mixing random variable G given in (2); $\mathbf{H}_t$ admits the decomposition (7) with constant conditional correlation; and the dynamics of the scale terms, $s_{k,t}$, are from (8). One can then evaluate this integral to yield the explicit expression

$$f_{\mathbf{Y}_t}\left(\mathbf{y}; \boldsymbol{\theta}_P^S, \boldsymbol{\theta}_D, \boldsymbol{\theta}_C\right) = C_t d_t^{-K/2+\lambda} \mathcal{K}_{\lambda-K/2}(d_t), \tag{13}$$

for $d_t = \sqrt{\left(\chi + (\mathbf{y}-\boldsymbol{\mu})' \mathbf{H}_t^{-1} (\mathbf{y}-\boldsymbol{\mu})\right)\psi}$ and the normalizing constant

$$C_t = \frac{\left(\sqrt{\chi\psi}\right)^{-\lambda} \psi^{K/2}}{(2\pi)^{K/2} |\mathbf{H}_t|^{1/2} \mathcal{K}_\lambda\left(\sqrt{\chi\psi}\right)}.$$

Direct maximization of the conditional (on $\boldsymbol{\Phi}_t$) likelihood function of $\mathbf{Y}$, denoted $L_{\mathbf{Y}}\left(\boldsymbol{\theta}_P^S, \boldsymbol{\theta}_D, \boldsymbol{\theta}_C\right)$, requires estimation of all the parameters in one step. This approach works for K very small, but becomes problematic as the number of assets increases, due to the quadratic increase in the number of parameters associated with the dispersion matrix, and the linear increase due to $\boldsymbol{\mu}$ and the univariate GARCH parameters. In the case with $\boldsymbol{\gamma} \neq \mathbf{0}$ and/or for large K, direct maximization is no longer feasible. However, estimation can be conducted via an Expectation-Conditional Maximization Either (ECME) algorithm from Liu and Rubin (1994). The standard ECME algorithm maximizes the likelihood function by an iterative procedure. It is a fixed-point algorithm and consists of the E-step, in which the realizations of the unobserved mixing variables $\{G_t\}_{t=1,\ldots,T}$ are imputed; and the CM-steps, in which all the parameters are updated by maximizing either the likelihood function $L_{\mathbf{Y}}(\boldsymbol{\theta}_P, \boldsymbol{\theta}_D, \boldsymbol{\theta}_C)$, or the conditional one, $L_{\mathbf{Y}|\mathbf{G}}(\boldsymbol{\theta}_P, \boldsymbol{\theta}_C)$. This is now detailed.

Consider the model with iid G_t, the complete log-likelihood function is

$$\log L_{\mathbf{Y},\mathbf{G}}(\boldsymbol{\theta}_P, \boldsymbol{\theta}_D, \boldsymbol{\theta}_C) = \log L_{\mathbf{Y}|\mathbf{G}}(\boldsymbol{\theta}_P, \boldsymbol{\theta}_C) + \log L_{\mathbf{G}}(\boldsymbol{\theta}_D), \tag{14}$$

where, based on the observed values $\mathbf{Y}_t = \mathbf{y}_t$ and conditional on $G_t = g_t = \widehat{G}_t$ (where the hatted notation indicates that G_t is estimated, and not observed), $t = 1, 2, \ldots, T$,

$$\log L_{\mathbf{Y}|\mathbf{G}}(\boldsymbol{\theta}_P, \boldsymbol{\theta}_C) = -\frac{1}{2}\sum_{t=1}^{T}\Big[K\log(2\pi) + \log|g_t \mathbf{S}_t \boldsymbol{\Gamma} \mathbf{S}_t| + g_t^{-1}(\mathbf{y}_t - \boldsymbol{\mu} - \boldsymbol{\gamma} g_t)' \mathbf{S}_t^{-1} \boldsymbol{\Gamma}^{-1} \mathbf{S}_t^{-1} (\mathbf{y}_t - \boldsymbol{\mu} - \boldsymbol{\gamma} g_t)\Big], \tag{15}$$

and $L_{\mathbf{G}}(\boldsymbol{\theta}_D)$ denotes the likelihood function of $(G_t \mid \boldsymbol{\Phi}_{t-1}) \sim \mathrm{GIG}(\lambda_t, \chi_t, \psi_t)$, $t = 1, 2, \ldots, T$. In a similar fashion to Bollerslev (1990), Engle and Sheppard (2002), and Pelletier (2006), we split (15) into a sum of two terms,

$$\log L_{\mathbf{Y}|\mathbf{G}}(\boldsymbol{\theta}_P, \boldsymbol{\theta}_C) = \log L_{\mathbf{Y}|\mathbf{G}}^{\mathrm{MV}}(\boldsymbol{\theta}_P) + \log L_{\mathbf{Y}|\mathbf{G}}^{\mathrm{Corr}}(\boldsymbol{\theta}_P, \boldsymbol{\theta}_C), \tag{16}$$

where $L_{\mathbf{Y}|\mathbf{G}}^{\mathrm{MV}}(\boldsymbol{\theta}_P)$ is the mean-volatility term given by

$$\log L_{\mathbf{Y}|\mathbf{G}}^{\mathrm{MV}}(\boldsymbol{\theta}_P) = -\frac{1}{2}\sum_{t=1}^{T}\Big[K\log(2\pi) + \log|\mathbf{S}_t|^2 + g_t^{-1}(\mathbf{y}_t - \boldsymbol{\mu} - \boldsymbol{\gamma} g_t)' \mathbf{S}_t^{-1} \mathbf{S}_t^{-1} (\mathbf{y}_t - \boldsymbol{\mu} - \boldsymbol{\gamma} g_t) + \log g_t^K\Big], \tag{17}$$

and $L_{\mathbf{Y}|\mathbf{G}}^{\mathrm{Corr}}(\boldsymbol{\theta}_P, \boldsymbol{\theta}_C)$ is the correlation term given by

$$\log L_{\mathbf{Y}|\mathbf{G}}^{\mathrm{Corr}}(\boldsymbol{\theta}_P, \boldsymbol{\theta}_C) = -\frac{1}{2}\sum_{t=1}^{T}\left[\log|\boldsymbol{\Gamma}| + \boldsymbol{e}_t' \boldsymbol{\Gamma}^{-1} \boldsymbol{e}_t - \boldsymbol{e}_t' \boldsymbol{e}_t\right], \tag{18}$$

where $\boldsymbol{e}_t = g_t^{-1/2} \mathbf{S}_t^{-1} \boldsymbol{\varepsilon}_t$ and $\boldsymbol{\varepsilon}_t = \mathbf{y}_t - \boldsymbol{\mu} - \boldsymbol{\gamma} g_t$ from (1).

Owing to the mixture structure of the MGHyp, $L_{\mathbf{Y}|\mathbf{G}}(\boldsymbol{\theta}_P, \boldsymbol{\theta}_C)$ is a multivariate Gaussian likelihood with a GARCH-type structure for the scales and a given conditional correlation model. As such, maximization of $L_{\mathbf{Y}|\mathbf{G}}(\boldsymbol{\theta}_P, \boldsymbol{\theta}_C)$ can be done in two steps. First, with the correlation structure ignored, the GARCH parameters in (8) are updated for each of the K assets separately (and concurrently with parallel computing) by maximizing $L_{\mathbf{Y}|\mathbf{G}}(\boldsymbol{\theta}_P \mid \boldsymbol{\theta}_C = \mathbf{I}_K)$, and, in the second step, the correlation parameters $\boldsymbol{\theta}_C$ are updated from the first-step de-volatilized residuals. This idea is not new, and is similar to the Gaussian setup in Bollerslev (1990), Engle (2002), and Pelletier (2006). Given the $\boldsymbol{\theta}_P$ and $\boldsymbol{\theta}_C$ estimates, denoted by $\widehat{\boldsymbol{\theta}}_P$ and $\widehat{\boldsymbol{\theta}}_C$, the mixing process parameters $\boldsymbol{\theta}_D$ are updated by maximizing $L_{\mathbf{Y}}\big(\boldsymbol{\theta}_D \mid \widehat{\boldsymbol{\theta}}_P, \widehat{\boldsymbol{\theta}}_C\big)$. Given all the parameter updates, we proceed with the next E-step update (see (19)) of the unobserved mixing random variable $\mathbf{G}$, and continue to iterate until convergence.

Observe that $L_{\mathbf{Y}|\mathbf{G}}(\boldsymbol{\theta}_P, \boldsymbol{\theta}_C)$ reduces to the mean-volatility component, $L^{\mathrm{MV}}_{\mathbf{Y}|\mathbf{G}}(\boldsymbol{\theta}_P)$, if and only if we assume zero correlations. As such, $L^{\mathrm{MV}}_{\mathbf{Y}|\mathbf{G}}(\boldsymbol{\theta}_P)$ corresponds to the likelihood of $\mathbf{Y}$ conditional on $\mathbf{G}$ for the model under zero correlations. Based on this decomposition, estimates of $\boldsymbol{\theta}_P$, $\boldsymbol{\theta}_C$ and $\boldsymbol{\theta}_D$ can be obtained by the following ECME algorithm.

E-step: Calculate $\mathbb{E}\left[\log L_{\mathbf{Y},\mathbf{G}} \mid \mathbf{Y}; \widehat{\boldsymbol{\theta}}_P, \widehat{\boldsymbol{\theta}}_C, \widehat{\boldsymbol{\theta}}_D\right]$.
The log-likelihood function (17) is linear with respect to g_t and g_t^{-1} ($\log g_t^K$ can be ignored without influencing the first order conditions). Hence the E-step involves replacing unobserved realizations of G_t and G_t^{-1} in (14) by the imputed values, $\widehat{G}_t$. Calculation shows that (see, e.g., Paolella, 2013, Eq. 35)

$$\left(G_t \mid \boldsymbol{\Phi}_t; \widehat{\boldsymbol{\theta}}_P, \widehat{\boldsymbol{\theta}}_C, \widehat{\boldsymbol{\theta}}_D\right) \sim \mathrm{GIG}\left(\lambda_t^*, \chi_t^*, \psi_t^*\right), \tag{19}$$

where

$$\lambda_t^* = \lambda_t - K/2,$$

$$\chi_t^* = \chi_t + (\mathbf{y}_t - \widehat{\boldsymbol{\mu}})' \mathbf{S}_t^{-1} \widehat{\boldsymbol{\Gamma}}^{-1} \mathbf{S}_t^{-1} (\mathbf{y}_t - \widehat{\boldsymbol{\mu}}) \quad \text{and}$$

$$\psi_t^* = \psi_t + \widehat{\boldsymbol{\gamma}}' \mathbf{S}_t^{-1} \widehat{\boldsymbol{\Gamma}}^{-1} \mathbf{S}_t^{-1} \widehat{\boldsymbol{\gamma}}.$$

The latent values of g_t and g_t^{-1} are then updated by their conditional expectations from (19), using the expression for the moments of the GIG random variable given in (43).

CM1-step: Update $\boldsymbol{\theta}_P$ and $\boldsymbol{\theta}_C$.

