



Institut canadien des dérivés
Canadian Derivatives Institute

L'Institut bénéficie du soutien financier de l'Autorité des marchés financiers ainsi que du ministère des Finances du Québec

Document de recherche

DR 19-14

Affine Multivariate GARCH Models

Publié Septembre 2019

Ce document de recherche a été rédigée par :

Marcos Escobar, University of Western Ontario
Javad Rastegari, University of Western Ontario
Lars Stentoft, University of Western Ontario

Affine Multivariate GARCH Models^{*}

Marcos Escobar, Javad Rastegari

*Department of Statistical and Actuarial Sciences, University of Western Ontario, London, ON, Canada,
N6A5B7*

Lars Stentoft

*Department of Statistical and Actuarial Sciences, University of Western Ontario, London, ON, Canada,
N6A5B7*

*Department of Economics, University of Western Ontario, Social Science Centre, London, ON, Canada,
N6A5C2*

Abstract

This paper introduces a class of Affine multivariate GARCH models. Our setting offers flexibility to accommodate stylized facts of asset returns like dynamic conditional correlation and a covariance dependent pricing kernel. The model admits a closed-form recursive representation for the moment generating function under both historical and risk-neutral measures, permitting efficient multi-asset option pricing and risk management calculations. We illustrate the applicability and impact of our framework on the five assets for which volatility indices are made publicly available, together with the S&P500. We demonstrate that our methodology can be remarkably faster than Monte Carlo simulation when pricing two-assets options. We confirm the importance of incorporating a covariance-dependent kernel compared to a linear kernel by reporting large and economically significant changes in the price of two-asset options. Similarly, our single-factor Index model structure for the marginal can lead to differences of up to 15% in the price of single-asset options.

Keywords: Multivariate GARCH models, Closed-form solutions, Covariance-dependent pricing kernel, Multi-asset derivative pricing

JEL: C15, G12, G13

^{*}Financial support from Canadian Derivatives Institute is acknowledged.

Email addresses: `marcos.escobar@uwo.ca` (Marcos Escobar), `jrastega@uwo.ca` (Javad Rastegari), `lars.stentoft@uwo.ca`, corresponding author (Lars Stentoft)

1. Introduction

This paper proposes and studies a general class of multivariate GARCH models built as linear combinations of one-dimensional Affine GARCH models. This construction is rooted in structures created via multivariate methods like factor analysis (FA) or principal component analysis (PCA). Both methods are very popular in finance for good reasons: they help in reducing the dimensionality hence controlling the size of the parametric space; they are very intuitive and economically sound (see Fama and French (2015)), and they can be conceived using well-known one-dimensional models as underlyings. There is a vast literature on FA and PCA GARCH models, see for instance Vrontos, Dellaportas, and Politis (2003) and Van der Weide (2002), and for a survey on multivariate GARCH models in general, see Bauwens, Laurent, and Rombouts (2006). The majority of the factor-driven models are non-Affine though and hence using them for option pricing and risk management is complicated typically requiring expensive simulation methods to be used. A recent exception is the work of Bégin, Dorion, and Gauthier (2019) where the authors extended the one-dimensional Levy jump affine model of Ornathanalai (2014) to a bivariate setting deriving closed form option pricing formulas. The aforementioned extension was set up in the form of a pairwise structure with essentially independent innovations, hence in the spirit of a Capital Asset Pricing model. In order to gain in flexibility and capture richer dependence structures, we create and study a full Affine multivariate model.

Our model construction offers critical flexibility to accommodate stylized facts of financial asset returns; in particular our model is the first affine multivariate GARCH model exhibiting covariance-dependent pricing kernel. This extends the work of Christoffersen, Heston, and Jacobs (2013), where the authors accommodate variance-dependent kernels (and therefore variance risk premium) in a one-dimensional affine GARCH model. The presence of correlation risk premium, in addition to variance risk premium, has been documented in the work of Driessen, Maenhout, and Vilkov (2009) and Buss and Vilkov (2012), among others. Moreover, our model not only leads to a dynamic conditional correlation structure, but

also allows for multiple volatility drivers which has been proven relevant in continuous-time settings to match the dynamics of implied volatility surfaces, see Christoffersen, Heston, and Jacobs (2009).¹ These benefits gain special relevance in the context of affine models as they admit closed-form recursive representation for the moment generating function (mgf). The efficient calculation of the mgf lead to quick multi-asset option pricing, speedy calibration to observable prices and quick risk management evaluations of large portfolios.

The literature on Affine one-dimensional models is relatively modest but it has evolved rapidly since the seminal work of Heston and Nandi (2000) with Gaussian innovations and linear pricing kernels. For example, Christoffersen, Heston, and Jacobs (2006) addressed the issue of non-gaussianity by considering an affine Inverse-Gaussian GARCH model, Orn-thanalai (2014) considered an affine Levy GARCH model as a relevant case, and Badescu, Cui, and Ortega (2018) characterized the conditions to have affine one-dimensional structures. Babaoğlu, Christoffersen, Heston, and Jacobs (2018) combined several key stylized facts in one-dimension for an integral treatment. Our modelling approach can be used to extend this body of literature to the multivariate situation.

With the objective of illustrating the feasibility of our model, we perform a full estimation/calibration exercise on the S&P 500 and its volatility index (VIX) jointly with each of the five assets for which volatility indices are reported by the Chicago Board Options Exchange. The assets are Apple, Amazon, Google, Goldman Sachs and IBM, each one with their corresponding volatility indices. We estimate the parameters in a bivariate Cholesky HN-MGARCH using daily returns and volatility index prices. We demonstrate the efficiency of our affine methodology in the pricing of two-assets options in a comparison with a Monte Carlo simulation approach. We also illustrate the importance of the covariance-dependent kernels in the pricing of two-asset options, and the relevance of combining multiple volatility drivers with the correlation to an index, in the pricing of single-asset derivatives.

¹Our modeling framework can capture more advanced dependence-related stylized facts like tail dependence (see Poon, Rockinger, and Tawn (2003)) and co-volatility movements. This will be explicitly covered in future research as part of a non-Gaussian estimation exercise.

We summarize our contributions as:

1. We introduce and study a general class of affine multivariate GARCH models. We use a Cholesky HN-MGARCH as the main example for the empirical study. This general class admits closed-form joint mgf's under both physical and risk-neutral measures.
2. This is the first paper in the affine literature that consider changes of measure with a full covariance-dependent pricing kernel, including variance-dependent marginals.
3. We highlight the richness in the dependence structure implied by the general model and derive the continuous time limits for a general HN-MGARCH case.
4. We estimate the bivariate Cholesky HN-MGARCH using asset returns and we calibrate the risk premiums using multiple volatility indexes. This is the first paper performing a successful multivariate calibration to multiple volatility indexes in the affine literature.
5. We benchmark the closed form solutions to existing methods demonstrating that the mgf-based methodology for the pricing of two-asset derivatives is up to 1,000 times faster than a Monte Carlo simulation approach.
6. We confirm the importance of incorporating a covariance-dependent kernel by reporting very large and economically important changes in the price of two-asset options.
7. We also show that the single-factor Index model structure for the marginal can lead to differences of up to 15% in the price of single-asset options.

The paper is structured as follows: Section 2 introduces the general multivariate affine factor GARCH, discusses its properties as well as a particular class, the Gaussian HN-MGARCH. Details about valid changes of measures are provided in Section 3 along with an analysis of the implications for risk neutral dynamics, in particular when it comes to the joint structure of volatility indexes. Section 4 reports the results from our empirical analysis to five individual stocks, detailing the estimation methodology and results, and assesses the impact of various key parameters on single-asset and two-assets derivatives as well as the efficiency of our method in a comparison to Monte Carlo. Section 5 concludes. All proofs can be found in Appendix A.

2. The model

We now present a general affine multivariate GARCH (MGARCH) model and demonstrate a number key properties and features of the model, in particular, the existence of a closed-form moment generating function obtained from a set of recursive equations. As our main example, we illustrate the multivariate generalization of Heston-Nandi GARCH model.

2.1. The affine multivariate GARCH model

Let r be the single-period risk-free rate, S_t be the vector of n asset spot prices, and define the vector of daily returns as $R_{t+1} = \log(S_{t+1}/S_t)$ with $R_{i,t+1}$ denoting the return of i -th asset over a single period. The affine MGARCH(1,1) model is defined as

$$R_{i,t} = r + \sum_{j=1}^n \lambda_{ij} h_{j,t} + \sum_{j=1}^n a_{ij} \varepsilon_{j,t} \quad (1)$$

$$h_{j,t} = g(h_{j,t-1}, \varepsilon_{j,t-1}), \quad (2)$$

or in vector form

$$R_t = r\mathbf{1} + \Lambda \Sigma_t \mathbf{1} + A \varepsilon_t,$$

where $\varepsilon_t | \mathcal{F}_{t-1}$ is an independent sequence of multivariate random variables with zero mean and a diagonal covariance matrix given by

$$\text{Var}(\varepsilon_t | \mathcal{F}_{t-1}) = \Sigma_t, \quad \Sigma_t = \text{diag}(h_{j,t}).$$

Here $A_{n \times n} = [a_{ij}]$ and $\Lambda_{n \times n} = [\lambda_{ij}]$ are matrices of parameters with A assumed invertible. As we will see later, the particular choice of innovations (ε_t) , covariance dynamics $g(h_{j,t-1}, \varepsilon_{j,t-1})$ and change of measure impose a structure on Λ .

One of the main advantages of the factor-like representation in the MGARCH model is that positive definiteness of the conditional covariance matrix is already satisfied without further parameter constraints in g . Even though the conditional variance matrices Σ_t is

diagonal, there exist correlation between various assets. The correlation is enforced by matrix A and the covariance matrix of returns over a single period is given by

$$H_t = \text{Var}(R_t|\mathcal{F}_{t-1}) = A \text{Var}(\varepsilon_t|\mathcal{F}_{t-1}) A' = A\Sigma_t A'.$$

The elements of H_t are given by $H_{ij,t} = \sum_{k=1}^n a_{ik}a_{jk}h_{k,t}$. One can assume further structure on matrix A . Namely, A can be an orthogonal matrix representing the spectral decomposition (PCA) of conditional covariance matrix H_t , or a lower triangular matrix representing the Cholesky decomposition.

To have an affine structure, one needs a specific covariance dynamics given a distribution of innovations. This can be formulated through a restriction on the joint cumulant generating function (cgf) of shocks and conditional covariances, as observed by Badescu, Cui, and Ortega (2018) in the univariate case. Let $C_{(\varepsilon_{j,t}, h_{j,t+1})|\mathcal{F}_{t-1}}(u, v)$ be the joint cgf of $(\varepsilon_{j,t}, h_{j,t+1})$, for $j = 1, \dots, n$, defined as

$$\exp(C_{(\varepsilon_{j,t}, h_{j,t+1})|\mathcal{F}_{t-1}}(u, v)) = E_{t-1}^{\mathbb{P}}[\exp(u\varepsilon_{j,t} + vh_{j,t+1})]. \quad (3)$$

Then we make the following assumption:

Assumption 2.1. *The joint cgf is in the form of*

$$C_{(\varepsilon_{j,t}, h_{j,t+1})|\mathcal{F}_{t-1}}(u, v) = \varphi_j(u, v) + \rho_j(u, v)h_{j,t}. \quad (4)$$

That is, $C_{(\varepsilon_{j,t}, h_{j,t+1})|\mathcal{F}_{t-1}}$ is affine in $h_{j,t}$ for all $j = 1, \dots, n$.

The first consequence of this assumption is that a closed form mgf exist for the Affine MGARCH(1,1) model as stated in the following proposition.

Proposition 2.2. *Let $m(\mathbf{u})$ be the joint mgf of X_T in the Affine MGARCH model in Equations (1) and (2). Under Assumption 2.1 on the conditional joint cgf of $(\varepsilon_{j,t}, h_{j,t+1})$ we*

have

$$m(u_1, \dots, u_n) = E_t[e^{\mathbf{u}'x_T}] = \exp\left(\sum_{j=1}^n u_j x_{j,t} + J_{T-t} + \sum_{j=1}^n K_{j,T-t} h_{j,t+1}\right), \quad (5)$$

where J_{T-t} and $K_{j,T-t}$ are given recursively by

$$J_k = J_{k-1} + r \sum_{j=1}^n u_j + \sum_{j=1}^n \varphi_j \left(\sum_{i=1}^n u_i a_{ij}, K_{j,k-1} \right) \quad (6)$$

$$K_{j,k} = \sum_{i=1}^n u_i \lambda_{ij} + \rho_j \left(\sum_{i=1}^n u_i a_{ij}, K_{j,k-1} \right), \quad j = 1, \dots, n, \quad (7)$$

for $k = 0, 1, \dots, T - t$ with initial values $J_0 = K_{1,0} = \dots = K_{n,0} = 0$.

