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 23-02

Covariance Dependent Kernels, a Q-Affine GARCH for Multi-Asset Option Pricing

Publié mai 2023

Ce document de recherche a été rédigée par :

Marcos Escobar-Anel, Western University
Javad Rastegari, Western University
Lars Stentoft, Western University

L'Institut canadien des dérivés n'assume aucune responsabilité liée aux propos tenus et aux opinions exprimées dans ses publications, qui n'engagent que leurs auteurs. De plus, l'Institut ne peut, en aucun cas être tenu responsable des conséquences dommageables ou financières de toute exploitation de l'information diffusée dans ses publications.

Covariance Dependent Kernels, a Q-Affine GARCH for multi-asset option pricing

Marcos Escobar-Anel
Department of Statistical and Actuarial Sciences, Western University, Canada

Javad Rastegari
Department of Statistical and Actuarial Sciences, Western University, Canada

Lars Stentoft
Department of Economics, Western University, Canada

Abstract

This paper introduces a class of multivariate GARCH models with sufficient flexibility to allow for pricing kernels dependent on variances and correlation. This extends the existing literature by explicitly modeling correlation dependent pricing kernels. A large subclass admits closed-form recursive solutions for the moment generating function under the risk-neutral measure, which permits efficient pricing of multi-asset options. We perform a full calibration to three bivariate series of index returns and their corresponding volatility indexes in a joint maximum likelihood estimation. The superiority of our model is captured by improvements in both the overall likelihood and the VIX-implied likelihoods, as well as a better fitting of marginal distributions with up to 15% less error on one-asset option prices. The new degree of freedom in the covariance-dependent kernel is also shown to significantly impact prices of two-asset correlation options leading to up to 53% differences for out-of-the-money claims with short maturity. The added flexibility is also noticeable in the shape of marginal and joint pricing kernels as demonstrated with our empirical estimates.

Keywords: Pricing, multi-asset options, GARCH models, Closed form solutions, Covariance dependent kernel, maximum likelihood estimation

JEL: C15, G12, G13

Electronic copy available at: https://ssrn.com/abstract=3953826

1 Introduction

The class of Generalized Autoregressive Conditional Heteroscedasticity, or GARCH, models of (Engle 1982) and (Bollerslev 1986), and their applications to derivative pricing and risk management has evolved remarkably over the past three decades. Among various specifications of GARCH models, of particular interest is the class of affine GARCH models. The moment generating function in an affine GARCH model can be written in closed form with coefficients given by a set of recursive algebraic equations. Hence, calculation of European option prices and their hedge ratios can be performed in closed-form which is considerably faster and more accurate than Monte Carlo methods, see (Badescu, Elliott, and Ortega 2015), (Menn and Rachev 2005), and (Date and Islyaev 2015) for examples of non-Affine GARCH option pricing.

The first affine GARCH specification was proposed by (Heston and Nandi 2000), henceforth the HN-GARCH model, as a univariate model with Gaussian innovations and a linear pricing kernel. Since then, a number of non-Gaussian affine models or extensions of the HN-GARCH model have been proposed by researchers. (Christoffersen, Heston, and Jacobs 2006) successfully obtained a closed-form moment generating function with Inverse Gaussian shocks. (Ornthanalai 2014) incorporated a combination of Gaussian innovations and Levy jumps in an affine setting, and (Bégin, Dorion, and Gauthier 2020) presented a bivariate extension of this Levy jump GARCH model. A general non-Gaussian conditional affine GARCH model was introduced in (Escobar-Anel, Rastegari, and Stentoft 2021).

The use of non-linear or variance-dependent pricing kernel in affine models was first introduced in (Christoffersen, Heston, and Jacobs 2013) which extends the classical HN-GARCH model. The change of measure in this model is driven by an additional variance premium parameter accounting for a shift in variance of innovations from the physical to the risk-neutral model. It also maintains the affine structure of the model under the equivalent martingale measure. (Badescu, Cui, and Ortega 2019) formulated a general affinity condition for univariate GARCH models which guarantees both a closed-form moment generating function and a variance-dependent kernel.

A general specification of an affine multivariate factor GARCH model is presented in the recent work by (Escobar-Anel, Rastegari, and Stentoft 2020). The model, henceforth the ERS-ARCH model, is based on decomposition of conditional covariance matrix such as spectral or Cholesky decomposition, and expresses the dynamics of the conditional covariance matrix as a linear combination of independent components driven by univariate affine GARCH models. Such factor GARCH models have been explored in (Vrontos, Dellaportas, and Politis 2003) and (Van der Weide 2002) in a non-affine setting. They exhibit cer-

Electronic copy available at: https://ssrn.com/abstract=3953826

tain advantages over other multivariate GARCH specifications, namely, guaranteed positive-definiteness of conditional covariance matrix, reduced number of parameters, straightforward estimation procedure, and meaningful economical interpretation.