(P) Update $\boldsymbol{\theta}_P$ by computing

$$\arg\max_{\boldsymbol{\theta}_P} \log L^{\mathrm{MV}}_{\mathbf{Y}|\mathbf{G}}(\boldsymbol{\theta}_P), \tag{20}$$

where $L^{\mathrm{MV}}_{\mathbf{Y}|\mathbf{G}}(\boldsymbol{\theta}_P)$ is a Gaussian likelihood with zero correlation, so we can estimate the parameters of each asset, $(\mu_k, \gamma_k, \omega_k, \alpha_k, \beta_k)$, separately by maximizing the corresponding likelihood function.

(C) Update $\boldsymbol{\theta}_C$ by computing the usual empirical correlation estimator (the MLE under normality) of the de-volatilized residuals $\widehat{G}_t^{-1/2} \mathbf{S}_t^{-1} \widehat{\boldsymbol{\varepsilon}}_t$, where $\widehat{\boldsymbol{\varepsilon}}_t = \mathbf{Y}_t - \widehat{\boldsymbol{\mu}} - \widehat{\boldsymbol{\gamma}} \widehat{G}_t$, $\widehat{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{\gamma}}$ are obtained in part (P) directly above, and $\widehat{G}_t$ is obtained in the E-step.

CM2-step: Given the CM1-step updates of $\boldsymbol{\theta}_P$ and $\boldsymbol{\theta}_C$, obtain new updates of $\boldsymbol{\theta}_D$ by maximizing the incomplete data log-likelihood function, i.e., compute

$$\arg\max_{\boldsymbol{\theta}_D} \log L_{\mathbf{Y}}\left(\boldsymbol{\theta}_D \mid \widehat{\boldsymbol{\theta}}_P, \widehat{\boldsymbol{\theta}}_C\right). \tag{21}$$

Iterate the above steps until convergence.

The hybrid GARCH-SV extension differs from the iid case because the random variable $(G_t \mid \mathbf{Y}_t)$ in (4) is not measurable with respect to $\boldsymbol{\Phi}_{t-1}$ but only with respect to $\boldsymbol{\Phi}_t$. However, this random variable conditional on $\boldsymbol{\Phi}_{t-1}$ or on $\boldsymbol{\Phi}_t$ is the same random variable. Hence, the above algorithm applied to the hybrid GARCH-SV model maximizes not only the conditional likelihood function $L_{\mathbf{Y}|(\mathbf{G}|\mathbf{Y})}$, but also the conditional (on $\boldsymbol{\Phi}_{t-1}$) likelihood function of the data, $L_{\mathbf{Y}}$. In fact, the two functions are equal. In Appendix A we prove the equality between the two conditional log-likelihoods, and the monotonic convergence of the proposed algorithm.

Remarks. (i) In contrast to the Gaussian distribution, setting $\boldsymbol{\Gamma}$ equal to the identity matrix does not imply independence when assuming MGHyp returns, due to the dependence induced by the GIG mixing variable G_t. But via the application of the ECME algorithm, which conditions on the realizations of G_t, we can estimate the parameters of the GARCH equations separately for each asset. This is the key to fast, simple, joint likelihood-based estimation of a multivariate non-normal model.

(ii) In the ECME algorithm, the E-step computation of the conditional expectations of G_t and G_t^{-1} involves computation of ratios of Bessel functions (3) for different arguments. In case of large arguments (χ_t^* and ψ_t^* in (19) can get very large because they involve a quadratic form of the inverse of the covariance matrix), numerical computation is subject to rounding error which affects estimates of all the parameters. We propose a method which increases numerical accuracy such that estimation for all data windows used in our empirical application was successful. It is given in Appendix E.

(iii) The method which increases the accuracy of the Bessel function computation given in Appendix E is also used in the CM2-step for evaluating $\log L_{\mathbf{Y}}$ for different $\boldsymbol{\theta}_D$ values in (21).

(iv) Here we state the starting values used for the estimation procedure. Those of the asymmetry parameter γ_k are taken to be zero, and the remaining ones in $\boldsymbol{\theta}_P$, $(\mu_k, \omega_k, \alpha_k, \beta_k)$, to be those values obtained from the normal-based GARCH estimates, using the method in Paolella and Polak (forthcoming) to avoid inferior local likelihood maxima. For $\boldsymbol{\theta}_C$, we use the empirical correlation matrix computed from the normal-based GARCH residuals. For $\boldsymbol{\theta}_D = (\lambda, \chi, \psi)$, we have confirmed that the likelihood is such that optimization is rather robust to the choice of starting values. For the special case of the multivariate asymmetric variance gamma distribution (MAVG) used in the empirical section below, we use $\lambda = 2$ as the starting value. For the hybrid GARCH-SV model we set, $c = 0.1$ and $\rho = 0.8$. In the estimation with a rolling window, we use also as starting values the previous window estimates; and take the final estimates to be those with the higher likelihood value.

(v) If the starting values are sufficiently close to those which maximize the likelihood, or if the likelihood function is unimodal in the parameter space, and the maximum is not on the boundary of the parameter space, then monotonicity in the likelihood values of the consecutive ECME estimates, which is shown in Appendix A, guarantees their convergence to the corresponding maximum likelihood parameters; see, e.g., McLachlan and Krishnan (2008). As such, under further standard regularity conditions on the likelihood, consistent and asymptotically normal point estimates are obtained.

(vi) As reported by Protassov (2004), and also confirmed by our studies, one or more of the MGHyp shape parameters can have a relatively flat likelihood (already after fixing χ or ψ for identification purposes), implying possible numeric problems when maximizing the likelihood. To this end, we focus our empirical analysis on a special case of the MGHyp distribution in which G_t is gamma distributed with shape parameter $\lambda > 0$ and unit scale parameter (the multivariate asymmetric variance gamma, MAVG, model). The MAVG distribution does not share the flat likelihood problem of the fully general MGHyp distribution, and so is considerably faster and numerically more reliable to estimate, but still retains the flexibility required for modeling asset returns by allowing for individual asset asymmetry parameters and also higher kurtosis than the normal distribution. (An empirical motivation of the MAVG distribution is given in the companion paper (Paolella and Polak, 2015b), where its performance is compared with other special cases of MGHyp distribution.)

(vii) For $\boldsymbol{\gamma} = 0$ and K small, we confirmed the accuracy of the proposed EM algorithm by comparing the EM estimates with the results based on the direct likelihood maximization of (13).

4. Option pricing

We present a feasible technique which allows for multivariate option pricing in the framework of our model. Since the work of Duan (1995), GARCH models have become increasingly popular in option pricing. More recent literature includes Heston and Nandi (2000), who derive a nearly closed-form pricing formula under normal return innovations and the valuation assumption from Duan (1995); Christoffersen et al. (2006), who propose a model with inverse Gaussian innovations which allows for conditional skewness; Barone-Adesi et al. (2008), who use filtered historical simulation; Christoffersen et al. (2010), who develop a theoretical framework for option valuation under very general assumptions which allow for conditional heteroskedasticity and non-normality; and Rombouts and Stentoft (2011), who consider multivariate option pricing in a model with a finite normal mixture. Our model is also multivariate and, as it allows for all the primary stylized facts of asset returns, it is expected to be a good candidate for option pricing, given a feasible calibration algorithm.

The proposed algorithm combines the equivalent martingale measure (EMM) technique in the presence of a GARCH structure as in Christoffersen et al. (2010), with a Monte Carlo simulation method. Like in Barone-Adesi et al. (2008), it does not focus on the analytical form of the change of measure. Barone-Adesi et al. (2008) utilize a Monte Carlo simulation based on QML estimates of model parameters under the historical measure and calibrate the EMM on the option prices. In contrast to their nonparametric method, our algorithm estimates the model parameters via maximizing the complete data likelihood function under the historical measure P, it changes the measure as if G_t were observed, and recovers the missing information about the future G_t values under the EMM from the option prices observed on the market. Crucially for the COMFORT model, the proposed algorithm does not suffer from the curse of dimensionality, and so, it is applicable in a multivariate setup with a large number of underlying assets.

Denote by $\mathbf{X}_t = \left(X_{t,1}, X_{t,2}, \dots, X_{t,K}\right)'$ a vector of prices of assets $k = 1, 2, \dots, K$ at time t. The price of an option contract at time t with maturity T and terminal payoff function $\varrho_\vartheta(\mathbf{X}_T)$, ϑ being the set of relevant parameters such as the strike price, can be computed as the following discounted expectation,

$$C_t(T, \varrho_\vartheta) = \exp\left(-(T-t)\,r\right)\mathbb{E}^P\left[\varrho_\vartheta(\mathbf{X}_T)\,\frac{\mathrm{d}Q^*}{\mathrm{d}P}\,\middle|\,\boldsymbol{\Phi}_t\right], \tag{22}$$

where $\frac{\mathrm{d}Q^*}{\mathrm{d}P} \mid \boldsymbol{\Phi}_t$ is the change of measure such that the discounted stock price process under Q^* is a martingale with respect to the $\boldsymbol{\Phi}_t$ filtration. This filtration is defined in Section 2 and it is associated with the incomplete conditional density function $f_{\mathbf{Y}_t}$. For the purpose of option pricing, we define a complete information filtration,[1] associated with the complete information density $f_{\mathbf{Y}_t, G_t}$, by $\mathcal{F}_t = \sigma(\{G_1, \mathbf{Y}_1, G_2, \mathbf{Y}_2, \dots, G_t, \mathbf{Y}_t\})$; and a second filtration, which includes information about the realization of G_{t+1}, $\mathcal{F}_t^{+G} = \sigma(\{G_1, \mathbf{Y}_1, G_2, \mathbf{Y}_2, \dots, G_t, \mathbf{Y}_t, G_{t+1}\})$.

The mixture property of the MGHyp distribution implies that $\mathbf{Y}_t \mid \mathcal{F}_{t-1}^{+G} \sim \mathrm{N}(\boldsymbol{\mu} + \boldsymbol{\gamma} G_t, G_t \mathbf{H}_t)$, hence the standard theory for option pricing under normality applies. In particular, following Christoffersen et al. (2010), and Rombouts and Stentoft (2011), if we impose the exponential affine form on the Radon–Nikodym derivative, with respect to $\mathcal{F}_t^{+G}$, then, under the corresponding measure Q^{+G}, as detailed in Appendix C, the dynamics of the returns remain Gaussian although with a shift in the mean.

[1] Note an important difference that the complete information filtration does not imply directly that the market is complete. Features like (i) the presence of time varying volatilities, and, as pointed by the referee, (ii) the discrete time nature of the model with the fact that asset prices can "jump" to infinitely many values from day to the next, and only a finite number of instruments (options, underlying asset prices, risk-free bond) are available in the market; they all induce incompleteness.