Proof. See Appendix A for the proof. □

Proposition 2.2 automatically produces the marginal mgf, $m_i(u_i) = E_t[e^{u_i x_{i,T}}]$, by setting $u_j = 0$ for all $j \neq i$, with recursive equations given by

$$J_k = J_{k-1} + r u_i + \sum_{j=1}^n \varphi_j (u_i a_{ij}, K_{j,k-1}) \quad (8)$$

$$K_{j,k} = u_i \lambda_{ij} + \rho_j (u_i a_{ij}, K_{j,k-1}), \quad j = 1, \dots, n, \quad (9)$$

The marginal mgf is in particular useful for pricing single-asset options.²

The marginal conditional moments can be found by noticing that conditional on \mathcal{F}_{t-1} , innovation terms $\varepsilon_{j,t}$ are independent with zero mean and conditional variance $h_{j,t}$. The we

²More generally, one could easily obtain a similar recursive representation for the joint mgf of x_T and h_{T+1} , which can be used to price covariance-related derivatives such as correlation-swaps and variance-swaps.

have that

$$\mu_{i,t} = E_{t-1}[R_{i,t}] = r + \sum_{j=1}^n \lambda_{ij} h_{j,t} \quad (10)$$

$$E_{t-1} [(R_{i,t} - \mu_{i,t})^2] = \sum_{j=1}^n a_{ij}^2 h_{j,t}. \quad (11)$$

$$E_{t-1} [(R_{i,t} - \mu_{i,t})^3] = \sum_{j=1}^n a_{ij}^3 E_{t-1} [\varepsilon_{j,t}^3], \quad (12)$$

$$E_{t-1} [(R_{i,t} - \mu_{i,t})^4] = \sum_{j=1}^n a_{ij}^4 E_{t-1} [\varepsilon_{j,t}^4] + 6 \sum_{j \neq k} a_{ij}^2 a_{ik}^2 h_{j,t} h_{k,t}. \quad (13)$$

The conditional covariance and correlation are given by

$$\begin{aligned} E_{t-1} [(R_{i,t} - \mu_{i,t})(R_{j,t} - \mu_{j,t})] &= \sum_{k=1}^n a_{ik} a_{jk} h_{k,t} \\ \frac{E_{t-1} [(R_{i,t} - \mu_{i,t})(R_{j,t} - \mu_{j,t})]}{\sqrt{E_{t-1} [(R_{i,t} - \mu_{i,t})^2] E_{t-1} [(R_{j,t} - \mu_{j,t})^2]}} &= \frac{\sum_{k=1}^n a_{ik} a_{jk} h_{k,t}}{\sqrt{\sum_{k=1}^n \sum_{l=1}^n a_{il}^2 a_{jl}^2 h_{k,t} h_{l,t}}} \end{aligned}$$

In general, the dependence structure is as rich as the innovations $\varepsilon_{j,t}$ permit. To see this note that the conditional copula (see Sklar (1959) and Burtschell, Gregory, and Laurent (2009)) of the vector of returns R_t can be expressed in terms of the copula of linear combination of innovations, $Y_{i,t} = \sum_{j=1}^n a_{ij} \varepsilon_{j,t}$, as

$$\begin{aligned} C_{R,t-1}(u_1, \dots, u_n) &= P(R_{1,t} \leq F_{R_{1,t-1}}^{-1}(u_1), \dots, R_{n,t} \leq F_{R_{n,t-1}}^{-1}(u_n)) \\ &= P(Y_{1,t} \leq F_{Y_{1,t-1}}^{-1}(u_1), \dots, Y_{n,t} \leq F_{Y_{n,t-1}}^{-1}(u_n)) \\ &= C_{Y,t-1}(u_1, \dots, u_n). \end{aligned}$$

where $F_{x,t-1}^{-1}(u_1)$ is the inverse of the cumulative distribution function of x_t given all the information up to $t - 1$. Therefore a non-Gaussian framework like Student-t or inverse Gaussian innovations could lead to, for example, non-zero values of lower and upper tail dependence defined as $\lim_{u \rightarrow 1^-} \frac{C(u,u)+1-2u}{1-u}$ and $\lim_{u \rightarrow 0^+} \frac{C(u,u)}{u}$, respectively.

2.2. Example: The HN-MGARCH model

A particular interesting example is the multivariate extension of the well-known model of Heston and Nandi (2000). In this specification, which we refer to as the HN-MGARCH model, $\varepsilon_t = \sqrt{\Sigma_t} z_t$, with $z_t \stackrel{\mathbb{P}}{\sim} N(0, I_{n \times n})$, such that ε_t contains n independent shock terms with $\varepsilon_{j,t} = \sqrt{h_{j,t}} z_{j,t}$, for $j = 1, \dots, n$. The HN-MGARCH model is given by

$$R_t = r\mathbf{1} + \Lambda \Sigma_t \mathbf{1} + A \sqrt{\Sigma_t} z_t \quad (14)$$

$$h_{j,t} = \omega_j + \beta_j h_{j,t-1} + \alpha_j \left(z_{j,t-1} - \gamma_j \sqrt{h_{j,t-1}} \right)^2, \quad j = 1, \dots, n, \quad (15)$$

when we express the returns in vector form.

For the HN-MGARCH model it is easy to show that the joint cgf $C_{(\varepsilon_{j,t}, h_{j,t+1})|\mathcal{F}_{t-1}}(u, v)$ is given by³

$$\begin{aligned} \exp \left(C_{(\varepsilon_{j,t}, h_{j,t+1})|\mathcal{F}_{t-1}}(u, v) \right) &= E_{t-1}^{\mathbb{P}} \left[\exp (u \varepsilon_{j,t} + v h_{j,t+1}) \right] \\ &= E_{t-1}^{\mathbb{P}} \left[\exp \left(u \sqrt{h_{j,t}} z_{j,t} + v \omega_j + v (\beta_j + \alpha_j \gamma_j^2) h_t - 2v \alpha_j \gamma_j \sqrt{h_{j,t}} z_{j,t} + v \alpha_j z_{j,t}^2 \right) \right] \\ &= \exp (v \omega_j + v (\beta_j + \alpha_j \gamma_j^2) h_{j,t}) \times \exp \left(-\frac{1}{2} \log(1 - 2v \alpha_j) + \frac{(u - 2v \alpha_j \gamma_j)^2}{2(1 - 2v \alpha_j)} h_{j,t} \right). \end{aligned}$$

Hence $C_{(\varepsilon_{j,t}, h_{j,t+1})|\mathcal{F}_{t-1}}(u, v)$ is indeed of the affine form in Assumption 2.1 with

$$\varphi_j(u, v) = v \omega_j - \frac{1}{2} \log(1 - 2v \alpha_j) \quad (16)$$

$$\rho_j(u, v) = v (\beta_j + \alpha_j \gamma_j^2) + \frac{(u - 2v \alpha_j \gamma_j)^2}{2(1 - 2v \alpha_j)}. \quad (17)$$

³When z is a standard normal random variable, $E[\exp(az^2 + bz)] = \exp \left(-\frac{1}{2} \log(1 - 2a) + \frac{b^2}{2(1 - 2a)} \right)$.

The above gives the recursive equations as

$$J_k = J_{k-1} + r \sum_{i=1}^n u_i + \sum_{j=1}^n \omega_j K_{j,k-1} - \frac{1}{2} \sum_{j=1}^n \log(1 - 2K_{j,k-1}\alpha_j) \quad (18)$$

$$K_{j,k} = \sum_{i=1}^n u_i \lambda_{ij} + K_{j,k-1}(\beta_j + \alpha_j \gamma_j^2) + \frac{(\sum_{i=1}^n u_i a_{ij} - 2K_{j,k-1}\alpha_j \gamma_j)^2}{2(1 - 2K_{j,k-1}\alpha_j)}, \quad j = 1, \dots, n. \quad (19)$$

Note that when $n = 1$, $a_{11} = 1$ and $\lambda_{11} = \lambda$, and we recover the recursive equations in the univariate HN-GARCH model.

To obtain the continuous time limit of our HN-MGARCH model we start from a transformation of returns dynamics. In particular, Equation (14) can be written as

$$R_t = A \left(\bar{r}\mathbf{1} + \bar{\Lambda}\Sigma_t\mathbf{1} + \sqrt{\Sigma_t}z_t \right) = A\bar{R}_t, \quad (20)$$

where $\bar{\Lambda} = A^{-1}\Lambda$ and $\bar{r} = A^{-1}r$. One can see that \bar{R}_t is a vector of independent affine GARCH models with variances $h_{j,t}$ of the HN type but with slightly different mean specifications, where the k 'th element is given by

$$\begin{aligned} \bar{R}_{k,t} &= \bar{r}_k + \sum_{j=1}^n \bar{\lambda}_{kj} h_{j,t} + \sqrt{h_{k,t}} z_{k,t} \\ h_{j,t} &= \omega_j + \beta_j h_{j,t-1} + \alpha_j \left(z_{j,t-1} - \gamma_j \sqrt{h_{j,t-1}} \right)^2, \quad j = 1, \dots, n. \end{aligned}$$

Note that if $\beta_j + \alpha_j \gamma_j < 1$ for $j = 1, \dots, n$ and the matrix A is invertible then the HN-MGARCH model is stationary.

Following along the lines of Heston and Nandi (2000), we can compute the conditional mean and the conditional variance of all covariances ($i, j = 1, \dots, n$) and the conditional

correlation among all variances for an increment Δ . These are given by

$$\begin{aligned}
E_{t-\Delta} [H_{ij,t+\Delta}] &= \sum_{k=1}^n a_{ik} a_{jk} E_{t-\Delta} [h_{k,t+\Delta}] = \sum_{k=1}^n a_{ik} a_{jk} (\omega_k + \alpha_k + (\beta_k + \alpha_k \gamma_k^2) h_{k,t}) \\
Var_{t-\Delta} [H_{ij,t+\Delta}] &= \sum_{k=1}^n a_{ik}^2 a_{jk}^2 Var_{t-\Delta} [h_{k,t+\Delta}] = \sum_{k=1}^n a_{ik}^2 a_{jk}^2 \alpha_k^2 (2 + 4\gamma_k^2) h_{k,t} \\
Corr_{t-\Delta} [H_{ii,t+\Delta}, H_{jj,t+\Delta}] &= \frac{\sum_{k=1}^n a_{ik}^2 a_{jk}^2 Var_{t-\Delta} [h_{k,t+\Delta}]}{\sqrt{\sum_{k=1}^n a_{ik}^4 Var_{t-\Delta} [h_{k,t+\Delta}]} \sqrt{\sum_{k=1}^n a_{jk}^4 Var_{t-\Delta} [h_{k,t+\Delta}]}}. \tag{21}
\end{aligned}$$

Note that Equation (21) reflects the co-movement in the variance (see Diebold and Nerlove (1989)).