The affine multivariate ERS-GARCH model is interestingly equipped with a covariance-dependent pricing kernel which maintains the closed-form structure of the moment generating function under risk-neutral process, thus generalizing the univariate model in (Christoffersen, Heston, and Jacobs 2013) and (Badescu, Cui, and Ortega 2019). However, due to the affine structure of the model, the pricing kernel is somewhat restricted. For example, in the bivariate case the pricing kernel is driven by only two free parameters accounting for changes in the three elements of the conditional covariance matrix (i.e. two variances and a correlation).

Our main objective in this paper is to present a full-parameter covariance-dependent pricing kernel in an affine setting, and to study the impact of the additional flexibility of this pricing kernel on model fit and option pricing. The key observation here is that from a pricing point of view, having a closed-form moment generating function under the risk-neutral measure is desirable but is not necessary in the physical process. Hence, we extend the bivariate ESR GARCH model in (Escobar-Anel, Rastegari, and Stentoft 2020) by relaxing the affinity condition on the physical process while ensuring that the model is affine under the risk-neutral dynamics. The new model admits a flexible covariance-dependent pricing kernel driven by at least three parameters to capture individual assets variance premia and the correlation premium.

To assess the impact of the proposed full-parameter covariance-dependent kernel, we perform a joint estimation on daily returns and volatility index values for three pairs of market indices. Our empirical results confirm statistical significance of all parameters in the pricing kernel, and show a clear improvement in returns likelihood and volatility index likelihood compared to the restricted ERS-GARCH model. Next, the effect on the marginal pricing kernels over periods of one day, one week and one month are examined and shown to be different from those given by existing models especially in the tails. Using single asset call and put options, we demonstrate that the new model has much better option pricing performance with up to 38% improvement in the implied volatility error. Finally, we consider the joint pricing kernel and show that this differs significantly from existing models. Comparing the price of two-asset correlation options calculated from our proposed model we observe a relative difference of up to 53% for out-of-the money short-maturity options when compared to the restricted ERS-GARCH model.

We summarize our contributions as follows:

- We introduce a multivariate model of log returns which admits both a variance and

Electronic copy available at: https://ssrn.com/abstract=3953826

a correlation dependent kernel. In particular, a general pricing kernel is created; this pricing kernel is driven by, at least, three free parameters accounting for shifts from the historical ($\mathbb{P}$) measure to the risk-neutral ($\mathbb{Q}$) measure in the conditional covariance matrix. The model is an extension of the ERS-GARCH model of (Escobar-Anel, Rastegari, and Stentoft 2020).

- Conditions ensuring a proper change from historical to risk-neutral measures are provided and though the model may be non-affine under the historical measure a special subclass allows for affine $\mathbb{Q}$ dynamics. The role of the new parameters in creating differences between the $\mathbb{P}$ and $\mathbb{Q}$ covariances is explained.

- A joint calibration on three bivariate pairs of indexes and their respective volatility indexes, chosen from S&P500 (SPX and VIX), Dow Jones (DJX and VXD), NASDAQ 100 (NDX and VXN) and Russell 2000 (RUT and RVX) is performed. The calibration improves, compared to the embedded ERS-GARCH model, on the likelihood of returns, the likelihood of volatility indexes, and the new parameters are, in most cases, statistically significant.

- The added flexibility of our model is further justified empirically when comparing the implied marginal and joint pricing kernel to the embedded models. This improvement results in significantly reduced errors in univariate option fitting. Moreover, when compared to the restricted ERS-GARCH model significant differences of up to 53% are found when pricing a popular two-assets correlation option.

The rest of the paper is organized as follows. Section 2 presents the general model, provides viable changes of measures along with a specific setting that ensures affine risk-neutral dynamics, and discusses the interpretation of the new parameter. Section 3 applies the model empirically providing parameter estimates from joint calibration of three bivariate series. The implications on the marginal and joint pricing kernels are studied here and the effect on fitting single asset options empirically and of pricing multi-asset options are discussed. Section 4 concludes. The online Appendix contains proofs.

2 Multivariate GARCH models with a covariance-dependent pricing kernel

In this section, we first present our multivariate GARCH model. We then provide general conditions for a viable changes of measure, as well as specific conditions that ensure an

Electronic copy available at: https://ssrn.com/abstract=3953826

affine risk-neutral structure. Finally, an in-depth analysis of the implications of our model for variance-dependent and correlation-dependent kernels is presented.