Next, we define a change of measure, under $\mathcal{F}_t$, by

$$\frac{\mathrm{d}Q}{\mathrm{d}P}\,\bigg|\,\mathcal{F}_t = \mathbb{E}^P\left[\left(\frac{\mathrm{d}Q^{+G}}{\mathrm{d}P}\,\bigg|\,\mathcal{F}_t^{+G}\right)\,\bigg|\,\mathcal{F}_t\right]. \tag{23}$$

This transformation defines a Radon–Nikodym derivative under $\mathcal{F}_t$ and the resulting measure Q is an EMM, under $\mathcal{F}_t$. Note that, from (36) the information contained in $\mathcal{F}_t$ is sufficient to construct $\frac{\mathrm{d}Q^{+G}}{\mathrm{d}P} \mid \mathcal{F}_t^{+G}$. Therefore, the change of measure (23) is available explicitly and equal to (36). Under the measure Q,

$$\mathbf{Y}_t \mid \mathcal{F}_{t-1}^{+G} \overset{Q}{\sim} N\left((\boldsymbol{\mu} - \mathbf{r}) + \boldsymbol{\gamma} G_t + \frac{1}{2}\mathrm{diag}\left(\mathbf{S}_t^2\right) G_t,\; G_t \mathbf{H}_t\right), \tag{24}$$

where $\mathbf{r} = (r, \dots, r)'$ is a $K \times 1$ vector of the risk free interest rates r.

Under no arbitrage conditions, the derivatives which are a function of the underlying stock price process can be priced as the expected value, under EMM, of their future cash flows discounted using the risk free interest rate, as given in (22). So far we have derived the necessary tools for pricing the option under the condition that the realizations of the G_t sequence are observed.

The algorithm in Appendix D mitigates this problem. Instead of analytically deriving the change of measure, it uses (24) together with simulated future G_t's to generate paths of the price processes, and calibrates the $\boldsymbol{\theta}_D^{Q^*}$ to match the option prices observed on the market. This method implicitly defines the change of measure, and it is feasible even in case of a large number of assets.

Having obtained, the parameter estimates under the risk neutral measure Q^* the conditional distribution of the returns is given by

$$\mathbf{Y}_t \mid \boldsymbol{\Phi}_{t-1} \overset{Q^*}{\sim} \mathrm{MGHyp}\left(\boldsymbol{\mu}^{Q^*}, \boldsymbol{\gamma}_t^{Q^*}, \mathbf{H}_t^{Q^*}, \boldsymbol{\theta}_D^{Q^*}\right), \tag{25}$$

where $\boldsymbol{\mu}^{Q^*} = \boldsymbol{\mu} - \mathbf{r}$ is the location vector, $\boldsymbol{\gamma}_t^{Q^*} = \boldsymbol{\gamma} + 1/2\mathrm{diag}\left(\left(\mathbf{S}_t^{Q^*}\right)^2\right)$ is the time-varying asymmetry vector, $\mathbf{H}_t^{Q^*} = \mathbf{S}_t^{Q^*}\boldsymbol{\Gamma}\mathbf{S}_t^{Q^*}$ is the dispersion matrix under Q^*, and $\mathbf{S}_t^{Q^*}$ is the diagonal matrix of the scale terms which differs from $\mathbf{S}_t$ in (7) because of the dependence on past $\boldsymbol{\gamma}_t^{Q^*}$ and past simulated values of $G_t^{Q^*}$.

Given the distribution of the returns under the risk neutral measure, one can price various options by using (22) and Monte Carlo simulation.

5. Analysis of option pricing

To demonstrate the applicability and competitiveness of the model in option pricing, we use the data set consisting of the 2767 daily returns of $K = 30$ components of the Dow Jones Industrial Index (DJ-30) from January 2nd, 2001, to December 30th, 2011 (based on the DJ-30 composition as of June 8th, 2009). Observe that this period covers the global financial crisis of 2008 (but does not extend further, past the end of 2011, because of the date of onset of this project). Returns for each asset are computed as continuously compounded percentage returns, given by $y_{k,t} = 100 \log\left(x_{k,t}/x_{k,t-1}\right)$, where $x_{k,t}$ is the price of asset k at time t.

We consider European call option on the average price of stocks from the DJ-30 with the price given by

$$C_t(T, \vartheta) = \exp(-r(T-t)) \times \mathbb{E}^{Q^*}\left[\max\left(\frac{1}{K}\sum_{k=1}^{K} X_{T,k} - \vartheta, 0\right)\,\middle|\,\boldsymbol{\Phi}_t\right], \tag{26}$$

with the interest rate r set to 0, ϑ being the strike, and $(T-t)$ the time to maturity. We price 21 such option contracts, with maturities 30, 60 and 90 days, moneyness 0.8, 0.9, 0.95, 1, 1.05,

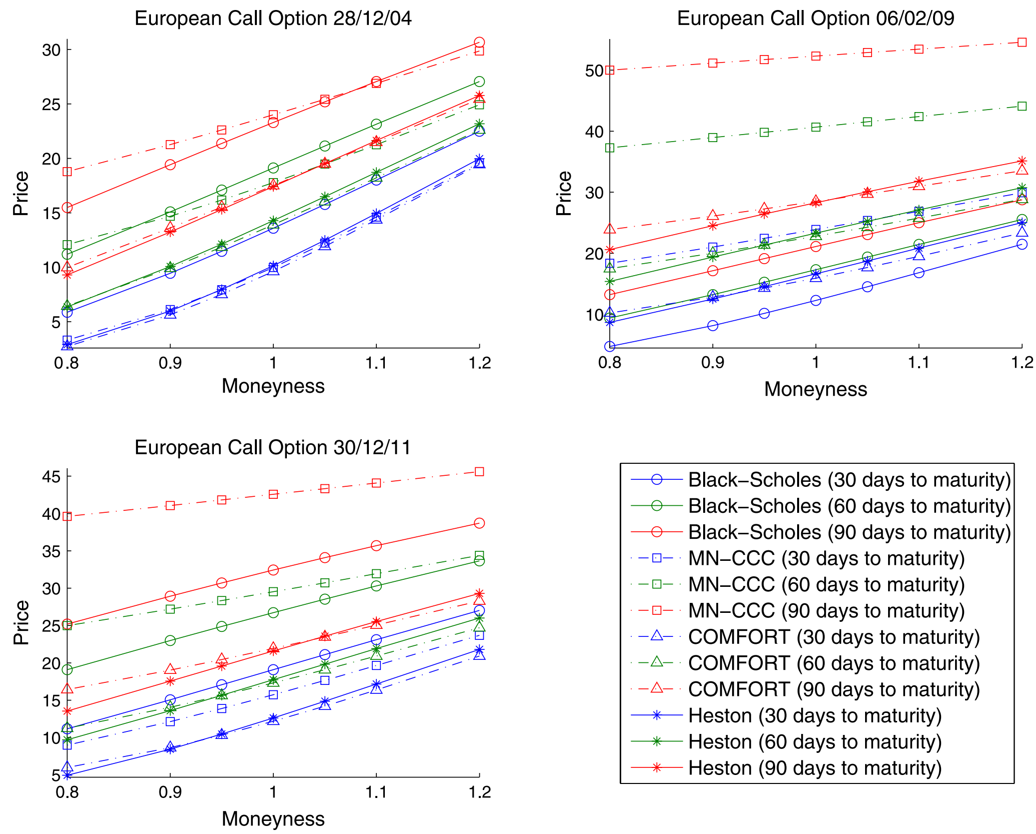

Fig. 1. Comparison of European call option prices, on the average of $K = 30$ stocks. The prices are implied by different models: (i) Black–Scholes model, (ii) Heston model, (iii) MN-CCC GARCH(1, 1) model, and (iv) MAVG-CCC GARCH(1, 1) model. On a given day, a single λ^{Q^*} parameter in the MAVG-CCC GARCH(1, 1) model is calibrated to match 21 option prices implied by Heston model for three maturities: 30, 60, and 90 days, and 7 levels of moneyness: 0.8, 0.9, 0.95, 1, 1.05, 1.1, and 1.2. Number of Monte Carlo simulations is set to $N = 20{,}000$. Each panel corresponds to a different day in the sample and different Q^* parameters. Upper left: A low volatility day (28/12/2004) with $\widehat{\lambda}^{Q^*} = 1.92$. Upper right: A high volatility day (06/02/2009) with $\widehat{\lambda}^{Q^*} = 1.35$. Bottom left: The last day in the sample (30/12/2011) with mid level of volatility and $\widehat{\lambda}^{Q^*} = 1.54$.

1.1, and 1.2, and $X_{t,k}$ is normalized to 100, for $k = 1, \ldots, K$. We use four different models: (i) Black–Scholes model of Black and Scholes (1973); (ii) MN-CCC GARCH(1, 1) model of Bollerslev (1990); (iii) Heston model of Heston (1993); and (iv) MAVG-CCC GARCH(1, 1) model.

Our interest centers on the quality of option prices implied by our model. Using a rolling window of 1000 observations, we estimate the parameters of MN-CCC GARCH(1, 1) model, and the MAVG-CCC GARCH(1, 1) model, under the historical measure P. The former are obtained by a two-step ML estimation method from Bollerslev (1990), and the latter by the ECME algorithm from Section 3. Given the parameter estimates under P, option prices implied by MN-CCC GARCH(1, 1) model are computed as in Rombouts and Stentoft (2011). The annualized asset price volatility in the Black–Scholes model and in the Heston model are set to the average across assets of long term volatilities implied by the MN-CCC GARCH(1, 1) model. The instantaneous volatility in the Heston model is set to the average across assets of conditional volatilities from the MN-CCC GARCH(1, 1) model. The rate of mean reversion of the volatility in the Heston model is matched with the mean reversion of the MN-CCC GARCH(1, 1) model by equating the half-life of models. Additional parameters in the Heston model such as volatility of the volatility, the risk premium for volatility, and the correlation between the underlying and the volatility, are set to 0.3, 0.03, and −0.6, respectively. The COMFORT option prices come from the calibration algorithm in Appendix D performed with $N = 20{,}000$ simulation paths, and with the prices implied by the Heston model set as the market prices.

The feasibility of the proposed algorithm and its flexibility is demonstrated by calibrating the MAVG-CCC GARCH(1, 1) model to replicate all 21 benchmark prices. Fig. 1 displays the results of the calibration for three different days in the sample which correspond to different market conditions. The MAVG-CCC GARCH(1, 1) model matches all the maturities and moneyness much closer than the MN-CCC GARCH(1, 1) model or the Black–Scholes model. Fig. 2 provides corresponding implied volatilities which are extracted using the Black–Scholes formula. The implied volatilities from the MAVG-CCC GARCH(1, 1) model are the closest to those from the Heston model. They are convex and decreasing with moneyness, and increasing with maturity. For deep out of the money (in the money) options, they are higher (lower) than those from the Heston model.