Let us define $v_{k,t} = \frac{h_{k,t}}{\Delta}$, $k = 1, \dots, n$. These stochastic processes are independent and given by

$$v_{k,t+\Delta} = \frac{\omega_k}{\Delta} + \beta_k v_{k,t} + \frac{\alpha_k}{\Delta} \left(z_{k,t} - \sqrt{\Delta} \gamma_k \sqrt{v_{k,t}} \right)^2, \tag{22}$$

where, using $\bar{R}_{k,t+\Delta} = \log \left(\frac{\bar{S}_{k,t+\Delta}}{\bar{S}_{k,t}} \right)$, we also have that

$$\log (\bar{S}_{k,t+\Delta}) = \log (\bar{S}_{k,t}) + \bar{r}_k + \sum_{j=1}^n \bar{\lambda}_{kj} v_{j,t} \Delta + \sqrt{v_{k,t} \Delta} z_{k,t}. \tag{23}$$

The pair in Equations (22) and (23) exhibits a correlation (leverage effect) driven by γ_k and given by

$$Corr_{t-\Delta} (v_{k,t+\Delta}, \log (\bar{S}_{k,t})) = \frac{-sign(\gamma_k) \sqrt{2\gamma_k^2 v_{k,t} \Delta}}{\sqrt{1 + 2\gamma_k^2 v_{k,t} \Delta}}.$$

With the specifications $\gamma_k(\Delta) = \frac{2}{\Delta \sigma_k} - \frac{\kappa_k}{\sigma_k}$, $\beta_k(\Delta) = 0$, $\alpha_k(\Delta) = \frac{\Delta^2 \sigma_k^2}{4}$ and $\omega_k(\Delta) = \left(\kappa_k \theta_k - \frac{\sigma_k^2}{4} \right) \Delta^2$, for $k = 1, \dots, n$, these converge weakly to processes similar to Heston (1993) except for a richer expected return structure given by

$$\begin{aligned}
d \log \bar{S}_{k,t} &= \left(\bar{r}_k + \sum_{j=1}^n \bar{\lambda}_{kj} v_{j,t} \right) dt + \sqrt{v_{k,t}} dz_{k,t} \\
dv_{j,t} &= \kappa_j (\theta_j - v_{j,t}) dt + \sigma_j \sqrt{v_{j,t}} dz_{j,t}, \quad j = 1, \dots, n.
\end{aligned}$$

Using $A\bar{\Lambda} = \Lambda$, it is simple to show that the multivariate process converges weakly to a continuous time process given by

$$d \log S_{i,t} = \left(r + \sum_{j=1}^n \lambda_{ij} v_{j,t} \right) dt + \sum_{j=1}^n a_{ij} \sqrt{v_{j,t}} dz_{j,t}, \quad (24)$$

$$dv_{j,t} = \kappa_j (\theta_j - v_{j,t}) dt + \sigma_j \sqrt{v_{j,t}} dz_{j,t}, \quad j = 1, \dots, n. \quad (25)$$

The marginals of this process resembles the multifactor volatility model in Christoffersen, Heston, and Jacobs (2009), while the full multivariate version coincides with that of De Col, Gnoatto, and Grasselli (2013) and Escobar, Ferrando, and Rubtsov (2017) under the assumption of perfect leverage effect.

3. Change of measure

Our proposed Affine MGARCH models can be used to price a variety of single-asset and multi-asset options. In this section we present the change of measure through a covariance dependent kernel, and we demonstrate that the general model maintains the affine structure under the equivalent martingale measure. We apply the change of measure to the Gaussian case and obtain the risk-neutral dynamics of the HN-MGARCH model. Then we provide further details of a bivariate Cholesky model where the first asset is readily considered as the market index. Finally, we study the volatility index (VIX) implied by this model.

3.1. Change of measure in affine MGARCH models

We now provide the multivariate generalization of the variance-dependent kernel introduced in Christoffersen, Heston, and Jacobs (2013).

Proposition 3.1. *Assume that asset returns follow the affine MGARCH model in Equations (1) and (2), and suppose Assumption 2.1 holds, that is the joint cgf of $(\varepsilon_{j,t}, h_{j,t+1})$, for*

$j = 1, \dots, n$, is of the form given in Equation (4). Let ξ_t and θ_t be vectors satisfying

$$\sum_{j=1}^n (-\varphi_j(\theta_{j,t}, \xi_{j,t}) + \varphi_j(\theta_{j,t} + a_{ij}, \xi_{j,t})) = 0 \quad (26)$$

$$-\rho_j(\theta_{j,t}, \xi_{j,t}) + \rho_j(\theta_{j,t} + a_{ij}, \xi_{j,t}) + \lambda_{ij} = 0, \quad j = 1, \dots, n, \quad (27)$$

for $t = 1, \dots, T$. Then the Radon-Nikodym derivative process

$$\left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_{\mathcal{F}_t} = \exp \left(\sum_{s=1}^t \sum_{j=1}^n \left(\theta_{j,s} \varepsilon_{j,s} + \xi_{j,s} h_{j,s+1} - C_{(\varepsilon_{j,s}, h_{j,s+1}) | \mathcal{F}_{s-1}}(\theta_{j,s}, \xi_{j,s}) \right) \right), \quad (28)$$

defines an equivalent martingale measure \mathbb{Q} .

Proof. See Appendix A for the proof. □

Remark 3.2. Note that the $2n$ parameter vector (θ, ξ) determines the market prices of risk for the factors (ε_t) and their variances (h_{t+1}) . We write $\theta_{j,t}$ and $\xi_{j,t}$ as possibly time-varying parameters since they may be time varying. This though does not occur in our HN-MGARCH model for example.

The following corollary shows that the joint cgf under \mathbb{Q} is affine in $h_{j,t}$, and therefore Proposition 2.2 remains valid under the risk-neutral dynamics.

Corollary 3.3. *Under the equivalent martingale measure in Proposition 3.1, the joint conditional cgf of $(\varepsilon_{j,t}, h_{j,t+1})$, for $j = 1, \dots, n$, is given by*

$$\begin{aligned} C_{(\varepsilon_{j,t}, h_{j,t+1}) | \mathcal{F}_{t-1}}^{\mathbb{Q}}(u, v) &= C_{(\varepsilon_{j,t}, h_{j,t+1}) | \mathcal{F}_{t-1}}(u + \theta_{j,t}, v + \xi_{j,t}) - C_{(\varepsilon_{j,t}, h_{j,t+1}) | \mathcal{F}_{t-1}}(\theta_{j,t}, \xi_{j,t}) \\ &= \varphi_j^{\mathbb{Q}}(u, v) + \rho_j^{\mathbb{Q}}(u, v) h_{j,t}, \end{aligned} \quad (29)$$

where

$$\varphi_j^{\mathbb{Q}}(u, v) = \varphi_j(u + \theta_{j,t}, v + \xi_{j,t}) - \varphi_j(\theta_{j,t}, \xi_{j,t}) \quad (30)$$

$$\rho_j^{\mathbb{Q}}(u, v) = \rho_j(u + \theta_{j,t}, v + \xi_{j,t}) - \rho_j(\theta_{j,t}, \xi_{j,t}). \quad (31)$$

Proof. See Appendix A for the proof. □

Remark 3.4. Let $\psi_{j,t}^{\mathbb{Q}}$ be the marginal cgf of j 'th shock $\varepsilon_{j,t}|\mathcal{F}_{t-1}$ under \mathbb{Q} . We have that

$$\psi_{j,t}^{\mathbb{Q}}(u) = \log E_{t-1}^{\mathbb{Q}} [e^{u\varepsilon_{j,t}}] = C_{(\varepsilon_{j,t}, h_{j,t+1})|\mathcal{F}_{t-1}}^{\mathbb{Q}}(u, 0) = \varphi_j^{\mathbb{Q}}(u, 0) + \rho_j^{\mathbb{Q}}(u, 0)h_{j,t}. \quad (32)$$

We can use this to obtain the distribution of shocks under the equivalent martingale measure.

We now apply the change of measure to HN-MGARCH model with φ_j and ρ_j given in Equations (16) and (17). Observe that $\varphi_j(u, v)$ does not depend on the first variable and hence Equation (26) is automatically satisfied. For each $i, j = 1, \dots, n$, Equation (27) gives

$$\theta_{j,t} = -\frac{1}{2}a_{ij} + 2\xi_{j,t}\alpha_j\gamma_j - \frac{\lambda_{ij}}{a_{ij}}(1 - 2\xi_{j,t}\alpha_j).$$

If we define

$$d_j = (1 - 2\xi_j\alpha_j)^{-1}, \quad b_j = d_j(\theta_j - 2\xi_j\alpha_j\gamma_j)$$

we obtain

$$\lambda_{ij} = -a_{ij}b_j - \frac{1}{2}a_{ij}^2d_j, \quad i, j = 1, \dots, n. \quad (33)$$

Therefore, the matrix Λ has the following structure

$$\Lambda = -AB - \frac{1}{2}A \circ AD, \quad (34)$$

where B and D are diagonal matrices with elements b_j and d_j defined above.

To obtain the distribution of $\varepsilon_{j,t}$ under \mathbb{Q} , we invoke Equations (30), (31) and (32) to get

$$\psi_{j,t}^{\mathbb{Q}}(u) = [\rho_j(u + \theta_j, \xi_j) - \rho_j(\theta_j, \xi_j)] h_{j,t} \quad (35)$$

$$= \frac{1}{2(1 - 2\xi_j\alpha_j)} [(u + \theta_j - 2\xi_j\alpha_j\gamma_j)^2 - (\theta_j - 2\xi_j\alpha_j\gamma_j)^2] h_{j,t} \quad (36)$$

$$= \frac{u^2}{2} d_j h_{j,t} + u b_j h_{j,t}. \quad (37)$$

This is the cgf of a normal random variable with mean $b_j h_{j,t}$ and variance $d_j h_{j,t}$. Therefore, under \mathbb{Q} , we can write $\varepsilon_{t,j}$ as $\varepsilon_{t,j} = \sqrt{d_j} \sqrt{h_{j,t}} z_{j,t}^* + b_j h_{j,t}$, where $z_{j,t}^* \stackrel{\mathbb{Q}}{\sim} N(0, 1)$. Equivalently, $z_{j,t} = \sqrt{d_j} z_{j,t}^* + b_j \sqrt{h_{j,t}}$, or in vector form $z_t = \sqrt{D} z_t^* + B \sqrt{\Sigma_t}$. Substituting we obtain the risk-neutral process given by

$$R_{i,t} = r - \sum_{j=1}^n \frac{1}{2} a_{ij}^2 h_{j,t}^* + \sum_{j=1}^n a_{ij} \sqrt{h_{j,t}^*} z_{j,t}^* \quad (38)$$

$$h_{j,t}^* = \omega_j^* + \beta_j h_{j,t-1}^* + \alpha_j^* \left(z_{j,t-1}^* - \gamma_j^* \sqrt{h_{j,t-1}^*} \right)^2, \quad j = 1, \dots, n, \quad (39)$$

where risk-neutral variables and parameters are defined as

$$h_{j,t}^* = d_j h_{j,t}, \quad \alpha_j^* = d_j^2 \alpha_j, \quad \omega_j^* = d_j \omega_j, \quad \gamma_j^* = \frac{\gamma_j - b_j}{d_j}. \quad (40)$$

The risk-neutral process in vector form is given by

$$R_t = r \mathbf{1} - \frac{1}{2} A \circ A \Sigma_t^* \mathbf{1} + A \sqrt{\Sigma_t^*} z_t^*,$$

from which it is straightforward to see that expected asset returns equals the risk free rate of return.

Note that an equivalent change of measure of the form $dz_{k,t}^* = dz_{k,t} - g_k \sqrt{v_{k,t}} dt$ lead to a discounted martingale for stocks if and only if $A^{-1} (AB + \frac{1}{2} A \circ A (D - \mathbf{1})) = G$, with G a diagonal matrix with components g_k . Hence the SDEs for the processes under this risk

neutral pricing measure \mathbb{Q} are given by

$$d \log S_{i,t} = \left(r - \frac{1}{2} \sum_{k=1}^n a_{ik}^2 v_{k,t} \right) dt + \sum_{k=1}^n a_{ik} \sqrt{v_{k,t}} dz_{k,t}^*,$$

$$dv_{j,t} = \kappa_j \left(\theta_j - \frac{(\kappa_j - \sigma_j g_j)}{\kappa_j} v_{j,t} \right) dt + \sigma_j \sqrt{v_{j,t}} dz_{j,t}^*, \quad j = 1, \dots, n.$$

From this it is seen that the flexibility of b as controller of equity premium and d as the driver of variance premium is lost to the parameter g in the continuous time limit. This is compatible with footnote 8 in Christoffersen, Heston, and Jacobs (2013) as any equivalent Girsanov-based change of measure in continuous time does not change variances.

3.2. Example: Cholesky affine MGARCH models

Similar to the full factor model of Vrontos, Dellaportas, and Politis (2003), the conditional covariance matrix of returns can be represented by its Cholesky decomposition. In this setting, A is a lower triangular matrix with 1's on the diagonal. Note that the order of assets in this model matters because of the triangular shape of A .