2.1 The model

For simplicity and to facilitate interpretations, we first present the bivariate model. Let r be the single-period risk-free rate, and let $S_{j,t}$ be the spot price, $x_{j,t} = \log(S_{j,t})$ be the log price, and $R_{j,t+1} = x_{j,t+1} - x_{j,t}$ denote the daily return over a single period of the j'th asset for $j = 1, 2$. We consider the following GARCH dynamics:

$$R_{1,t} = r + \lambda_1 h_{1,t} + \sqrt{h_{1,t}} z_{1,t} \tag{1}$$

$$R_{2,t} = r + \lambda_{21} h_{1,t} + \left(a\sqrt{h_{1,t}} - \eta_t \sqrt{h_{2,t}}\right) z_{1,t} + \lambda_2 h_{2,t} + \sqrt{h_{2,t}} z_{2,t} \tag{2}$$

$$h_{1,t} = \omega_1 + \beta_1 h_{1,t-1} + \alpha_1 \left(z_{1,t-1} - \gamma_1 \sqrt{h_{1,t-1}}\right)^2 \tag{3}$$

$$h_{2,t} = \omega_2 + \beta_2 h_{2,t-1} + \alpha_2 \left(z_{2,t-1} - \eta_{t-1} z_{1,t-1} - \gamma_2 \sqrt{h_{2,t-1}}\right)^2, \tag{4}$$

where $z_{1,t}$ and $z_{2,t}$ are independent sequences of standard normal random variables with $\omega_j, \beta_j, \alpha_j > 0$ to ensure positivity of the variance components $h_{j,t}$. The time-dependent parameter η_t is of the form $\eta_t = g(h_{1,t}, h_{2,t})$ for some generic function g to be specified later. The dynamics of the first asset is the well-known HN-GARCH model of (Heston and Nandi 2000), while the dynamics of the second asset is as prescribed by the ERS-GARCH model of (Escobar-Anel, Rastegari, and Stentoft 2020) except for inclusion of the parameter η_t.

As per the representation above, one interpretation of this model is along the CAPM setting. The first asset can be seen as a market factor like an index. In such case z_1 drives the systematic risk, and the second asset could be a stock, which permits z_2 to be interpreted as the driver of the idiosyncratic risk. For simplicity, we will use this language of systematic and idiosyncratic to refer to z_1 and z_2 although in general the model creates a multivariate structure that does not necessarily fit into the setting of CAPM.

The returns dynamics can be written in matrix form as follows:

$$R_t = r\mathbf{1} + \Lambda \operatorname{diag}(h_t)\mathbf{1} + A_t z_t, \tag{5}$$

where

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 \\ \lambda_{21} & \lambda_2 \end{bmatrix}, \quad \operatorname{diag}(h_t) = \begin{bmatrix} h_{1,t} & 0 \\ 0 & h_{2,t} \end{bmatrix}, \text{and}$$

Electronic copy available at: https://ssrn.com/abstract=3953826

$$A_t = \begin{bmatrix} \sqrt{h_{1,t}} & 0 \\ a\sqrt{h_{1,t}} - \eta_t\sqrt{h_{2,t}} & \sqrt{h_{2,t}} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ a & 1 \end{bmatrix} \operatorname{diag}\left(\sqrt{h_t}\right) \begin{bmatrix} 1 & 0 \\ -\eta_t & 1 \end{bmatrix} = A \operatorname{diag}\left(\sqrt{h_t}\right) \widetilde{\eta}_t,$$

and where $z_t|F_{t-1}$ is an independent sequence of bivariate standard normal random variables. The covariance matrix of returns over a single period is given by

$$H_t = Var(R_t|\mathcal{F}_{t-1}) = A_t\, A_t'.$$

Note that the positive definiteness of the conditional covariance matrix is already satisfied without further parameter constraints in the model.

An alternative representation that helps understand a n-dimensional representation can be obtained by a change of variable $\widetilde{z}_t = \widetilde{\eta}_t z_t$ leading to:

$$R_{1,t} = r + \lambda_1 h_{1,t} + \sqrt{h_{1,t}}\tilde{z}_{1,t} \tag{6}$$

$$R_{2,t} = r + \lambda_{21} h_{1,t} + \lambda_2 h_{2,t} + a\sqrt{h_{1,t}}\tilde{z}_{1,t} + \sqrt{h_{2,t}}\tilde{z}_{2,t} \tag{7}$$

$$h_{1,t} = \omega_1 + \beta_1 h_{1,t-1} + \alpha_1 \left(\tilde{z}_{1,t-1} - \gamma_1\sqrt{h_{1,t-1}}\right)^2 \tag{8}$$

$$h_{2,t} = \omega_2 + \beta_2 h_{2,t-1} + \alpha_2 \left(\tilde{z}_{2,t-1} - \gamma_2\sqrt{h_{2,t-1}}\right)^2. \tag{9}$$

This looks structurally the same as the ERS-GARCH model, albeit with correlated innovations $\widetilde{z}_t$. Hence a n-dimensional generalization can be written as:

$$\begin{aligned} R_t &= r\mathbf{1} + \Lambda \operatorname{diag}(h_t)\mathbf{1} + A_t z_t \\ &= r\mathbf{1} + \Lambda \operatorname{diag}(h_t)\mathbf{1} + A \operatorname{diag}(h_t)\widetilde{z}_t, \end{aligned} \tag{10}$$

where $\widetilde{\eta}_t$ and A are the lower diagonal matrices defined implicitly above.