The flexibility of the COMFORT model in replicating options prices is demonstrated in Fig. 3. It depicts prices of 30, 60 and 90 days to maturity European call options implied by the MAVG-CCC GARCH(1, 1) model for various values of the λ^{Q^*} parameter. In particular, increase in λ^{Q^*} implies an increase in option prices for all moneyness and maturities. Low values of λ^{Q^*} correspond to a very steep and convex increase of the option price when the contract goes from out of the money to in the money levels. Higher values of λ^{Q^*} correspond to higher prices, and to less steep and more linear increase of the option price with the moneyness. This is intuitive because in the MAVG-CCC GARCH(1, 1) model $\mathbb{E}^{Q^*}[G_t] = \lambda^{Q^*}$, and a higher λ^{Q^*} corresponds to the risk neutral dynamics which associate more risk premium to the spikes in the returns, and options which are far out of the money still have some value because of the Q^*-probability of large spikes being higher.

Fig. 4 illustrates the role of the λ^{Q^*} parameter in the MAVG-CCC GARCH(1, 1) model from the implied volatility perspective. In the left column, the remaining Q^* parameters are the same as

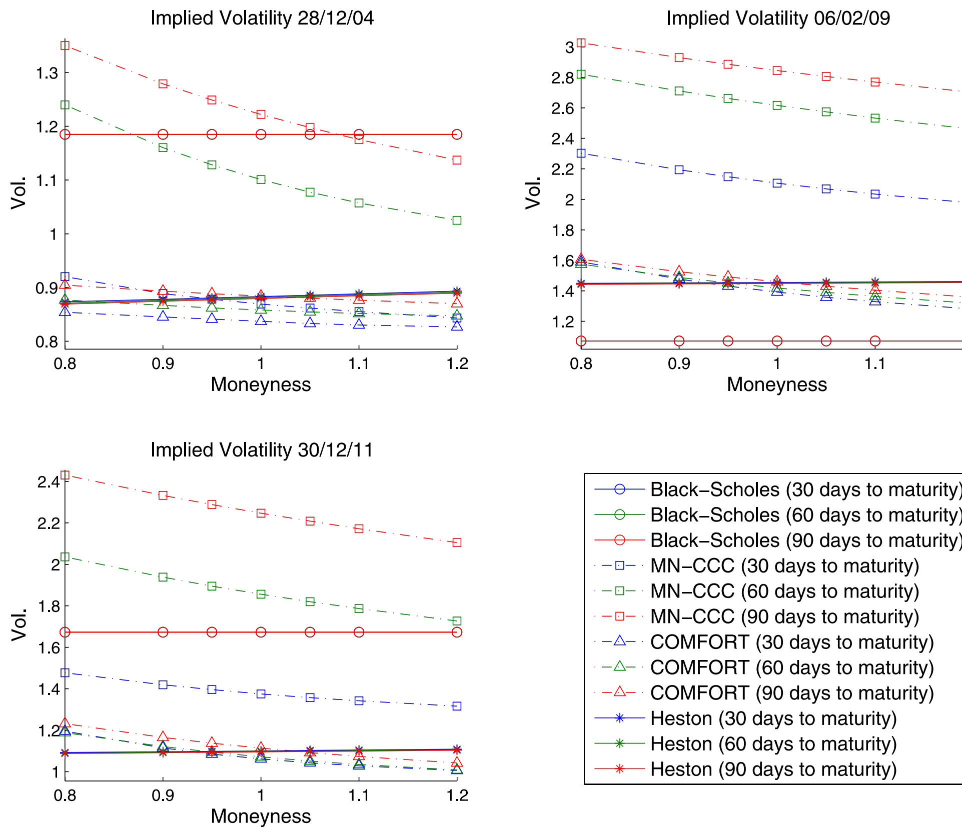

Fig. 2. Comparison of European call option in terms of implied volatilities which are extracted using the Black–Scholes formula with prices implied by different models. Analogous to Fig. 1. Upper left: A low volatility day (28/12/2004) with $\widehat{\lambda}^{Q^*} = 1.92$. Upper right: A high volatility day (06/02/2009) with $\widehat{\lambda}^{Q^*} = 1.35$. Bottom left: The last day in the sample (30/12/2011) with mid level of volatility and $\widehat{\lambda}^{Q^*} = 1.54$.

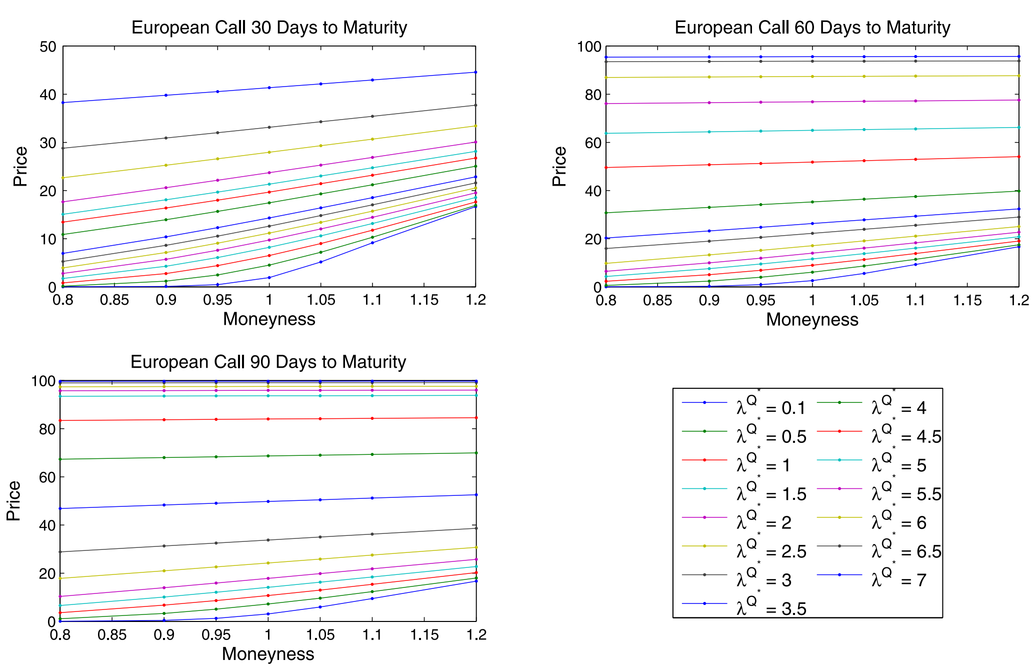

Fig. 3. Prices of 30, 60 and 90 days to maturity European call options on the average of $K = 30$ stocks based on MAVG-CCC GARCH(1, 1) model for 7 levels of moneyness: 0.8, 0.9, 0.95, 1, 1.05, 1.1, and 1.2, and different λ^{Q^*} parameters (remaining Q^* parameters are the same as on 28/12/04). Number of Monte Carlo simulations is set to $N = 20{,}000$.

for the options priced on 28/12/2004, and in the right column the remaining Q^* parameters are the same as for the option priced on 06/02/2009. Depending on λ^{Q^*}, the implied volatility surface exhibits different shapes. It is (i) for $\lambda^{Q^*} < 1$, decreasing in maturity, convex in moneyness; (ii) for $\lambda^{Q^*} = 1$, slowly decaying in maturity and almost flat in moneyness; and (iii) for $\lambda^{Q^*} > 1$, increasing in maturity, slightly convex in moneyness, and skewed toward lower moneyness. By comparing vertically the panels in

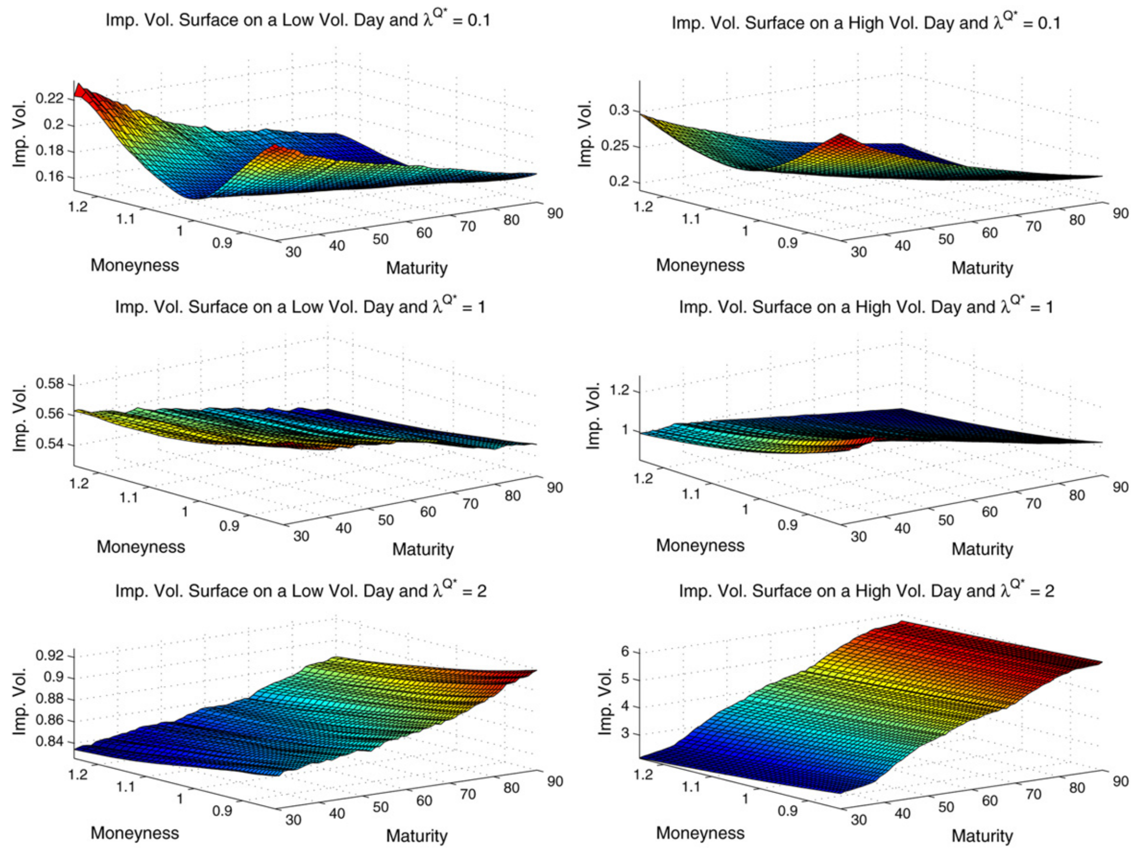

Fig. 4. Different shapes of the implied volatility surface for the MAVG-CCC GARCH(1, 1) model (extracted using the Black–Scholes formula) depending on $\lambda^{Q^*} = 0.1, 1, 2$ parameter. Number of Monte Carlo simulations is set to $N = 20{,}000$. Left column: Remaining Q^* parameters are the same as on 28/12/04 when the volatility is low. Right column: Remaining Q^* parameters are the same as on 06/02/09 when the volatility is high.

Fig. 4 we see that it is higher and also more steep along maturity on the day with a higher volatility, i.e., higher $s_{k,t}$ for $k = 1, \ldots, K$.