3.2.1. A bivariate index model

For simplicity and due to its usage in the empirical section, we next show the explicit form of equations for the HN-MGARCH with $n = 2$ under a Cholesky decomposition. In this case we have that

$$A = \begin{bmatrix} 1 & 0 \\ a & 1 \end{bmatrix}, \quad B = \begin{bmatrix} b_1 & 0 \\ 0 & b_2 \end{bmatrix}, \quad D = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}$$

$$\Lambda = \begin{bmatrix} -b_1 - \frac{1}{2}d_1 & 0 \\ -ab_1 - \frac{1}{2}a^2d_1 & -b_1 - \frac{1}{2}d_2 \end{bmatrix}, \quad H_t = \begin{bmatrix} h_{1,t} & ah_{1,t} \\ ah_{1,t} & a^2h_{1,t} + h_{2,t} \end{bmatrix}.$$

The physical process can be written as

$$R_{1,t} = r + \lambda_{11}h_{1,t} + \sqrt{h_{1,t}}z_{1,t} = r + \left(-b_1 - \frac{1}{2}d_1\right)h_{1,t} + \sqrt{h_{1,t}}z_{1,t} \quad (41)$$

$$\begin{aligned} R_{2,t} &= r + \lambda_{21}h_{1,t} + \lambda_{22}h_{2,t} + a\sqrt{h_{1,t}}z_{1,t} + \sqrt{h_{2,t}}z_{2,t} \\ &= r + \left(-ab_1 - \frac{1}{2}a^2d_1\right)h_{1,t} + \left(-b_2 - \frac{1}{2}d_2\right)h_{2,t} + a\sqrt{h_{1,t}}z_{1,t} + \sqrt{h_{2,t}}z_{2,t} \end{aligned} \quad (42)$$

$$h_{i,t} = \omega_i + \beta_i h_{i,t-1} + \alpha_i \left(z_{i,t-1} - \gamma_i \sqrt{h_{i,t-1}}\right)^2, \quad i = 1, 2,$$

with risk-neutral dynamics given by

$$R_{1,t} = r - \frac{1}{2}h_{1,t}^* + \sqrt{h_{1,t}^*}z_{1,t}^* \quad (43)$$

$$R_{2,t} = r - \frac{1}{2}a^2h_{1,t}^* - \frac{1}{2}h_{2,t}^* + a\sqrt{h_{1,t}^*}z_{1,t}^* + \sqrt{h_{2,t}^*}z_{2,t}^* \quad (44)$$

$$h_{i,t}^* = \omega_i^* + \beta_i h_{i,t-1}^* + \alpha_i^* \left(z_{i,t-1}^* - \gamma_i^* \sqrt{h_{i,t-1}^*}\right)^2.$$

Equation (44) demonstrates clearly the flexibility of our MGARCH framework in generating realistic correlations spills over to the risk neutral world.

Recall that the marginal variances and absolute correlation can be represented as follows under \mathbb{P} , shown in the first row, and under \mathbb{Q} , shown in the second row

$$\begin{array}{lll} H_{1,t} = h_{1,t} & H_{2,t} = a^2h_{1,t} + h_{2,t} & |\rho_t| = \sqrt{\frac{a^2h_{1,t}}{a^2h_{1,t} + h_{2,t}}} \\ H_{1,t}^* = h_{1,t}^* = d_1h_{1,t} & H_{2,t}^* = d_1 \left(a^2h_{1,t} + \frac{d_2}{d_1}h_{2,t}\right) & |\rho_t^*| = \sqrt{\frac{a^2h_{1,t}}{a^2h_{1,t} + \frac{d_2}{d_1}h_{2,t}}} \end{array}$$

It is easy to see that if $1 < d_2 < d_1$ then variances and correlations in the risk neutral world are greater than historical values. If $1 < d_2 = d_1$ then variances are larger under \mathbb{Q} but the correlation is the same under both measures. On the other hand, if $d_2 < 1 < d_1$ then the variance of the first asset and the correlation are clearly larger under \mathbb{Q} . The variance of the second asset would be larger under \mathbb{Q} at any time t where $\frac{h_{1,t}}{h_{2,t}} > \frac{1-d_2}{a^2(d_1-1)}$.

The case $d_1 = d_2 = 1$ corresponds to a linear affine pricing kernel and leads to no change in covariance. Note that when $(-ab_1 - 0.5a^2(1 - d_1)) \frac{h_{1,t}}{h_{2,t}} < b_2 + 0.5(1 - d_2)$, the risk premium for the second asset can be negative.

3.2.2. Implications for volatility indices

In a GARCH setting, the volatility index (VIX) of the j -th asset can be defined as

$$VIX_{j,t} = \frac{1}{100} \sqrt{252 \overline{H^*}_{j,t}} \quad (45)$$

where $\overline{H^*}_{j,t}$ is the average of expected conditional variance over τ future trading days (where $\tau = 21$ for the standard VIX) under the risk-neutral measure. This is given by

$$\overline{H^*}_{j,t} = \frac{1}{\tau} \sum_{k=1}^{\tau} E_t^{\mathbb{Q}}[H^*_{j,t+k}] \quad (46)$$

where $H^*_{j,t}$ is the conditional variance of daily returns of the j -th asset. In the bivariate Cholesky HN-GARCH model we have $H^*_{1,t} = h_{1,t}$ and $H^*_{2,t} = a^2 h^*_{1,t} + h^*_{2,t}$ which gives

$$\overline{H^*}_{1,t} = \frac{1}{\tau} \sum_{k=1}^{\tau} E_t^{\mathbb{Q}}[h^*_{1,t+k}] \quad (47)$$

$$\overline{H^*}_{2,t} = \frac{1}{\tau} \sum_{k=1}^{\tau} E_t^{\mathbb{Q}}[a^2 h^*_{1,t+k} + h^*_{2,t+k}]. \quad (48)$$

Let p_j and $\mu_{j,h}$ be the persistence and long-run value of each factor variance $h_{j,t}$ for $j = 1, 2$, given by

$$p_j = \beta_j^* + \alpha_j^* \gamma_j^{*2}, \quad \mu_{j,h} = \frac{\omega_j + \alpha_j}{1 - p_j}, \quad (49)$$

and observe that $E_t^{\mathbb{Q}}[h_{j,t+2}^*] = (\omega_j + \alpha_i) + p_j h_{j,t+1}^*$. Using iterated expectations and averaging over τ days, we obtain

$$\overline{H^*}_{1,t} = \mu_{1,h} + \frac{1 - p_1^\tau}{\tau(1 - p_1)} (h_{1,t+1}^* - \mu_{1,h}) \quad (50)$$

$$\overline{H^*}_{2,t} = a^2 \mu_{h,1} + a^2 \frac{1 - p_1^\tau}{\tau(1 - p_1)} (h_{1,t+1}^* - \mu_{h,1}) + \mu_{h,2} + \frac{1 - p_2^\tau}{\tau(1 - p_2)} (h_{2,t+1}^* - \mu_{h,2}). \quad (51)$$

which together with (45) gives the VIX formula implied by our model.

Note that the relation between a bivariate VIX and the vector h_t^* is captured by the following equations

$$\begin{aligned} \frac{VIX_{1,t}^2}{c} &= \mu_{h,1} + \frac{1 - p_1^\tau}{\tau(1 - p_1)} \left(\omega_1^* + \beta_1 h_{1,t}^* + \alpha_1^* \left(z_{1,t}^* - \gamma_1^* \sqrt{h_{1,t}^*} \right)^2 - \mu_{h,1} \right) \\ \frac{VIX_{2,t}^2}{c} &= a^2 \frac{VIX_{1,t}^2}{c} + \mu_{h,2} + \frac{1 - p_2^\tau}{\tau(1 - p_2)} \left(\omega_2^* + \beta_2 h_{2,t}^* + \alpha_2^* \left(z_{2,t}^* - \gamma_2^* \sqrt{h_{2,t}^*} \right)^2 - \mu_{h,2} \right), \end{aligned}$$

where $c = 252(100)^2$. This represents the “model-implied” multivariate VIX dynamics. Our proposed model permits a degree of dependence among VIXs influenced by the correlation between the underlying assets. This is compatible with the stylized fact of co-movements in variance.

4. Empirical analysis

We estimate/calibrate the parameters of our proposed bivariate Cholesky HN-MGARCH model using daily returns and VIX values. Data under the risk neutral measure, like, e.g., option data or implied volatilities, are required to identify the variance premium parameters, d_i . The first asset in all the bivariate models is the S&P 500 Index. As for the second asset we choose the 5 stocks for which implied volatility index values are provided by the Chicago Board Options Exchange, CBOE, namely Apple, Amazon, Google, Goldman Sachs and IBM. The historical daily values of the VIX equivalent for these stocks are available on the CBOE website and our data set consists of the full sample of these with daily returns and volatility

index values for 2243 trading days from June 2010 to April 2019.

We first describe the estimation methodology used in our multivariate setting and report the estimates for the five bivariate cases considered. Besides the bivariate model, we also present the univariate HN-GARCH estimation for each of the 5 selected stocks. We then study the impact of some key parameters in our model on the implied volatility surfaces on single-asset European call options. To demonstrate the full impact of our modelling approach we also consider the case of a popular two dimensional product; the correlation option. Specifically, we study the efficiency of our closed form formulas for pricing that are available with the affine MGARCH model as well as the impact of the covariance-dependent kernel on the price of these bivariate options.

4.1. Estimation methodology and results

The most obvious method for estimating the parameters in our model is to use Maximum Likelihood based estimation using a sample of historical returns on the relevant assets. To illustrate this approach, let $\Theta = (\omega, \alpha, \beta, \gamma, A, B, D)$ denote the parameter set in the multivariate Cholesky HN-GARCH, and assume $\{R_t : t = 1, \dots, T\}$ is the observed set of daily return vectors for n assets. By assumption, the innovations are Gaussian and we have that $R_t | \mathcal{F}_{t-1} \sim N(r\mathbf{1} + \Lambda\Sigma_t\mathbf{1}, A\Sigma_tA')$. Thus, the vector of residuals is given by

$$z_t = (A\sqrt{\Sigma_t})^{-1}(R_t - r\mathbf{1} - \Lambda\Sigma_t\mathbf{1}) = \sqrt{\Sigma_t^{-1}A^{-1}}(R_t - r\mathbf{1} - \Lambda\Sigma_t\mathbf{1}),$$

and entries of Σ_t are updated as

$$h_{j,t} = \omega_j + \beta_j h_{j,t} + \alpha_j \left(z_{j,t-1} - \gamma_j \sqrt{h_{j,t-1}} \right)^2.$$

Moreover, as $\det A = 1$ the likelihood function is given by

$$\begin{aligned} L(\Theta) &= -\frac{1}{2} \sum_{t=1}^T (n \log(2\pi) + \log(\det(\Sigma_t)) + z_t' z_t) \\ &= -\frac{1}{2} \sum_{t=1}^T \left(n \log(2\pi) + \sum_{i=1}^n \log(h_{i,t}) + \sum_{i=1}^n z_{i,t}^2 \right). \end{aligned} \quad (52)$$

The likelihood function of the univariate HN-GARCH is obtained by setting $n = 1$.

In the bivariate model ($n = 2$) there are 13 parameters to estimate which we denote explicitly with $\Theta = (\omega_1, \omega_2, \alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2, b_1, b_2, d_1, d_2, a)$. Note that when estimating the parameters using only returns data, we set $d_1 = d_2 = 1$. To estimate d_1 and d_2 we incorporate VIX data in the estimation. In particular, we extend the VIX likelihood in Hao and Zhang (2013), Kannianen, Lin, and Yang (2014) and Badescu, Chen, Couch, and Cui (2019) to multivariate models. To illustrate this method, let VIX_t^{Mkt} denote the vector of VIX values observed in the market at time t , and VIX_t^{Mod} denote the vector of VIX values implied by our model in Section 3.2.2. We then define the vector of VIX errors as

$$\varepsilon_t^V = \frac{VIX_t^{Mkt} - VIX_t^{Mod}}{100\sqrt{252}},$$

and we assume the individual error components are independent and normally distributed with zero mean such that $\varepsilon_t^V \sim N(0, \Sigma_V)$, where $\Sigma_V = \text{diag}(\sigma_1^2, \dots, \sigma_n^2)$ is a diagonal matrix of error variances. Then the VIX likelihood function is

$$\begin{aligned} L_V(\Theta^*) &= \sum_{t=1}^T \sum_{i=1}^n \log \left(\frac{1}{\sqrt{2\pi\sigma_{i,t}^2}} \exp \left(-\frac{(\varepsilon_{i,t}^V)^2}{2\sigma_{i,t}^2} \right) \right) \\ &= -\frac{1}{2} \sum_{t=1}^T \sum_{i=1}^n \left(\log(2\pi\sigma_i^2) + \frac{1}{\sigma_i^2} (\varepsilon_{i,t}^V)^2 \right). \end{aligned} \quad (53)$$

Here $\Theta^* = (\omega^*, \alpha^*, \beta^*, \gamma^*)$ denotes the risk-neutral parameters which are mapped from the physical parameters by Equation (40). For convenience, and since we do not need these

Table 1: Estimation results for then HN-GARCH model fitted to S&P500 data

Data	$\alpha_1(\times 10^{-6})$	β_1	$\gamma_1(\times 100)$	b_1	d_1	Lik
Returns	4.733 (0.4)	0.762 (0.012)	1.932 (0.125)	-5.976 (2.343)	1	7727
Returns and VIX	1.442 (0.063)	0.773 (0.007)	3.798 (0.107)	-6.163 -	1.374 (0.034)	18749

This table shows the estimation results for various univariate models fitted on S&P500 returns and VIX data from 2010 through 2019.

parameters for option pricing, the variance of the VIX errors is replaced by the sample variance $\sigma_i^2 \approx s_i^2 = Var[\varepsilon_{i,t}^V]$. Maximizing the total likelihood

$$L_T = L_R(\Theta) + L_V(\Theta^*),$$

allow us to estimate all the model parameters with the available return and VIX data.

Our model is formulated in such a way that the first asset impacts the dynamics of the second asset but not the other way around. This makes economically sense if we consider the first asset being, e.g., the market index which we take to be represented by the S&P 500 Index. To ensure consistency across the stocks we estimate the parameters for the index once only in a first step and keep them constant when estimating the individual stock parameters in a second step. The results from this estimation are shown in Table 1.⁴ The table presents parameter estimates based on historical returns alone, in which case we only estimate the equity risk parameter b_1 , and based on historical returns and VIX together.⁵ The table shows that irrespective of which data is used, the parameters are estimated significantly different from zero but that including the VIX in estimation increases the precision of the estimated dynamic parameters. The estimated parameter values are consistent with the

⁴In this and all other estimations we use a annualized risk free interest rate of 0.49%, which is the average of 3-month T-Bill rate over this period, and to initiate the conditional variance we use a one year warm-up period with returns from June 2009 to May 2010.

⁵We follow Christoffersen, Heston, and Jacobs (2013) by setting ω equal to zero and by keeping the equity risk parameter, b_1 , fixed at the value estimated from returns alone when estimating parameters based on option data.

existing empirical literature demonstrating high persistency in the conditional volatility and important asymmetries. The estimated value of b_1 indicates a meaningful size of the equity risk premium and the value of d_1 , which is estimated to be significantly larger than one when using VIX data, implies a large wedge between the level of the risk neutral and physical volatility, something which has also been documented empirically.

Given the estimated parameter values for the index in Table 1 it is straightforward to use a similar approach to estimate the corresponding parameters for the individual stocks. The results are shown in Table 2, where each panel reports the results for AAPL, AMZN, GOOGL, GS, and IBM, respectively. For completeness, Table 3 reports the results for the corresponding univariate models, i.e. when we fix $a = 0$ and there is no correlation allowed in the model.⁶ The first thing to note from the table is that not only is a always statistically significant but the loss in likelihood value from neglecting correlation, shown in brackets in right most column, is very large. Interestingly, both the estimated value of a , and hence the implied correlation, and the loss in likelihood value from assuming independence is larger when including both return and VIX data in the estimation. Across the 5 stock the results show that most of the parameters are estimated significantly different from zero, or in the case of d_2 from one. The exception to this is the estimated equity risk parameter, b_2 , which is significantly different from zero for only AAPL and AMZN and the estimated variance risk premium, d_2 , which is significantly different from one for only AMZN, GOOGL, and GS. Note that, as it was the case with the index, including VIX in the estimation increases the precision in the estimated dynamic parameters.

4.2. Option pricing

One of the main arguments for working with models for which the mgf of the returns is known, is that European options can then be priced efficiently. We now consider single-asset

⁶It is not possible to directly compare the estimated parameters, in particular the ones related to the dynamics of the conditional variance, between the two panels, since in the independent specification these are related to total variance and not just the asset specific part of this as is the case in the bivariate model.

Table 2: Estimation results for bivariate HN-GARCH model fitted to individual stock data

Data	$\alpha_2(\times 10^{-6})$	β_2	$\gamma_2(\times 100)$	b_2	d_2	a	Lik
Panel A: AAPL							
Returns	2.055 (0.197)	0.961 (0.004)	1.15 (0.121)	-3.234 (1.592)	1	0.981 (0.027)	6567 [391]
Returns and VIX	1.178 (0.131)	0.993 (0.001)	0.148 (0.095)	-3.213 -	0.958 (0.028)	1.143 (0.015)	16181 [601]
Panel B: AMSN							
Returns	1.914 (0.167)	0.981 (0.002)	0.778 (0.077)	-3.024 (1.384)	1	1.218 (0.029)	6064 [377]
Returns and VIX	0.652 (0.065)	0.985 (0.002)	1.398 (0.127)	-2.879 -	0.709 (0.023)	1.453 (0.017)	15299 [807]
Panel C: GOOGL							
Returns	0.016 (0.01)	0.973 (0.001)	13.061 (4.199)	1.259 (0.993)	1	1.075 (0.026)	6815 [582]
Returns and VIX	0.04 (0.003)	0.93 (0.001)	13.195 (0.576)	1.362 -	0.795 (0.025)	1.128 (0.012)	16922 [1009]
Panel D: GS							
Returns	1.755 (0.28)	0.984 (0.002)	0.204 (0.128)	2.297 (1.905)	1	1.262 (0.02)	7006 [782]
Returns and VIX	0.267 (0.026)	0.923 (0.006)	5.283 (0.404)	2.215 -	1.164 (0.047)	1.25 (0.013)	16508 [814]
Panel E: IBM							
Returns	0.012 (0.005)	0.976 (0.001)	13.917 (2.721)	-1.636 (1.252)	1	0.84 (0.021)	7255 [561]
Returns and VIX	0.01 (0.002)	0.993 (0.001)	8.497 (0.901)	-1.636 -	1.001 (0.025)	0.926 (0.008)	17936 [1163]

This table shows the parameter estimates (and standard errors in parenthesis) for the bivariate models fitted on the time series of stock returns and implied volatility data from 2010 through 2019 keeping the parameters for the index fixed at their respective values from Table 1. In brackets in the right most column we show differences in likelihood values between this table and Table 3.

as well two-asset correlation option pricing in the bivariate affine HN-MGARCH model.

Table 3: Estimation results for univariate HN-GARCH model fitted to individual stock data

Data	$\alpha_2(\times 10^{-6})$	β_2	$\gamma_2(\times 100)$	b_2	d_2	a	Lik
Panel A: AAPL							
Returns	22.717 (1.664)	0.859 (0.008)	0.508 (0.038)	-3.79 (1.33)	1	0	6176
Returns and VIX	5.815 (0.331)	0.958 (0.002)	0.57 (0.036)	-3.88 -	1.191 (0.024)	0	15580
Panel B: AMZN							
Returns	6.13 (0.464)	0.924 (0.006)	0.991 (0.058)	-3.499 (1.101)	1	0	5687
Returns and VIX	3.184 (0.196)	0.979 (0.002)	0.632 (0.055)	-3.559 -	1.121 (0.019)	0	14492
Panel C: GOOGL							
Returns	1.109 (0.118)	0.956 (0.007)	1.891 (0.227)	-3.401 (1.435)	1	0	6233
Returns and VIX	1.387 (0.075)	0.954 (0.003)	1.703 (0.095)	-3.443 -	1.083 (0.016)	0	15913
Panel D: GS							
Returns	9.66 (1.022)	0.934 (0.008)	0.521 (0.077)	-1.301 (1.336)	1	0	6224
Returns and VIX	4.806 (0.278)	0.91 (0.004)	1.208 (0.051)	-1.42 -	1.239 (0.031)	0	15694
Panel E: IBM							
Returns	1.044 (0.119)	0.961 (0.006)	1.765 (0.184)	-1.495 (1.761)	1	0	6694
Returns and VIX	1.105 (0.073)	0.99 (0.001)	0.5 (0.086)	-1.579 -	1.203 (0.021)	0	16773

This table shows the parameter estimates (and standard errors in parenthesis) for the univariate models fitted on the time series of stock returns and implied volatility data from 2010 through 2019.

4.2.1. Single-asset options

The value of a European call option on the j 'th asset with strike price K and time to maturity T is given by

$$C = e^{-r(T-t)} E_t^{\mathbb{Q}}[\max(e^{x_{j,T}} - K, 0)] \quad (54)$$

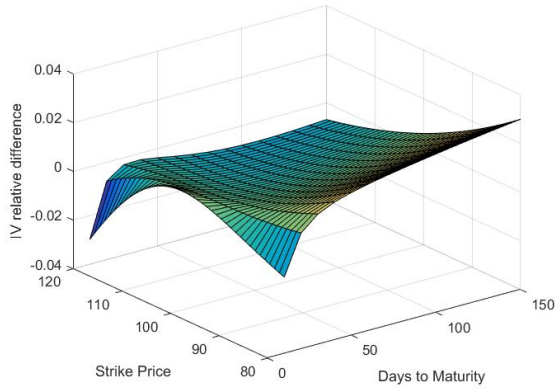
$$= e^{-r(T-t)} \left(\int_{\log K}^{+\infty} e^{x_{j,T}} f_j^{\mathbb{Q}}(x_{j,T}) dx_{j,T} - K \int_{\log K}^{+\infty} f_j^{\mathbb{Q}}(x_{j,T}) dx_{j,T} \right), \quad (55)$$

where $f_j^{\mathbb{Q}}$ is the marginal density of the logarithm of j -th asset price under the risk-neutral measure. Using well-known Fourier inversion methods, the above integrals can be written in terms of the marginal mgf of $x_{j,T}$, $m_j^{\mathbb{Q}}(u_j)$, and we can thus price a European call option with strike price K and time to maturity T as

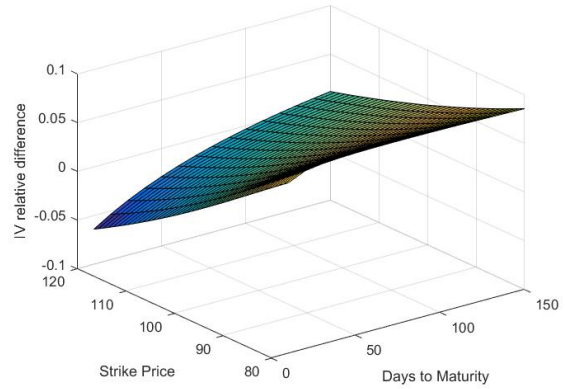
$$C = \frac{S_{j,t} - Ke^{-r(T-t)}}{2} + \frac{e^{-r(T-t)}}{\pi} \int_0^{\infty} \operatorname{Re} \left[\frac{K^{-iu_j} (m_j^{\mathbb{Q}}(iu_j + 1) - Km_j^{\mathbb{Q}}(iu_j))}{iu_j} \right] du_j,$$

where $\operatorname{Re}[\cdot]$ denotes the real part of a complex number. An alternative method based on the Fast Fourier Transform is provided in Carr and Madan (1999). The price of a European put option can be obtained from the put-call parity relation.