The hybrid GARCH-SV extension of the model has an extra degree of flexibility, which for contracts with long expiration, is difficult to control in calibration. However, for the purpose of pricing options with short maturities, the dynamics in (4) allow to reproduce more complex shapes of implied volatility surface. It is shown in Gatheral (2006, Ch. 3) that the Heston model fails to fit the observed implied volatility surface for short expirations. As a solution to this problem, when calibrating to short maturity contracts, they propose to add jumps to the model. The hybrid GARCH-SV extension has a similar role in reproducing an implied volatility surface for short maturities in the COMFORT model. Fig. 5 illustrates this with an example of two implied volatility surfaces for contracts with maturity between 5 and 30 days. The c^{Q^*} and ρ^{Q^*} parameters are set to keep the unconditional mean value of G_t in both panels the same, the difference is in the mean reversion speed of (4). The left panel, with slow mean reversion, provides a u-shaped volatility surface. The right panel, with faster mean reversion, results in a v-shaped volatility surface which is more common for short term contracts (see, e.g., Gatheral, 2006, Ch. 3). Changing the c^{Q^*} and ρ^{Q^*} parameters along the unconditional mean value of G_t has the same effect on implied volatility surface as changing the λ^{Q^*} in the model with iid G_t's.

6. Conclusions

We introduce a new class of models which combines GARCH-type dynamics with an SV structure (hybrid GARCH-SV class). The former captures the asset-specific volatility clustering effects and the latter is responsible for common market shocks. The proposed model also allows for a new type of dynamic in the dependency structure leading to additional dynamics in the higher-order moments. Maximum likelihood estimation is numerically reliable and fast, and can be used with a large number of assets. It yields consistent and asymptotically normal estimates of the parameters. The model lends itself to multivariate option pricing by combining the equivalent martingale measure technique in the presence of a GARCH structure with a Monte Carlo simulation to match the option prices observed on the market. It is demonstrated empirically that the model matches the option prices implied by a Heston model for European call options on the average of the stocks. Future work will pursue the performance of the proposed option pricing algorithm with market data on different option contracts.

An important property of the model is that it delivers a non-Gaussian predictive distribution with a tractable sum of margins, and so could be applied to minimum expected shortfall portfolio optimization, provided the expected shortfall can be computed fast and accurately. In a companion paper, Paolella and Polak (2015b), we develop fast methods for this computation. In addition, we modify the CCC structure in the same vein as in the DCC model of Engle (2002) and the VC model of Tse and Tsui (2002), and demonstrate superior performance of these models over Gaussian competitors in terms of multivariate density forecasting, risk prediction, and portfolio optimization.

Appendix A. Convergence of the ECME algorithm

The conditional on $\boldsymbol{\Phi}_{t-1}$ incomplete data log-likelihood function of $\mathbf{Y}$ is given by

$$\log L_{\mathbf{Y}}(\boldsymbol{\theta}) = \sum_{t=1}^{T} \log f_{\mathbf{Y}_t|\boldsymbol{\Phi}_{t-1}}(\mathbf{y}_t; \boldsymbol{\theta}), \tag{27}$$

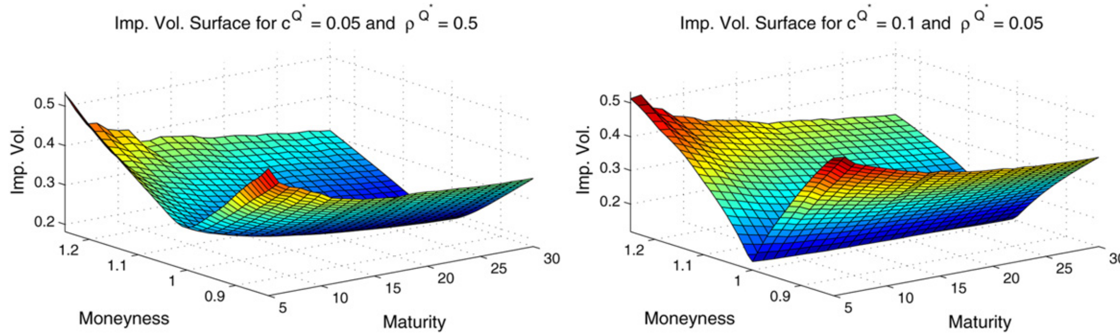

Fig. 5. Different shapes of the implied volatility surface for the MAVG-CCC hybrid GARCH(1, 1)-SV model (extracted using the Black–Scholes formula) depending on c^{Q^*} and ρ^{Q^*} parameters. Number of Monte Carlo simulations is set to $N = 20{,}000$. Remaining Q^* parameters are the same as on 28/12/04 when the volatility is low. Left column: Slowly mean reverting dynamics of $\lambda_t^{Q^*}$. Right column: Fast mean reversion in the dynamics of $\lambda_t^{Q^*}$.

where $\boldsymbol{\theta} = (\boldsymbol{\theta}_P, \boldsymbol{\theta}_C, \boldsymbol{\theta}_D)$ is defined in (12). By Bayes' theorem, we have

$$\log f_{\mathbf{Y}_t|\boldsymbol{\Phi}_{t-1}}(\mathbf{y};\boldsymbol{\theta}) = \log f_{\mathbf{Y}_t|G_t,\boldsymbol{\Phi}_{t-1}}(\mathbf{y} \mid g;\boldsymbol{\theta}_P,\boldsymbol{\theta}_C) + \log f_{G_t|\boldsymbol{\Phi}_{t-1}}(g;\boldsymbol{\theta}_D) - \log f_{G_t|\boldsymbol{\Phi}_t}(g \mid \mathbf{y};\boldsymbol{\theta}). \tag{28}$$

Taking expectations of both sides of (28) with respect to the conditional distribution of $(\mathbf{Y}_t, G_t)$ given $\boldsymbol{\Phi}_t$, using the ℓ iteration fit $\boldsymbol{\theta}^{[\ell]} = \left(\boldsymbol{\theta}_P^{[\ell]}, \boldsymbol{\theta}_C^{[\ell]}, \boldsymbol{\theta}_D^{[\ell]}\right)$ for $\boldsymbol{\theta}$, we have that

$$\begin{aligned}\log f_{\mathbf{Y}_t|\boldsymbol{\Phi}_{t-1}}(\mathbf{y};\boldsymbol{\theta}) &= \mathbb{E}\left[\log f_{\mathbf{Y}_t|G_t,\boldsymbol{\Phi}_{t-1}}(\mathbf{y} \mid g;\boldsymbol{\theta}_P,\boldsymbol{\theta}_C) \mid \boldsymbol{\Phi}_t, \boldsymbol{\theta}^{[\ell]}\right] \\ &+ \mathbb{E}\left[\log f_{G_t|\boldsymbol{\Phi}_{t-1}}(g;\boldsymbol{\theta}_D) \mid \boldsymbol{\Phi}_t, \boldsymbol{\theta}^{[\ell]}\right] \\ &- \mathbb{E}\left[\log f_{G_t|\boldsymbol{\Phi}_t}(g \mid \mathbf{y};\boldsymbol{\theta}) \mid \boldsymbol{\Phi}_t, \boldsymbol{\theta}^{[\ell]}\right].\end{aligned}$$

Evaluating the last expression at $\mathbf{Y}_t = \mathbf{y}_t$, and summing it over $t = 1, \ldots, T$, gives

$$\begin{aligned}\log L_{\mathbf{Y}}(\boldsymbol{\theta}) &= \sum_{t=1}^{T} \mathbb{E}\left[\log f_{\mathbf{Y}_t|G_t,\boldsymbol{\Phi}_{t-1}}(\mathbf{y}_t \mid g;\boldsymbol{\theta}_P,\boldsymbol{\theta}_C) \mid \boldsymbol{\Phi}_t, \boldsymbol{\theta}^{[\ell]}\right] \\ &+ \sum_{t=1}^{T} \mathbb{E}\left[\log f_{G_t|\boldsymbol{\Phi}_{t-1}}(g;\boldsymbol{\theta}_D) \mid \boldsymbol{\Phi}_t, \boldsymbol{\theta}^{[\ell]}\right] \\ &- \sum_{t=1}^{T} \mathbb{E}\left[\log f_{G_t|\boldsymbol{\Phi}_t}(g \mid \mathbf{y}_t;\boldsymbol{\theta}) \mid \boldsymbol{\Phi}_t, \boldsymbol{\theta}^{[\ell]}\right].\end{aligned} \tag{29}$$

We need to show that

$$\log L_{\mathbf{Y}}\left(\boldsymbol{\theta}^{[\ell+1]}\right) \geq \log L_{\mathbf{Y}}\left(\boldsymbol{\theta}^{[\ell]}\right). \tag{30}$$

The CM1-step of the algorithm maximizes the first term in (29) and does not change the second term. The CM2-step finds $\boldsymbol{\theta}_D^{[\ell+1]}$ such that

$$\log L_{\mathbf{Y}}\left(\boldsymbol{\theta}_P^{[\ell+1]}, \boldsymbol{\theta}_C^{[\ell+1]}, \boldsymbol{\theta}_D^{[\ell+1]}\right) \geq \log L_{\mathbf{Y}}\left(\boldsymbol{\theta}_P^{[\ell+1]}, \boldsymbol{\theta}_C^{[\ell+1]}, \boldsymbol{\theta}_D\right), \quad \text{for all } \boldsymbol{\theta}_D. \tag{31}$$

Finally, for the last term in (29), by Jensen's inequality and the concavity of the logarithmic function, for all $\boldsymbol{\theta}$,

$$\begin{aligned}&\mathbb{E}\left[\log f_{G_t|\boldsymbol{\Phi}_t}(g \mid \mathbf{y}_t;\boldsymbol{\theta}) \mid \boldsymbol{\Phi}_t, \boldsymbol{\theta}^{[\ell]}\right] - \mathbb{E}\left[\log f_{G_t|\boldsymbol{\Phi}_t}\left(g \mid \mathbf{y}_t;\boldsymbol{\theta}^{[\ell]}\right) \mid \boldsymbol{\Phi}_t, \boldsymbol{\theta}^{[\ell]}\right] \\ &= \mathbb{E}\left[\log \frac{f_{G_t|\boldsymbol{\Phi}_t}(g \mid \mathbf{y}_t;\boldsymbol{\theta})}{f_{G_t|\boldsymbol{\Phi}_t}\left(g \mid \mathbf{y}_t;\boldsymbol{\theta}^{[\ell]}\right)} \,\middle|\, \boldsymbol{\Phi}_t, \boldsymbol{\theta}^{[\ell]}\right] \\ &\leq \log \mathbb{E}\left[\frac{f_{G_t|\boldsymbol{\Phi}_t}(g \mid \mathbf{y}_t;\boldsymbol{\theta})}{f_{G_t|\boldsymbol{\Phi}_t}\left(g \mid \mathbf{y}_t;\boldsymbol{\theta}^{[\ell]}\right)} \,\middle|\, \boldsymbol{\Phi}_t, \boldsymbol{\theta}^{[\ell]}\right] \\ &= \log \int \frac{f_{G_t|\boldsymbol{\Phi}_t}(g \mid \mathbf{y}_t;\boldsymbol{\theta})}{f_{G_t|\boldsymbol{\Phi}_t}\left(g \mid \mathbf{y}_t;\boldsymbol{\theta}^{[\ell]}\right)} f_{G_t|\boldsymbol{\Phi}_t}\left(g \mid \mathbf{y}_t;\boldsymbol{\theta}^{[\ell]}\right) \mathrm{d}g \\ &= \log \int f_{G_t|\boldsymbol{\Phi}_t}(g \mid \mathbf{y}_t;\boldsymbol{\theta})\, \mathrm{d}g = 0.\end{aligned}$$