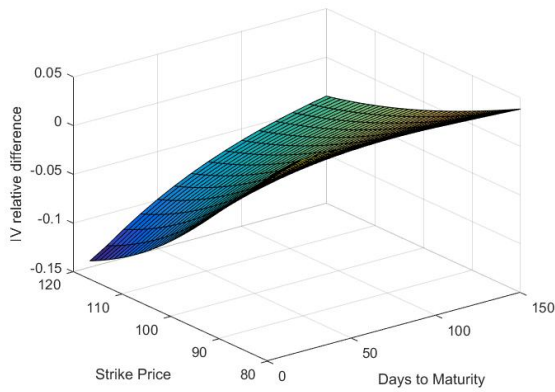
Given the parameter estimates we can now use these formulas to price a sample of single-asset options written on the individual stocks with varying moneyness and maturity. Figure 1 displays the relative differences in the implied volatility surface between a (univariate) HN-GARCH model with variance-dependent kernel and our full model as described by Equation (44) for single asset options written on each of the 5 stocks considered. The figure shows that the difference can be as large as 20% in some cases and that these differences varies importantly across moneynesses and maturity as well as between assets. For example, the relative differences are much smaller for AAPL and GS than for the other assets. For AMZN, GOOGL and IBM the plots show a similar pattern documenting that neglecting correlation when pricing options, even single asset options, leads to significant relative mispricing. For example, the plots show that out of the money call options (and hence in the money put options) are priced too high when using the restricted univariate model whereas the in the money call options (and hence out of the money put options) are priced too low with this model compared to when using the appropriate bivariate model. In other words, a potential explanation for the underpricing of out of the money put options, which has been empirically documented in several papers, could be explained by neglected correlation between the underlying asset and, e.g., the market index. Note that the differences are most important



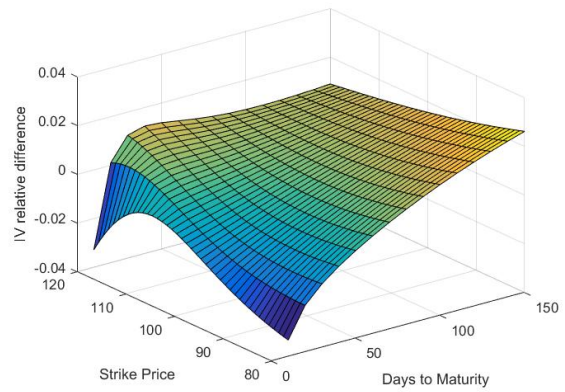
(a) AAPL



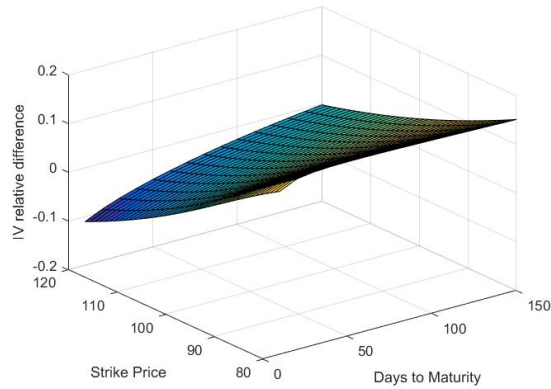
(b) AMZN



(c) GOOGL



(d) GS



(e) IBM

Figure 1: Relative differences in implied volatilities for single asset options

This figure plots the differences in implied volatilities for options priced with a one dimensional model and with a two dimensional model relative to the latter, i.e. divided by the implied volatility from the one dimensional option price.

for short term options and as the maturity increases these relative differences tend to vanish resulting in a clear term structure pattern in the pricing errors.

4.2.2. Two-asset options

The value of a Correlation call option on two assets with strike prices K_1 and K_2 , and time to maturity T as defined in Bakshi and Madan (2000) is given by

$$C = e^{-r(T-t)} E_t^{\mathbb{Q}} [\max(e^{x_{1,T}} - K_1, 0) \max(e^{x_{2,T}} - K_2, 0)] \quad (56)$$

$$= e^{-r(T-t)} \left(\int_{\log K_1}^{+\infty} \int_{\log K_2}^{+\infty} \max((e^{x_{1,T}} - K_1)(e^{x_{2,T}} - K_2), 0) f^{\mathbb{Q}}(x_{1,T}, x_{2,T}) dx_{1,T} dx_{2,T} \right), \quad (57)$$

where $f^{\mathbb{Q}}$ is the joint density of logarithm of the asset prices under the risk-neutral measure. Using Fourier inversion methods, the above integrals can be similarly simplified and written in terms of the joint mgf $m^{\mathbb{Q}}(u_1, u_2)$, hence the price can be expressed as follows (see Dempster and Hong (2002)):

$$C = \frac{e^{-\alpha_1 \log K_1 - \alpha_2 \log K_2}}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-i(u_1 \log K_1 + u_2 \log K_2)} \Psi(u_1, u_2) du_1 du_2,$$

$$\Psi(u_1, u_2) = \frac{e^{-r(T-t)} m^{\mathbb{Q}}(u_1 - (\alpha_1 + 1)i, u_2 - (\alpha_2 + 1)i)}{(\alpha_1 + iu_1)(\alpha_1 + 1 + iu_1)(\alpha_2 + iu_2)(\alpha_2 + 1 + iu_2)}.$$

Since the joint mgf is available in our HN-MGARCH model closed form option pricing is feasible.

In the univariate case it is well known that the closed form solution available in the Affine model is orders of magnitude faster than what would be obtained with, e.g., Monte Carlo simulation methods. Before we consider the actual price estimates for multivariate options we consider the relative efficiency of our mgf based method compared to crude Monte Carlo simulation. Figure 2 plots the root mean squared error, RMSE, against computational time for the recursive method in two dimensions and compares this to what is obtained with Monte Carlo simulation. The top plots show the results for at the money options with

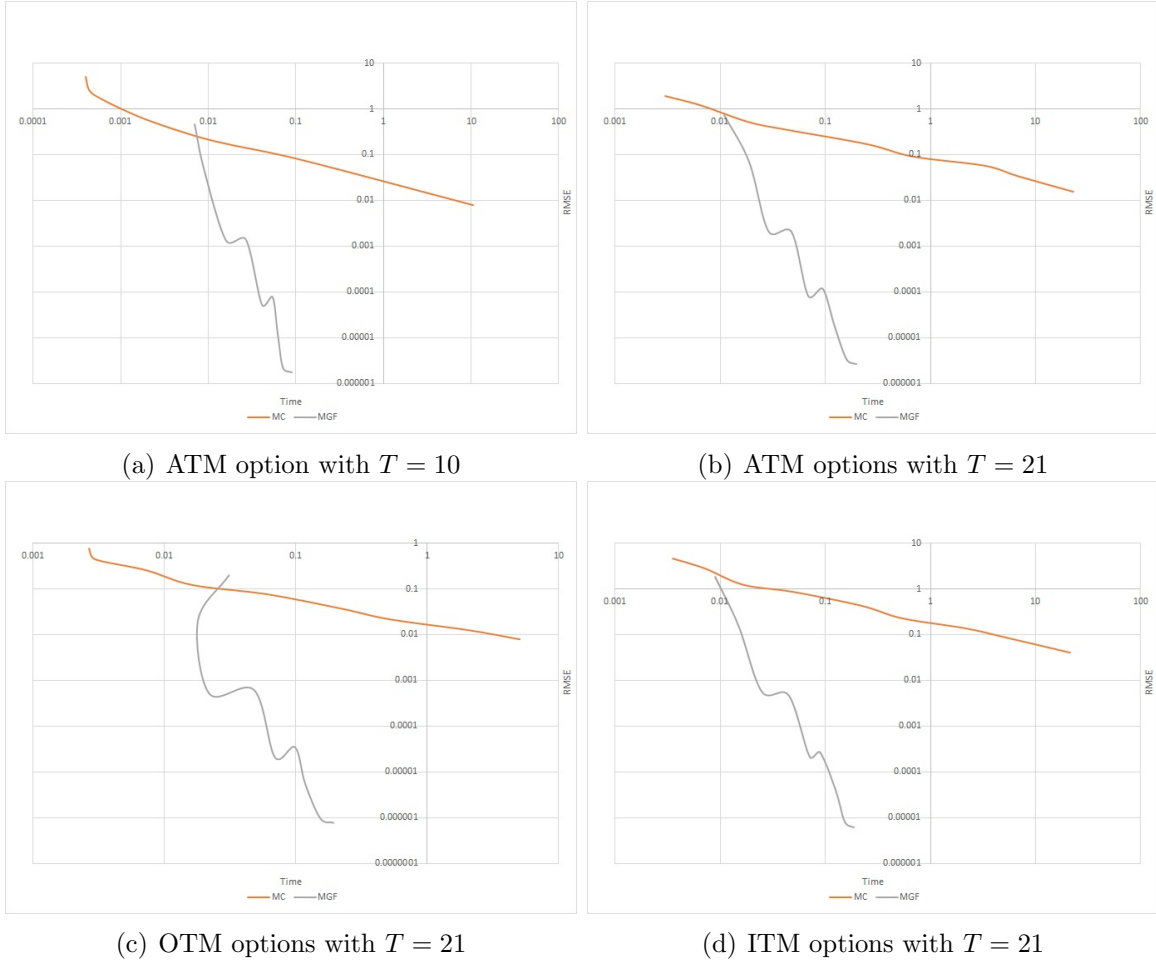


Figure 2: RMSE against computational time for recursive mgf and Monte Carlo methods

This figure plots the RMSE of our recursive mgf method compared to that of Monte Carlo simulation methods. For the recursive mgf method the results are based on increasing the number of points over which the integral is approximated. For the Monte Carlo simulation the results are based on 100 independent simulations with increasing number of paths.

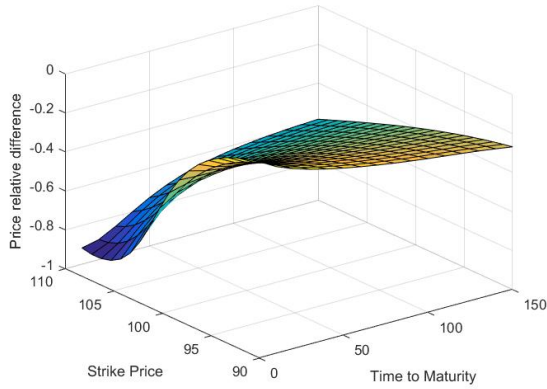
two different maturities demonstrating that the mgf based recursive method is indeed much more efficient than simulation in two dimensions. In particular, the efficiency increases with option maturity and for one month options, i.e. with $T = 21$ trading days, to obtain a one cent error would roughly require 1,000 times more computational time with a simulation method than with the recursive mgf method we develop in Section 3. The bottom plots show that the relative efficiency increases with moneyness and it may be several thousand times faster to obtain a certain precision with the mgf based method than with a simulation method. Note also that for any reasonable level of precision the mgf method performs the

best irrespective of the moneyness and maturity of the option.

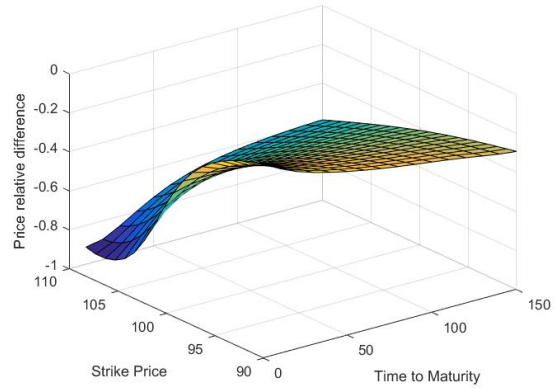
We now consider the pricing of multivariate options, which our model allows for in closed form. For convenience, we consider the correlation call option only, although pricing options with other payoffs would be straightforward. We are particularly interested in the effect of allowing for a variance dependent pricing kernel, something that has not been considered in the previous literature. In Figure 3 we report the relative differences in the price of correlation call options when considering a linear pricing kernel and a variance dependent pricing kernel. To be specific, we plot the difference between the price obtained with a linear kernel and a variance dependent kernel relative to the latter. The plot first of all shows that all the prices obtained with a linear kernel are lower than those obtained with a variance dependent kernel and neglecting the price of variance risk can lead to significantly underpricing of options. This is expected since the variance risk parameters d_1 and d_2 are what drives the wedge between risk neutral and physical volatilities and, as Section 3.2.1 shows, risk neutral and physical correlations. The underpricing is particularly severe for out of the money options which only have meaningful prices when a variance dependent pricing kernel is used but have prices very close to zero when neglecting this premium. However, underpricing is seen to occur across almost all values of moneyness and maturity and it is only for the deep in the money options with very short maturity where price differences are small and potentially negligible.