Combining these three arguments proves (30) and establishes monotonicity of the ECME algorithm for model (1) with iid G_t's. In case of hybrid GARCH-SV dynamics, the ECME algorithm maximizes

$$\log L_{\mathbf{Y}|(\mathbf{G}|\mathbf{Y})}(\boldsymbol{\theta}) = \sum_{t=1}^{T} \log f_{\mathbf{Y}_t|(G_t|\boldsymbol{\Phi}_t),\boldsymbol{\Phi}_{t-1}}(\mathbf{y}_t \mid g_t;\boldsymbol{\theta}),$$

where g_t, for $t = 1, \ldots, T$, are fixed at each iteration to the values filtered in the E-step. However,

$$\begin{aligned}f_{\mathbf{Y}_t|\boldsymbol{\Phi}_{t-1}}(\mathbf{y};\boldsymbol{\theta}) &= \frac{f_{\mathbf{Y}_t|(G_t|\boldsymbol{\Phi}_t),\boldsymbol{\Phi}_{t-1}}(\mathbf{y} \mid g;\boldsymbol{\theta})\, f_{(G_t|\boldsymbol{\Phi}_t)|\boldsymbol{\Phi}_{t-1}}(g;\boldsymbol{\theta})}{f_{(G_t|\boldsymbol{\Phi}_t)|\mathbf{Y}_t,\boldsymbol{\Phi}_{t-1}}(g \mid \mathbf{y};\boldsymbol{\theta})} \\ &= f_{\mathbf{Y}_t|(G_t|\boldsymbol{\Phi}_t),\boldsymbol{\Phi}_{t-1}}(\mathbf{y} \mid g;\boldsymbol{\theta}),\end{aligned} \tag{32}$$

for all $g > 0$ and $\mathbf{y}$, because $(G_t \mid \boldsymbol{\Phi}_t) \mid \boldsymbol{\Phi}_{t-1}$ and $(G_t \mid \boldsymbol{\Phi}_t) \mid \mathbf{Y}_t, \boldsymbol{\Phi}_{t-1}$ are the same random variables. Therefore, the algorithm also maximizes $\log L_{\mathbf{Y}}(\boldsymbol{\theta})$.

Appendix B. Link to the Taylor (1982) SV model

The univariate model proposed in Taylor (1982) is given by

$$Y_t = \exp(Q_t/2)\, Z_t, \quad \text{with } Q_t = c + \rho Q_{t-1} + \sigma \eta_{t-1}, \tag{33}$$

for $Z_t \overset{\text{iid}}{\sim} N(0, 1)$; $\eta_t \overset{\text{iid}}{\sim} N(0, 1)$; and Z_t and η_t are independent. Define Q_t by $G_t = \exp(Q_t)$. Then (1) can be rewritten as

$$\mathbf{Y}_t = \boldsymbol{\mu} + \boldsymbol{\gamma} \exp(Q_t) + \mathbf{H}_t^{1/2} \exp(Q_t/2)\, \mathbf{Z}_t. \tag{34}$$

Switching from $(G_t \mid \boldsymbol{\Phi}_{t-1})$ dynamics in (4) to $(Q_t \mid \boldsymbol{\Phi}_{t-1})$ dynamics, we get

$$\mathbb{E}\left[Q_t^r \mid \boldsymbol{\Phi}_{t-1}\right] = c_r + \rho_r \mathbb{E}\left[Q_{t-1}^r \mid \boldsymbol{\Phi}_{t-2}\right] + \tilde{\zeta}_{r,t}, \tag{35}$$

where $\tilde{\zeta}_{r,t} = \mathbb{E}\left[Q_t^r \mid \boldsymbol{\Phi}_t\right] - \mathbb{E}\left[Q_t^r \mid \boldsymbol{\Phi}_{t-1}\right]$. Setting $\boldsymbol{\mu} = \mathbf{0}$, $\boldsymbol{\gamma} = \mathbf{0}$, $r = 1$ and $\mathbf{H}_t = 1$ in (34) (for $K = 1$) results in (34) reduce to (33), with η_{t-1} replaced by $\tilde{\zeta}_{1,t}$ and $\mathbb{E}[Q_t \mid \boldsymbol{\Phi}_{t-1}]$ instead of Q_t.

Our model differs from the SV model in three additional aspects. Firstly, we replace the past shock η_{t-1} by the current shock η_t in (33). In the SV literature, one has to use a lag shock (together with $\mathrm{Corr}(Z_t, \eta_t) < 0$) to obtain the key feature of SV models, the asymmetric return-volatility relation, often called a statistical leverage effect. In our setup, we can incorporate the asymmetry or the leverage effect through the scale-term dynamics.

Note that, if we were to use $\zeta_{r,t-1}$ instead of $\zeta_{r,t}$ in our model, then there would be a one-period shift between the filtered $\widehat{G}_t$ values, and the moments of G_t conditional on $\boldsymbol{\Phi}_{t-1}$. Use of $\zeta_{r,t}$ avoids this and allows for shocks to the volatility to have an immediate impact on returns, a feature which is absent in discrete time SV models.

The second difference is: we work with G_t instead of $\log G_t$ because the former has tractable moment expressions, and we can still guarantee positive values of $\mathbb{E}[G_t \mid \boldsymbol{\Phi}_{t-1}]$ by imposing the constraints $c_r > 0$ and $\rho_r \geq 0$.

Thirdly, the dynamics in our model are written in terms of the conditional moments $\mathbb{E}[G_t \mid \boldsymbol{\Phi}_{t-1}]$ and not in terms of G_t itself because keeping the dynamics of G_t only in terms of conditional expectations allows us to maintain, without any extra conditions, the monotonic increase of the incomplete-data likelihood which is a key property of the ECME algorithm, and because we are able to filter out $\mathbb{E}[G_t \mid \boldsymbol{\Phi}_t]$ through our ECME algorithm without an additional computational burden.

Appendix C. Derivation of the Q^{+G}-dynamics for option pricing

Following Christoffersen et al. (2010) and a multivariate extension given in Rombouts and Stentoft (2011), we impose the exponential affine form on the Radon–Nikodym derivative, with respect to $\mathcal{F}_t^{+G}$. Hence, by the law of iterative expectation, we get:

Lemma C.1. *For any K-dimensional sequence $\mathbf{v}_s$,*

$$\frac{dQ^{+G}}{dP} \mid \mathcal{F}_t^{+G} = \exp\left(-\sum_{s=1}^{t}\left(\mathbf{v}_s'\boldsymbol{\varepsilon}_s + \frac{1}{2}\mathbf{v}_s'\mathbf{H}_s\mathbf{v}_s G_s\right)\right) \tag{36}$$

is a Radon–Nikodym derivative with respect to $\mathcal{F}_t^{+G}$.

Proof. We need to show that $\frac{dQ^{+G}}{dP} \mid \mathcal{F}_t^{+G} > 0$ and $\mathbb{E}_0^P\left[\frac{dQ^{+G}}{dP} \mid \mathcal{F}_t^{+G}\right] = 1$. Non-negativity is an immediate consequence of the exponential form. For the second condition we use the law of iterated expectations, with respect to $\mathcal{F}_t^{+G}$, to obtain

$$\begin{aligned}
\mathbb{E}_0^P\left[\frac{dQ^{+G}}{dP}\,\middle|\,\mathcal{F}_t^{+G}\right] &= \mathbb{E}_0^P\left[\exp\left(-\sum_{s=1}^{t}\left(\mathbf{v}_s'\boldsymbol{\varepsilon}_s + \frac{1}{2}\mathbf{v}_s'\mathbf{H}_s\mathbf{v}_s G_s\right)\right)\right]\\
&= \mathbb{E}_0^P\left[\mathbb{E}_1^P,\ldots,\mathbb{E}_{t-1}^P \exp\left(\sum_{s=1}^{t}\left(-\mathbf{v}_s'\boldsymbol{\varepsilon}_s\right) - \sum_{s=1}^{t}\frac{1}{2}\mathbf{v}_s'\mathbf{H}_s\mathbf{v}_s G_s\right)\right]\\
&= \mathbb{E}_0^P\Bigg[\mathbb{E}_1^P,\ldots,\mathbb{E}_{t-2}^P\\
&\quad\times \exp\left(\sum_{s=1}^{t-1}\left(-\mathbf{v}_s'\boldsymbol{\varepsilon}_s\right) - \sum_{s=1}^{t}\frac{1}{2}\mathbf{v}_s'\mathbf{H}_s\mathbf{v}_s G_s\right)\mathbb{E}_{t-1}^P \exp\left(-\mathbf{v}_t'\boldsymbol{\varepsilon}_t\right)\Bigg]\\
&= \mathbb{E}_0^P\Bigg[\mathbb{E}_1^P,\ldots,\mathbb{E}_{t-2}^P\\
&\quad\times \exp\left(\sum_{s=1}^{t-1}\left(-\mathbf{v}_s'\boldsymbol{\varepsilon}_s\right) - \sum_{s=1}^{t}\frac{1}{2}\mathbf{v}_s'\mathbf{H}_s\mathbf{v}_s G_s\right)\mathbb{E}_{t-1}^P \exp\left(\frac{1}{2}\mathbf{v}_t'\mathbf{H}_t\mathbf{v}_t G_t\right)\Bigg]\\
&= \mathbb{E}_0^P\left[\mathbb{E}_1^P,\ldots,\mathbb{E}_{t-2}^P \exp\left(\sum_{s=1}^{t-1}\left(-\mathbf{v}_s'\boldsymbol{\varepsilon}_s\right) - \sum_{s=1}^{t-1}\frac{1}{2}\mathbf{v}_s'\mathbf{H}_s\mathbf{v}_s G_s\right)\right],
\end{aligned}$$

where the second last equality follows from the normality of $\boldsymbol{\varepsilon}_t$ conditional on $\mathcal{F}_{t-1}^{+G}$. Iterating on this yields the required result. □

Having a valid candidate for the change of measure, we proceed to find conditions for the sequence $\mathbf{v}_s$ under which the proposed Radon–Nikodym derivative defines an EMM under $\mathcal{F}_t^{+G}$. Denote by $\mathbf{r} = (r, r, \ldots, r)'$ a K vector of risk free interest rates, then

Proposition C.1. *The probability measure Q^{+G} defined by the Radon–Nikodym derivative in (36) is an EMM under $\mathcal{F}_t^{+G}$ if and only if*

$$\mathbf{v}_t = \mathbf{H}_t^{-1}\left(\left(\boldsymbol{\mu} - \mathbf{r}\right)G_t^{-1} + \boldsymbol{\gamma} + \frac{1}{2}\mathrm{diag}\left(\mathbf{S}_t^2\right)\right).$$

Proof. We need to show that, for all $k = 1, 2, \ldots, K$, $\mathbb{E}^{Q^{+G}}\left[\frac{X_{t,k}}{X_{t-1,k}} \mid \mathcal{F}_{t-1}^{+G}\right] = \exp(r)$, where r is the risk free interest rate, and $X_{t,k}$ is the price of stock k at time t. We have

$$\begin{aligned}
&\mathbb{E}^{Q^{+G}}\left[\frac{X_{t,k}}{X_{t-1,k}}\exp(-r) \mid \mathcal{F}_{t-1}^{+G}\right]\\
&= \mathbb{E}^P\left[\left(\frac{\frac{dQ^{+G}}{dP} \mid \mathcal{F}_t^{+G}}{\frac{dQ^{+G}}{dP} \mid \mathcal{F}_t}\right)\left(\frac{\frac{dQ^{+G}}{dP} \mid \mathcal{F}_t}{\frac{dQ^{+G}}{dP} \mid \mathcal{F}_{t-1}^{+G}}\right)\frac{X_{t,k}}{X_{t-1,k}}\exp(-r) \mid \mathcal{F}_{t-1}^{+G}\right]\\
&= \mathbb{E}^P\left[\left(\frac{\frac{dQ^{+G}}{dP} \mid \mathcal{F}_t^{+G}}{\frac{dQ^{+G}}{dP} \mid \mathcal{F}_{t-1}^{+G}}\right)\frac{X_{t,k}}{X_{t-1,k}}\exp(-r) \mid \mathcal{F}_{t-1}^{+G}\right]\\
&= \mathbb{E}^P\Bigg[\exp\left(-\mathbf{v}_t'\boldsymbol{\varepsilon}_t - \frac{1}{2}\mathbf{v}_t'\mathbf{H}_t\mathbf{v}_t G_t\right)\\
&\qquad \times \exp\left(\mu_k + \gamma_k G_t + \varepsilon_{t,k}\right)\exp(-r) \mid \mathcal{F}_{t-1}^{+G}\Bigg]\\
&= \mathbb{E}^P\Bigg[\exp\Bigg(-\frac{1}{2}\mathbf{v}_t'\mathbf{H}_t\mathbf{v}_t G_t + \mu_k + \gamma_k G_t\\
&\qquad - r - \mathbf{v}_t'\boldsymbol{\varepsilon}_t + \mathbf{e}_k'\boldsymbol{\varepsilon}_t\Bigg) \mid \mathcal{F}_{t-1}^{+G}\Bigg]\\
&= \exp\Bigg(-\frac{1}{2}\mathbf{v}_t'\mathbf{H}_t\mathbf{v}_t G_t + \mu_k + \gamma_k G_t - r\\
&\qquad + \frac{1}{2}\left(\mathbf{v}_t - \mathbf{e}_k\right)'\mathbf{H}_t\left(\mathbf{v}_t - \mathbf{e}_k\right)G_t\Bigg)\\
&= \exp\left(-\mathbf{e}_k'\mathbf{H}_t\mathbf{v}_t G_t + \frac{1}{2}\mathbf{e}_k'\mathbf{H}_t\mathbf{e}_k G_t + \mu_k + \gamma_k G_t - r\right),
\end{aligned}$$

where $\mathbf{e}_k = (0, \ldots, 0, 1, 0, \ldots, 0)'$ is a vector of zeros with one at position k. Thus, if we ensure that

$$-\mathbf{e}_k'\mathbf{H}_t\mathbf{v}_t G_t + \frac{1}{2}\mathbf{e}_k'\mathbf{H}_t\mathbf{e}_k G_t + \mu_k + \gamma_k G_t - r = 0$$

for all $k = 1, 2, \ldots, K$, by choosing the vector series $\mathbf{v}_t$, then the probability measure Q^{+G} is an EMM since it makes discounted asset prices martingales. Solving it for $\mathbf{v}_t$, in vector form we obtain

$$\mathbf{v}_t = \mathbf{H}_t^{-1}\left(\left(\boldsymbol{\mu} + \boldsymbol{\gamma}G_t - \mathbf{r}\right)G_t^{-1} + \frac{1}{2}\mathrm{diag}\left(\mathbf{S}_t^2\right)\right). \quad \square$$

Thus, if we knew the realizations of G_s for $s = 1, 2, \ldots, t$, then it is possible to solve explicitly for the respective $\mathbf{v}_s$, and Proposition C.1 guarantees that the corresponding measure is an EMM under $\mathcal{F}_t^{+G}$.

Use of the $\mathcal{F}_t^{+G}$ filtration implies that, although we are working within an incomplete market framework, there is only one source of randomness. Therefore, constraining the Radon–Nikodym derivative to be of the exponential affine form, as in (36), allows us to derive a unique measure under which the discounted asset prices are Q^{+G}-martingales. Moreover, because $\mathbf{Y}_t \mid G_t$ is Gaussian, we can characterize the change of measure and the Q^{+G}-dynamics corresponding to model (1) explicitly. Denote by $\Psi_t(\mathbf{u})$ the logarithm of the conditional moment generating function of $\boldsymbol{\varepsilon}_t$ given $\mathcal{F}_{t-1}^{+G}$, i.e.,

$$\Psi_t(\mathbf{u}) = \frac{1}{2}\mathbf{u}'\mathbf{H}_t\mathbf{u}G_t. \tag{37}$$

In order to obtain the Q^{+G}-dynamics of our model, we derive the analogue of Ψ_t under Q.

Corollary C.1. *The logarithm of the conditional (on $\mathcal{F}_{t-1}^{+G}$) moment generating function of $\boldsymbol{\varepsilon}_t$ under Q^{+G} is given by*

$$\begin{aligned}\Psi_t^{Q^{+G}}(\mathbf{u}) &= \log \mathbb{E}^{Q^{+G}}\left[\exp\left(-\mathbf{u}'\boldsymbol{\varepsilon}_t\right) \mid \mathcal{F}_{t-1}^{+G}\right] \\ &= \Psi_t(\mathbf{v}_t+\mathbf{u}) - \Psi_t(\mathbf{v}_t).\end{aligned} \tag{38}$$

Proof. By change of measure and rearranging we get

$$\begin{aligned}&\mathbb{E}^{Q^{+G}}\left[\exp\left(-\mathbf{u}'\boldsymbol{\varepsilon}_t\right) \mid \mathcal{F}_{t-1}^{+G}\right] \\ &= \mathbb{E}^{P}\left[\left(\frac{\frac{dQ^{+G}}{dP} \mid \mathcal{F}_{t}^{+G}}{\frac{dQ^{+G}}{dP} \mid \mathcal{F}_{t-1}^{+G}}\right)\exp\left(-\mathbf{u}'\boldsymbol{\varepsilon}_t\right) \mid \mathcal{F}_{t-1}^{+G}\right] \\ &= \mathbb{E}^{P}\left[\exp\left(-\mathbf{v}_t'\boldsymbol{\varepsilon}_t - \Psi_t(\mathbf{v}_t)\right)\exp\left(-\mathbf{u}'\boldsymbol{\varepsilon}_t\right) \mid \mathcal{F}_{t-1}^{+G}\right] \\ &= \mathbb{E}^{P}\left[\exp\left(-(\mathbf{v}_t+\mathbf{u})'\boldsymbol{\varepsilon}_t - \Psi_t(\mathbf{v}_t)\right) \mid \mathcal{F}_{t-1}^{+G}\right] \\ &= \exp\left(\Psi_t(\mathbf{v}_t+\mathbf{u}) - \Psi_t(\mathbf{v}_t)\right).\end{aligned}$$

Taking log of both sides completes the proof. □

Substituting (37) into (38), we get

$$\Psi_t^{Q^{+G}}(\mathbf{u}) = \mathbf{u}'\mathbf{H}_t\mathbf{v}_tG_t + \frac{1}{2}\mathbf{u}'\mathbf{H}_t\mathbf{u}G_t.$$

Now, using the expression for $\mathbf{v}_t$ given in Proposition C.1,

$$\begin{aligned}\Psi_t^{Q^{+G}}(\mathbf{u}) &= \mathbf{u}'(\boldsymbol{\mu}-r) + \mathbf{u}'\boldsymbol{\gamma}G_t \\ &\quad + \frac{1}{2}\mathbf{u}'\operatorname{diag}\left(\mathbf{S}_t^2\right)G_t + \frac{1}{2}\mathbf{u}'\mathbf{H}_t\mathbf{u}G_t.\end{aligned} \tag{39}$$

So, the Q^{+G}-dynamics of the returns remain Gaussian with a shift in the mean, as in (24).

Appendix D. Option pricing calibration algorithm

The measure Q defined by the change of measure in (23) is an EMM, only if the realizations of the common market factor are observed. In order to price options we need to incorporate the latency of G_t. For this purpose we combine the dynamics under EMM Q from (24) with a calibration of the common market factor parameters in $\boldsymbol{\theta}_D^{Q^*}$ by a Monte Carlo simulation. The estimates of risk neutral parameters in $\boldsymbol{\theta}_D^{Q^*}$ are obtained by minimizing the mean square error between prices of the options observed on the market C_t^m, and the corresponding simulated option prices $C_t^{(N)}$, i.e., the optimal $\widehat{\boldsymbol{\theta}}_D^{Q^*}$ satisfies

$$\begin{aligned}\widehat{\boldsymbol{\theta}}_D^{Q^*} &= \arg\min_{\boldsymbol{\theta}_D^{Q^*}} \frac{1}{M}\sum_{m=1}^{M}\Big\{C_t^{(N)}\left(T, \varrho_{\vartheta_m}, \widehat{\boldsymbol{\theta}}_P^{Q}, \widehat{\boldsymbol{\theta}}_C^{Q}, \boldsymbol{\theta}_D^{Q^*}\right) \\ &\quad - C_t^m\left(T, \varrho_{\vartheta_m}\right)\Big\}^2,\end{aligned} \tag{40}$$

where

$$\begin{aligned}&C_t^{(N)}\left(T, \varrho_{\vartheta_m}, \widehat{\boldsymbol{\theta}}_P^{Q}, \widehat{\boldsymbol{\theta}}_C^{Q}, \boldsymbol{\theta}_D^{Q^*}\right) \\ &= \frac{1}{N}\exp\left(-r(T-t)\right)\sum_{n=1}^{N}\varrho_{\vartheta_m}\left(\mathbf{X}_T^{(n)}\left(\widehat{\boldsymbol{\theta}}_P^{Q}, \widehat{\boldsymbol{\theta}}_C^{Q}, \boldsymbol{\theta}_D^{Q^*}\right)\right),\end{aligned} \tag{41}$$

is an estimate of the option price based on N simulations from the COMFORT model, with $\mathbf{X}_T^{(n)}\left(\widehat{\boldsymbol{\theta}}_P^{Q}, \widehat{\boldsymbol{\theta}}_C^{Q}, \boldsymbol{\theta}_D^{Q^*}\right)$ being the nth simulated vector of spot prices at the option maturity. In the simulation we use the Q-measure estimates of the $\boldsymbol{\theta}_P^{Q}$ and $\boldsymbol{\theta}_C^{Q}$ parameters obtained as in Appendix C, with unobserved future values of $G_{t+s}^{Q^*}$, for $s = 1, \ldots, T-t$ simulated according to the model dynamics. For the GARCH case we draw a random sample from $\mathrm{GIG}\left(\lambda^{Q^*}, \chi^{Q^*}, \psi^{Q^*}\right)$, and for the hybrid GARCH-SV extension we use $\mathbb{E}^P\left[G_t^r \mid \boldsymbol{\Phi}_t\right]$ as a starting value and, for $s > 0$, recursively recover the GIG parameters from

$$\mathbb{E}^{Q^*}\left[G_{t+s}^r \mid \boldsymbol{\Phi}_{t+s-1}\right] = c_r^{Q^*} + \rho_r^{Q^*}\mathbb{E}^{Q^*}\left[G_{t+s-1}^r \mid \boldsymbol{\Phi}_{t+s-2}\right], \tag{42}$$

by (43), and sample accordingly from $\mathrm{GIG}\left(\lambda_{t+s}^{Q^*}, \chi_{t+s}^{Q^*}, \psi_{t+s}^{Q^*}\right)$, for $s = 1, \ldots, T-t$. Finally, in order to guarantee that the price process satisfies the martingale restriction and to reduce the Monte Carlo simulation error we follow the Empirical Martingale Simulation (EMS) procedure by Duan and Simonato (1998).