5. Conclusion

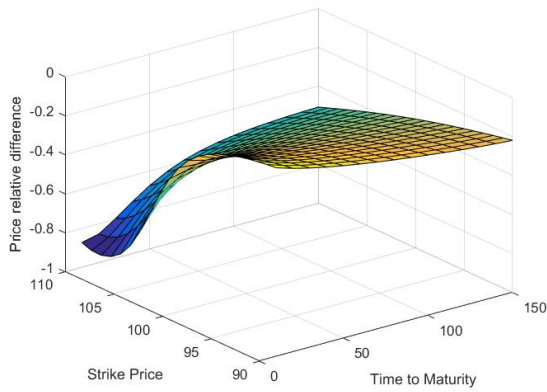
This paper introduces a class of flexible multivariate GARCH models motivated via factor analysis decomposition, which allow a recursive representation for the moment generating function (mgf) under both historical and risk-neutral pricing measures. We describe the statistical properties of this model, in particular the recursive mgf, marginal and joint moments and covariance dynamics. The continuous time version of the model is reported, together with the implications for the dynamics of a vector of volatility indexes. An empirical anal-



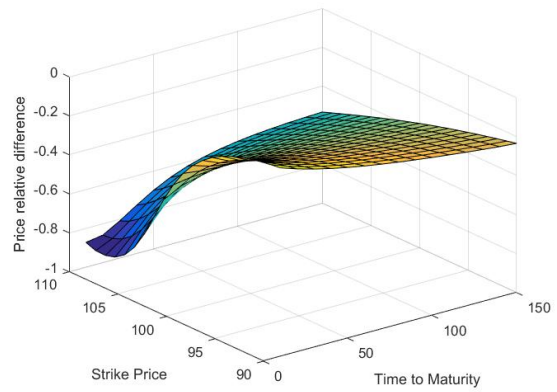
(a) AAPL



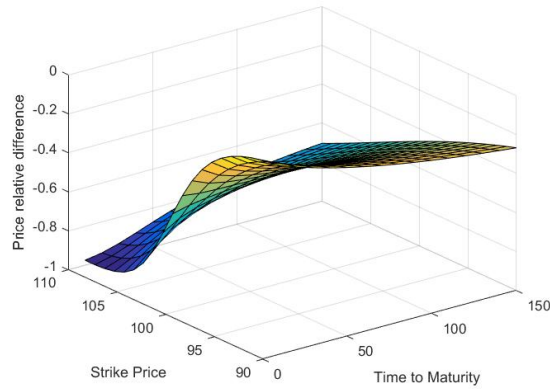
(b) AMZN



(c) GOOGL



(d) GS



(e) IBM

Figure 3: Relative differences in prices for multi asset correlation options

This figure plots the differences in prices for multiple asset correlation options priced with a linear kernel and with a variance dependent kernel relative to the latter, i.e. divided by the price from the option price with a variance dependent kernel.

ysis is conducted on the five pairs of asset returns for which volatility indices are reported by Chicago Board of Options Exchange. The implications of the model’s newly captured features, namely the dynamic correlation structure and covariance-dependent kernel, on the pricing of one and two-assets derivatives is highlighted via changes of up to 15% in the implied volatility surface and large and economically significant changes in the price of correlation options, respectively. The methodology is proven to be highly efficient in the pricing of two-asset products which can be done up to 1,000 times faster than with Monte Carlo simulation.

References

- BABAOĞLU, K., P. CHRISTOFFERSEN, S. HESTON, AND K. JACOBS (2018): “Option Valuation with Volatility Components, Fat Tails, and Nonmonotonic Pricing Kernels,” *The Review of Asset Pricing Studies*, 8(2), 183–231.
- BADESCU, A., Y. CHEN, M. COUCH, AND Z. CUI (2019): “Variance Swaps Valuation under Non-affine GARCH models and Their Diffusion Limits,” *Quantitative Finance*, 19(2), 227–246.
- BADESCU, A., Z. CUI, AND J.-P. ORTEGA (2018): “Closed-form Variance Swap Prices under General Affine GARCH Models and Their Continuous-time Limits,” *forthcoming in Annals of Operations Research*.
- BAKSHI, G., AND D. MADAN (2000): “Spanning and derivative-security valuation,” *Journal of financial economics*, 55(2), 205–238.
- BAUWENS, L., S. LAURENT, AND J. V. ROMBOUTS (2006): “Multivariate GARCH models: A Survey,” *Journal of Applied Econometrics*, 21(1), 79–109.
- BÉGIN, J.-F., C. DORION, AND G. GAUTHIER (2019): “Idiosyncratic Jump Risk Matters: Evidence from Equity Returns and Options,” *forthcoming in The Review of Financial Studies*.

- BURTSCHHELL, X., J. GREGORY, AND J.-P. LAURENT (2009): “A Comparative Analysis of CDO Pricing Models under the Factor Copula Framework,” *The Journal of Derivatives*, 16(4), 9–37.
- BUSS, A., AND G. VILKOV (2012): “Measuring Equity Risk with Option-implied Correlations,” *The Review of Financial Studies*, 25(10), 3113–3140.
- CARR, P., AND D. MADAN (1999): “Option Valuation using the Fast Fourier Transform,” *Journal of Computational Finance*, 2(4), 61–73.
- CHRISTOFFERSEN, P., S. HESTON, AND K. JACOBS (2006): “Option Valuation with Conditional Skewness,” *Journal of Econometrics*, 131(1-2), 253–284.
- CHRISTOFFERSEN, P., S. HESTON, AND K. JACOBS (2009): “The Shape and Term Structure of the Index Option Smirk: Why Multifactor Stochastic Volatility Models work so well,” *Management Science*, 55(12), 1914–1932.
- CHRISTOFFERSEN, P., S. HESTON, AND K. JACOBS (2013): “Capturing Option Anomalies with a Variance-Dependent Pricing Kernel,” *The Review of Financial Studies*, 26(8), 1962–2006.
- DE COL, A., A. GNOATTO, AND M. GRASSELLI (2013): “Smiles all Around: FX Joint Calibration in a Multi-Heston Model,” *Journal of Banking & Finance*, 37(10), 3799–3818.
- DEMPSTER, M. A. H., AND S. G. HONG (2002): “Spread Option Valuation and the Fast Fourier Transform,” in *Mathematical Finance-Bachelier Congress 2000*, pp. 203–220. Springer.
- DIEBOLD, F. X., AND M. NERLOVE (1989): “The Dynamics of Exchange Rate Volatility: A Multivariate Latent Factor ARCH Model,” *Journal of Applied Econometrics*, 4(1), 1–21.
- DRIESSEN, J., P. J. MAENHOUT, AND G. VILKOV (2009): “The Price of Correlation Risk: Evidence from Equity Options,” *The Journal of Finance*, 64(3), 1377–1406.

- ESCOBAR, M., S. FERRANDO, AND A. RUBTSOV (2017): “Optimal Investment under Multi-factor Stochastic Volatility,” *Quantitative Finance*, 17(2), 241–260.
- FAMA, E. F., AND K. R. FRENCH (2015): “A Five-Factor Asset Pricing Model,” *Journal of Financial Economics*, 116(1), 1–22.
- HAO, J., AND J. E. ZHANG (2013): “GARCH Option Pricing Models, the CBOE VIX, and Variance Risk Premium,” *Journal of Financial Econometrics*, 11(3), 556–580.
- HESTON, S. L. (1993): “A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options,” *The Review of Financial Studies*, 6(2), 327–343.
- HESTON, S. L., AND S. NANDI (2000): “A Closed-form GARCH Option Valuation Model,” *The Review of Financial Studies*, 13(3), 585–625.
- KANNIAINEN, J., B. LIN, AND H. YANG (2014): “Estimating and Using GARCH Models with VIX data for Option Valuation,” *Journal of Banking & Finance*, 43, 200–211.
- ORNTHANALAI, C. (2014): “Levy Jump Risk: Evidence from Options and Returns,” *Journal of Financial Economics*, 112(1), 69–90.
- POON, S.-H., M. ROCKINGER, AND J. TAWN (2003): “Extreme Value Dependence in Financial Markets: Diagnostics, Models, and Financial Implications,” *The Review of Financial Studies*, 17(2), 581–610.
- SKLAR, M. (1959): “Fonctions de repartition an dimensions et leurs marges,” *Publ. inst. statist. univ. Paris*, 8, 229–231.
- VAN DER WEIDE, R. (2002): “GO-GARCH: A Multivariate Generalized Orthogonal GARCH Model,” *Journal of Applied Econometrics*, 17(5), 549–564.
- VRONTOS, I. D., P. DELLAPORTAS, AND D. N. POLITIS (2003): “A Full-Factor Multivariate GARCH model,” *The Econometrics Journal*, 6(2), 312–334.

Appendix A. Proofs of main propositions

Proof of Proposition 2.2.

The proof is based on the law of iterated expectations. We demonstrate the calculations for X_{t+1} and X_{t+2} . The proof then follows recursively. Starting with mgf of X_{t+1} we have

$$\begin{aligned}
 E_t[e^{\mathbf{u}'X_{t+1}}] &= E_t \left[\exp \left(\sum_{i=1}^n u_i x_{i,t+1} \right) \right] \\
 &= E_t \left[\exp \left(\sum_{i=1}^n u_i x_{i,t} + r \sum_{i=1}^n u_i + \sum_{i=1}^n \sum_{j=1}^n u_i \lambda_{ij} h_{j,t+1} + \sum_{i=1}^n \sum_{j=1}^n u_i a_{ij} \varepsilon_{j,t+1} \right) \right] \\
 &= \exp \left(\sum_{i=1}^n u_i x_{i,t} + r \sum_{i=1}^n u_i + \sum_{j=1}^n \sum_{i=1}^n u_i \lambda_{ij} h_{j,t+1} \right) \times \prod_{j=1}^n E_t \left[\exp \left(\left(\sum_{i=1}^n u_i a_{ij} \right) \varepsilon_{j,t+1} \right) \right].
 \end{aligned}$$

where we used the fact that ε_{t+1} is a vector of independent shocks. Using definition of joint cgf in Equations (3) and the assumption in (4) we get

$$\begin{aligned}
 E_t[e^{\mathbf{u}'X_{t+1}}] &= \exp \left(\sum_{i=1}^n u_i x_{i,t} + r \sum_{i=1}^n u_i + \sum_{j=1}^n \sum_{i=1}^n u_i \lambda_{ij} h_{j,t+1} \right) \times \prod_{j=1}^n \exp \left(\varphi_j \left(\sum_{i=1}^n u_i a_{ij}, 0 \right) + \rho_j \left(\sum_{i=1}^n u_i a_{ij}, 0 \right) h_{j,t+1} \right) \\
 &= \exp \left(\sum_{i=1}^n u_i x_{i,t} + J_1 + \sum_{j=1}^n K_{j,1} h_{j,t+1} \right),
 \end{aligned}$$

with J_1 and $K_{j,1}$ given by

$$\begin{aligned}
 J_1 &= r \sum_{i=1}^n u_i + \sum_{j=1}^n \varphi_j \left(\sum_{i=1}^n u_i a_{ij}, 0 \right) \\
 K_{j,1} &= \sum_{i=1}^n u_i \lambda_{ij} + \rho_j \left(\sum_{i=1}^n u_i a_{ij}, 0 \right), \quad j = 1, \dots, n.
 \end{aligned}$$

In the second step we show the mgf of X_{t+2} is of the form in Equation (5). We have

$$\begin{aligned}
E_t[e^{\mathbf{u}'X_{t+2}}] &= E_t[E_{t+1}[e^{\mathbf{u}'X_{t+2}}]] \\
&= E_t \left[\exp \left(\sum_{i=1}^n u_i x_{i,t} + r \sum_{i=1}^n u_i + \sum_{i=1}^n \sum_{j=1}^n u_i \lambda_{ij} h_{j,t+1} + \sum_{i=1}^n \sum_{j=1}^n u_i a_{ij} \varepsilon_{j,t+1} \right. \right. \\
&\quad \left. \left. + J_1 + \sum_{j=1}^n K_{j,1} h_{j,t+2} \right) \right] \\
&= \exp \left(\sum_{i=1}^n u_i x_{i,t} + J_1 + r \sum_{i=1}^n u_i + \sum_{j=1}^n \sum_{i=1}^n u_i \lambda_{ij} h_{j,t+1} \right) \\
&\quad \times \prod_{j=1}^n E_t \left[\exp \left(\left(\sum_{i=1}^n u_i a_{ij} \right) \varepsilon_{j,t+1} + K_{j,1} h_{j,t+2} \right) \right],
\end{aligned}$$

where we used the independence of shocks $\varepsilon_{j,t+1}$ for $j = 1, \dots, n$. Again, Equations (3) and (4) give

$$\begin{aligned}
E_t[e^{\mathbf{u}'X_{t+2}}] &= \exp \left(\sum_{i=1}^n u_i x_{i,t} + J_1 + r \sum_{i=1}^n u_i + \sum_{j=1}^n \sum_{i=1}^n u_i \lambda_{ij} h_{j,t+1} \right) \times \prod_{j=1}^n \exp \left(C_{(\varepsilon_{j,t+1}, h_{j,t+2}) | \mathcal{F}_t} \left(\sum_{i=1}^n u_i a_{ij}, K_{j,1} \right) \right) \\
&= \exp \left(\sum_{i=1}^n u_i x_{i,t} + J_1 + r \sum_{i=1}^n u_i + \sum_{j=1}^n \sum_{i=1}^n u_i \lambda_{ij} h_{j,t+1} \right) \\
&\quad \times \prod_{j=1}^n \exp \left(\varphi_j \left(\sum_{i=1}^n u_i a_{ij}, K_{j,1} \right) + \rho_j \left(\sum_{i=1}^n u_i a_{ij}, K_{j,1} \right) h_{j,t+1} \right) \\
&= \exp \left(\sum_{i=1}^n u_i x_{i,t} + J_2 + \sum_{j=1}^n K_{j,2} h_{j,t+1} \right),
\end{aligned}$$

with J_1 and $K_{j,1}$ given by

$$\begin{aligned}
J_2 &= J_1 + r \sum_{i=1}^n u_i + \sum_{j=1}^n \varphi_j \left(\sum_{i=1}^n u_i a_{ij}, K_{j,1} \right) \\
K_{j,2} &= \sum_{i=1}^n u_i \lambda_{ij} + \rho_j \left(\sum_{i=1}^n u_i a_{ij}, K_{j,1} \right), \quad j = 1, \dots, n.
\end{aligned}$$