Appendix E. Evaluation of the Bessel function

Let $G \sim \mathrm{GIG}(\lambda, \chi, \psi)$, for $\chi > 0$ and $\psi > 0$. Then it may be shown (see, e.g., Paolella, 2007, Ch. 9) that

$$\mathbb{E}\left[G^{\alpha}\right] = \left(\frac{\chi}{\psi}\right)^{\alpha/2}\frac{\mathcal{K}_{\lambda+\alpha}\left(\sqrt{\chi\psi}\right)}{\mathcal{K}_{\lambda}\left(\sqrt{\chi\psi}\right)}, \quad \alpha \in \mathbb{R}, \tag{43}$$

which involves a ratio of Bessel functions $\mathcal{K}_v(z)$ as given in (3).

It is possible to compute the limit of the Bessel function ratio for some cases. Let $\mathbf{Y} \mid (G = g) \sim \mathrm{N}(\boldsymbol{\mu} + \boldsymbol{\gamma}g, g\boldsymbol{\Sigma})$. We are interested in the expectations of $G^{\pm 1} \mid (\mathbf{Y} = \mathbf{y})$ which, for $\psi \neq 0$ or $\gamma \neq 0$, are always positive and have their limits given by

$$\begin{aligned}&\lim_{m\to\infty}\left(\frac{m+\chi}{\psi + \boldsymbol{\gamma}\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}'}\right)^{\pm 1/2} \\ &\times \frac{\mathcal{K}_{\lambda-K/2\pm 1}\left(\sqrt{(m+\chi)\left(\psi+\boldsymbol{\gamma}\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}'\right)}\right)}{\mathcal{K}_{\lambda-K/2}\left(\sqrt{(m+\chi)\left(\psi+\boldsymbol{\gamma}\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}'\right)}\right)} = 0,\end{aligned}$$

where $m = (\mathbf{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})$.

In our model, these ratios are responsible for proper weights, in the E-step and CM1-step of the ECME algorithm. We thus require a highly accurate approximation of Bessel function ratios for large v or z. This can be done by using the Watson (1922, p. 202) asymptotic expansion of $\mathcal{K}_v(z)$ given by

$$\mathcal{K}_v(z) = \sqrt{\frac{\pi}{2z}}\exp(-z)E(v,z), \tag{44}$$

where

$$E(v,z) = 1 + \sum_{k=1}^{\infty}\frac{\prod_{l=1}^{k}\left(4v^2-(2l-1)^2\right)}{k!\,(8z)^k}. \tag{45}$$

Inspection of (44) reveals that, for a ratio of Bessel functions, a numerically problematic $\exp(z)$ cancels out. In order to use (44), we have to truncate the series at some finite K which causes, for z small ($z < 10$), some loss of accuracy, but as z increases, the accuracy grows very rapidly because of z to the k power in terms of the series (45).

References

Aït-Sahalia, Y., Cacho-Diaz, J., Hurd, T.R., 2009. Portfolio choice with jumps: A closed-form solution. Ann. Appl. Probab. 19 (2), 556–584.

Aït-Sahalia, Y., Jacod, J., 2011. Testing whether jump have finite or infinite activity. Ann. Statist. 39 (3), 1689–1719.

Andersen, T.G., Bollerslev, T., Diebold, F.X., 2007. Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. Rev. Econ. Statist. 89 (4), 701–702.

Asai, M., McAleer, M., 2009. The structure of dynamic correlations in multivariate stochastic volatility models. J. Econometrics 150, 182–192.

Asai, M., McAleer, M., Yu, J., 2006. Multivariate stochastic volatility: A review. Econometric Rev. 25, 145–175.

Barone-Adesi, G., Engle, R.F., Mancini, L., 2008. A GARCH option pricing model with filtered historical simulations. Rev. Financ. Stud. 21 (3), 1223–1258.

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

Bollerslev, T., 1990. Modeling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH approach. Rev. Econ. Statist. 72, 498–505.

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

Bollerslev, T., Todorov, V., Li, S.Z., 2013. Jump tails, extreme dependencies, and the distribution of stock returns. J. Econometrics 172 (2), 307–324.

Bos, C.S., 2012. In: Bauwens, L., Hafner, C., Laurent, S. (Eds.), Relating Stochastic Volatility Estimation Methods. In: Handbook of Volatility Models and Their Applications, John Wiley & Sons Inc., pp. 147–175.

Carr, P., Geman, H., Madan, D.B., Yor, M., 2002. The fine structure of asset returns: An empirical investigation. J. Bus. 75 (2), 305–333.

Chan, W.H., Maheu, J.M., 2002. Conditional jump dynamics in stock market returns. J. Bus. Econom. Statist. 20 (3), 377–389.

Chernov, M., Gallant, A.R., Ghysels, E., Tauchen, G., 2003. Alternative models for stock price dynamics. J. Econometrics 116 (3), 225–257.

Christoffersen, P., Elkamhi, R., Feunou, B., Jacobs, K., 2010. Option valuation with conditional heteroskedasticity and nonnormality. Rev. Financ. Stud. 23 (5), 2139–2183.

Christoffersen, P., Heston, S., Jacobs, K., 2006. Option valuation with conditional skewness. J. Econometrics 131, 253–284.

Duan, J.-C., 1995. The GARCH option pricing model. Math. Finance 5 (1), 13–32.

Duan, J.-C., Simonato, J.-G., 1998. Empirical martingale simulation for asset prices. Manag. Sci. 44 (9), 1218–1233.

Eberlein, E., Keller, U., 1995. Hyperbolic distributions in finance. Bernoulli 1, 281–299.

Eberlein, E., Keller, U., Prause, K., 1998. New insights into smile, mispricing, and value at risk: the hyperbolic model. J. Bus. 38, 371–405.

Engle, R.F., 2002. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. J. Bus. Econom. Statist. 20, 339–350.

Engle, R.F., Sheppard, K., 2002. Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH. SSRN-id:1296441.

Eraker, B., 2004. Do stock prices and volatility jump? Reconciling evidence from spot and option prices. J. Finance 59 (3), 1367–1403.

Eraker, B., Johannes, M., Polson, N., 2003. The impact of jumps in volatility and returns. J. Finance 58 (3), 1269–1300.

Gatheral, J., 2006. The Volatility Surface: A Practitioners Guide. John Wiley & Sons.

Gilder, D., Shackleton, M.B., Taylor, S.J., 2014. Cojumps in stock prices: Empirical evidence. J. Bank. Finance 40 (C), 443–459.

Heston, S.L., 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. Rev. Financ. Stud. 6 (2), 327–343.

Heston, S.L., Nandi, S., 2000. A closed-form GARCH option valuation model. Rev. Financ. Stud. 13 (3), 585–625.

Liu, C., Rubin, D.B., 1994. The ECME algorithm: A simple extension of EM and ECM with faster monotone convergence. Biometrika 81 (4), 633–648.

Madan, D.B., Carr, P.P., Chang, E.C., 1998. The variance gamma process and option pricing. Eur. Finance Rev. 2 (1), 79–105.

Madan, D.B., Seneta, E., 1990. The variance gamma (V. G.) model for share market returns. J. Bus. 63 (4), 511–524.

Maheu, J.M., McCurdy, T.H., 2004. News arrival, jump dynamics, and volatility components for individual stock returns. J. Finance 59 (2), 755–793.

McLachlan, G.J., Krishnan, T., 2008. The EM Algorithm and Extensions, second ed. John Wiley & Sons, Hoboken, New Jersey.

McNeil, A.J., Frey, R., Embrechts, P., 2005. Quantitative Risk Management: Concepts, Techniques and Tools. Princton University Press.

Mencía, J., Sentana, E., 2009. Multivariate location-scale mixtures of normals and mean–variance-skewness portfolio allocation. J. Econometrics 153 (2), 105–121.

Paolella, M.S., 2007. Intermediate Probability: A Computational Approach. Wiley–Interscience.

Paolella, M.S., 2013. Multivariate asset return prediction with mixture models. Eur. J. Finance.

Paolella, M.S., Polak, P., 2015a. ALRIGHT: Asymmetric LaRge-Scale (I)GARCH with Hetero-Tails. Int. Rev. Econ. Finance (forthcoming).

Paolella, M.S., Polak, P., 2015b. Portfolio Selection with Active Risk Monitoring. Research Paper, Swiss Finance Institute.

Pelletier, D., 2006. Regime switching for dynamic correlations. J. Econometrics 131, 445–473.

Protassov, R., 2004. EM-based maximum likelihood parameter estimation for multivariate generalized hyperbolic distributions with fixed lambda. Stat. Comput. 14 (1), 67–77.

Rombouts, J.V., Stentoft, L., 2011. Multivariate option pricing with time varying volatility and correlations. J. Bank. Finance 35, 2267–2281.

Scott, D.J., Würtz, D., Dong, C., Tran, T.T., 2011. Moments of the generalized hyperbolic distribution. Comput. Statist. 26, 459–476.

Taylor, S.J., 1982. Financial Returns Modelled by the Product of Two Stochastic Processes—a Study of Daily Sugar Prices 1961–79. In: Time Series Analysis: Theory and Practice, vol. 1. North-Holland, Amsterdam, pp. 203–226.

Todorov, V., Tauchen, G., 2011. Volatility jumps. Am. Stat. Assoc. 29 (3), 356–371.

Tse, Y.K., Tsui, A.K. C., 2002. A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. J. Bus. Econom. Statist. 20 (3), 351–362.

Watson, G.N., 1922. A Treatise on the Theory of Bessel Functions. Cambridge University Press, Cambridge.