Continuing recursively we obtain

$$J_m = J_{m-1} + r \sum_{i=1}^n u_i + \sum_{j=1}^n \varphi_j \left(\sum_{i=1}^n u_i a_{ij}, K_{j,m-1} \right)$$

$$K_{j,m} = \sum_{i=1}^n u_i \lambda_{ij} + \rho_j \left(\sum_{i=1}^n u_i a_{ij}, K_{j,m-1} \right), \quad j = 1, \dots, n,$$

for $m = 0, 1, \dots, T - t$ with initial values $J_0 = K_{1,0} = \dots = K_{n,0} = 0$. \square

Proof of Proposition 3.1.

The Radon-Nikodym derivative is a positive function. Using the definition of joint cgf $C_{(\varepsilon_{j,t}, h_{j,t+1})|\mathcal{F}_{t-1}}$ in Equation (3) we have

$$\begin{aligned} E_{t-1}^{\mathbb{P}} \left[\frac{dQ}{d\mathbb{P}} \Big|_{\mathcal{F}_t} \right] &= E_{t-1}^{\mathbb{P}} \left[\frac{dQ}{d\mathbb{P}} \Big|_{\mathcal{F}_{t-1}} \exp \left(\sum_{j=1}^n \left(\theta_{j,t} \varepsilon_{j,t} + \xi_{j,t} h_{j,t+1} - C_{(\varepsilon_{j,t}, h_{j,t+1})|\mathcal{F}_{t-1}}(\theta_{j,t}, \xi_{j,t}) \right) \right) \right] \\ &= \frac{dQ}{d\mathbb{P}} \Big|_{\mathcal{F}_{t-1}} \exp \left(- \sum_{j=1}^n C_{(\varepsilon_{j,t}, h_{j,t+1})|\mathcal{F}_{t-1}}(\theta_{j,t}, \xi_{j,t}) \right) \prod_{j=1}^n E_{t-1}^{\mathbb{P}} [\exp(\theta_{j,t} \varepsilon_{j,t} + \xi_{j,t} h_{j,t+1})] \\ &= \frac{dQ}{d\mathbb{P}} \Big|_{\mathcal{F}_{t-1}} \exp \left(- \sum_{j=1}^n C_{(\varepsilon_{j,t}, h_{j,t+1})|\mathcal{F}_{t-1}}(\theta_{j,t}, \xi_{j,t}) + \sum_{j=1}^n C_{(\varepsilon_{j,t}, h_{j,t+1})|\mathcal{F}_{t-1}}(\theta_{j,t}, \xi_{j,t}) \right) \\ &= \frac{dQ}{d\mathbb{P}} \Big|_{\mathcal{F}_{t-1}}. \end{aligned}$$

This shows the Radon-Nikodym derivative is a \mathbb{P} -martingale. Moreover, by continuing recursively we obtain $E_0^{\mathbb{P}} \left[\frac{dQ}{d\mathbb{P}} \Big|_{\mathcal{F}_t} \right] = 1$. Therefore, Equation (28) defines a valid Radon-Nikodym derivative.

Next, we show that each discounted asset price is a \mathbb{Q} -martingale. It is sufficient to show

that $E_{t-1}^{\mathbb{Q}} \left[\frac{S_{i,t}}{S_{i,t-1}} \right] = e^r$.

$$\begin{aligned}
E_{t-1}^{\mathbb{Q}} \left[\frac{S_{i,t}}{S_{i,t-1}} \right] &= E_{t-1}^{\mathbb{P}} \left[\frac{S_{i,t}}{S_{i,t-1}} \frac{d\mathbb{Q}/d\mathbb{P} | \mathcal{F}_t}{d\mathbb{Q}/d\mathbb{P} | \mathcal{F}_{t-1}} \right] \\
&= E_{t-1}^{\mathbb{P}} \left[\exp \left(R_{i,t} + \sum_{j=1}^n \left(\theta_{j,t} \varepsilon_{j,t} + \xi_{j,t} h_{j,t+1} - C_{(\varepsilon_{j,t}, h_{j,t+1}) | \mathcal{F}_{t-1}}(\theta_{j,t}, \xi_{j,t}) \right) \right) \right] \\
&= \exp \left(- \sum_{j=1}^n C_{(\varepsilon_{j,t}, h_{j,t+1}) | \mathcal{F}_{t-1}}(\theta_{j,t}, \xi_{j,t}) \right) \times \exp \left(r + \sum_{j=1}^n \lambda_{ij} h_{j,t} \right) \\
&\times E_{t-1}^{\mathbb{P}} \left[\exp \left(\sum_{j=1}^n \theta_{j,t} \varepsilon_{j,t} + \sum_{j=1}^n \xi_{j,t} g(h_{j,t}, \varepsilon_{j,t}) + \sum_{j=1}^n a_{ij} \varepsilon_{j,t} \right) \right].
\end{aligned}$$

Using independence of $\varepsilon_{j,t} | \mathcal{F}_{t-1}$ for $j = 1, \dots, n$, and definition of joint cgf $C_{(\varepsilon_{j,t}, h_{j,t+1}) | \mathcal{F}_{t-1}}$ in Equation (3) we get

$$\begin{aligned}
E_{t-1}^{\mathbb{Q}} \left[\frac{S_{i,t}}{S_{i,t-1}} \right] &= \exp \left(- \sum_{j=1}^n C_{(\varepsilon_{j,t}, h_{j,t+1}) | \mathcal{F}_{t-1}}(\theta_{j,t}, \xi_{j,t}) + r + \sum_{j=1}^n \lambda_{ij} h_{j,t} \right) \\
&\times \prod_{j=1}^n E_{t-1}^{\mathbb{P}} \left[\exp \left((\theta_{j,t} + a_{ij}) \varepsilon_{j,t} + \xi_{j,t} g(h_{j,t}, \varepsilon_{j,t}) \right) \right] \\
&= e^r \exp \left(\sum_{j=1}^n \left(-C_{(\varepsilon_{j,t}, h_{j,t+1}) | \mathcal{F}_{t-1}}(\theta_{j,t}, \xi_{j,t}) + \lambda_{ij} h_{j,t} + C_{(\varepsilon_{j,t}, h_{j,t+1}) | \mathcal{F}_{t-1}}(\theta_{j,t} + a_{ij}, \xi_{j,t}) \right) \right).
\end{aligned}$$

The above is equal to e^r if

$$\sum_{j=1}^n \left(-C_{(\varepsilon_{j,t}, h_{j,t+1}) | \mathcal{F}_{t-1}}(\theta_{j,t}, \xi_{j,t}) + \lambda_{ij} h_{j,t} + C_{(\varepsilon_{j,t}, h_{j,t+1}) | \mathcal{F}_{t-1}}(\theta_{j,t} + a_{ij}, \xi_{j,t}) \right) = 0.$$

Invoking Equation (4), the above can be written as

$$\sum_{j=1}^n \left(-\varphi_j(\theta_{j,t}, \xi_{j,t}) + \varphi_j(\theta_{j,t} + a_{ij}, \xi_{j,t}) \right) + \sum_{j=1}^n \left(-\rho_j(\theta_{j,t}, \xi_{j,t}) + \rho_j(\theta_{j,t} + a_{ij}, \xi_{j,t}) + \lambda_{ij} \right) h_{j,t} = 0.$$

Sufficient conditions for this equation to hold are

$$-\varphi(\theta_{j,t}, \xi_{j,t}) + \varphi(\theta_{j,t} + a_{ij}, \xi_{j,t}) = 0 \quad (\text{A.1})$$

$$-\rho(\theta_{j,t}, \xi_{j,t}) + \rho(\theta_{j,t} + a_{ij}, \xi_{j,t}) + \lambda_{ij} = 0, \quad j, i = 1, \dots, n, \quad (\text{A.2})$$

which gives Equations (26) and (27). \square

Proof of Corollary 3.3.

For each $j = 1, \dots, n$, the conditional joint cgf $C_{(\varepsilon_{j,t}, h_{j,t+1})|\mathcal{F}_{t-1}}$ under \mathbb{Q} is calculated as

$$\begin{aligned} \exp\left(C_{(\varepsilon_{j,t}, h_{j,t+1})|\mathcal{F}_{t-1}}^{\mathbb{Q}}(u, v)\right) &= E_{t-1}^{\mathbb{Q}}[\exp(u\varepsilon_{j,t} + vh_{j,t+1})] \\ &= E_{t-1}^{\mathbb{P}}\left[\exp(u\varepsilon_{j,t} + vh_{j,t+1}) \frac{d\mathbb{Q}/d\mathbb{P}|_{\mathcal{F}_t}}{d\mathbb{Q}/d\mathbb{P}|_{\mathcal{F}_{t-1}}}\right] \\ &= E_{t-1}^{\mathbb{P}}\left[\exp\left(u\varepsilon_{j,t} + vh_{j,t+1} + \sum_{k=1}^n (\theta_{k,t}\varepsilon_{k,t} + \xi_{k,t}h_{k,t+1} - C_{(\varepsilon_{k,t}, h_{k,t+1})|\mathcal{F}_{t-1}}(\theta_{k,t}, \xi_{k,t}))\right)\right] \\ &= \exp\left(-\sum_{k=1}^n C_{(\varepsilon_{k,t}, h_{k,t+1})|\mathcal{F}_{t-1}}(\theta_{k,t}, \xi_{k,t})\right) \\ &\times E_{t-1}^{\mathbb{P}}\left[\exp\left((u + \theta_{j,t})\varepsilon_{j,t} + (v + \xi_{j,t})h_{j,t+1} + \sum_{k \neq j} (\theta_{k,t}\varepsilon_{k,t} + \xi_{k,t}h_{k,t+1})\right)\right]. \end{aligned}$$

Since $\varepsilon_{j,t}$'s are independent for $j = 1, \dots, n$, we get

$$\begin{aligned} \exp\left(C_{(\varepsilon_{j,t}, h_{j,t+1})|\mathcal{F}_{t-1}}^{\mathbb{Q}}(u, v)\right) &= \exp\left(-\sum_{k=1}^n C_{(\varepsilon_{k,t}, h_{k,t+1})|\mathcal{F}_{t-1}}(\theta_{k,t}, \xi_{k,t})\right) \\ &\times \exp\left(C_{(\varepsilon_{j,t}, h_{j,t+1})|\mathcal{F}_{t-1}}(u + \theta_{j,t}, v + \xi_{j,t}) + \sum_{k \neq j} C_{(\varepsilon_{k,t}, h_{k,t+1})|\mathcal{F}_{t-1}}(\theta_{k,t}, \xi_{k,t})\right) \\ &= \exp\left(C_{(\varepsilon_{j,t}, h_{j,t+1})|\mathcal{F}_{t-1}}(u + \theta_{j,t}, v + \xi_{j,t}) - C_{(\varepsilon_{j,t}, h_{j,t+1})|\mathcal{F}_{t-1}}(\theta_{j,t}, \xi_{j,t})\right), \end{aligned}$$

which gives Equation (29). Equations (30) and (31) now follow from (4). \square