The representations above show that one of the novelties in our model is the parameter η_t. As mentioned above, this object can be constant or time dependent, e.g. a function of variance components and we write this as $\eta_t = g(h_{1,t}, h_{2,t})$. The case $\eta_t = 0$ corresponds to the ERS-GARCH model and this new parameter creates two effects. First, the innovations of the first asset (z_1 or the systematic factor) impact the dynamics of the variance of the idiosyncratic factor, see Equation (4). In other words, it captures a correlation between the variances of the systematic and idiosyncratic factors. Secondly, the representation in Equation (10) demonstrates that η_t also captures a left-over correlation between the systematic and idiosyncratic innovations.

The general GARCH process in Equations (1)-(4) is not necessarily affine under the physical measure as this property depends on the choice of η_t. To see this, observe that the

Electronic copy available at: https://ssrn.com/abstract=3953826

conditional moment generating function (mgf) over a single period can be calculated as

$$\begin{aligned}
m(\vec{u},\vec{v}) &= E_t^{\mathbb{P}}[e^{ux_{1,t+1}+vx_{2,t+1}}] \\
&= \exp\left(ux_{1,t}+vx_{2,t}+ur+u\lambda_1 h_{1,t}+vr+v\lambda_{21}h_{1,t}+v\lambda_2 h_{2,t}\right) \\
&\times E_t^{\mathbb{P}}\left[\exp\left(u\sqrt{h_{1,t}}z_{1,t}+v\left(a\sqrt{h_{1,t}}-\eta_t\sqrt{h_{2,t}}\right)z_{1,t}+v\sqrt{h_{2,t}}z_{2,t}\right)\right] \\
&= \exp\left(ux_{1,t}+vx_{2,t}+ur+u\lambda_1 h_{1,t}+vr+v\lambda_{21}h_{1,t}+v\lambda_2 h_{2,t}\right) \\
&\times \exp\left(\frac{1}{2}(u+va)^2 h_{1,t}+\frac{1}{2}v^2(1+\eta_t^2)h_{2,t}-(u+va)\eta_t\sqrt{h_{1,t}h_{2,t}}\right).
\end{aligned}$$

For most specifications of η_t, this moment generating function would not be of the affine exponential form mainly due to the last term involving $\sqrt{h_{1,t}h_{2,t}}$. However, as we demonstrate in the following it is possible to propose a general change of measure for which the risk-neutral dynamics of our model can be made affine under specific conditions. This captures a key contribution of our paper; proposing a model which is not necessarily affine under the historical measure, and thus more flexible for a better fit to empirical data, but can be made affine under the risk-neutral measure, and therefore continues to allow for efficient pricing of derivatives.

2.2 Correlation-dependent pricing kernel

We now provide a general correlation-dependent pricing kernel for the model in Equations (1)-(4), derive the risk-neutral distribution and dynamics, and present a subclass under which the risk-neutral dynamics are affine. First, we present our proposed Radon-Nikodym derivative in the following Proposition.

Proposition 2.1. *Assume that asset returns follow the bivariate GARCH model in Equations* (1)-(4)*, and let* $C_{t-1} = C_{t-1}(\theta_{1,t},\theta_{2,t},\xi_{1,t},\xi_{2,t},\xi_{3,t})$ *be an* $\mathcal{F}_{t-1}$*-adaptable process defined as*

$$\exp(C_{t-1}) = E_{t-1}^{\mathbb{P}}\left[\exp\left(\theta_{1,t}z_{1,t}+\theta_{2,t}z_{2,t}+\xi_{3,t}z_{1,t}z_{2,t}+\xi_{1,t}h_{1,t+1}+\xi_{2,t}h_{2,t+1}\right)\right], \tag{11}$$

where the parameters $\theta_{1,t}$*,* $\theta_{2,t}$*,* $\xi_{1,t}$*,* $\xi_{2,t}$*, and* $\xi_{3,t}$ *can be functions of* $h_{1,t}$ *and* $h_{2,t}$*. Suppose for* $t = 1,\ldots,T$ *that the following conditions hold:*

$$\lambda_1 h_{1,t} - C_{t-1}(\theta_{1,t},\theta_{2,t},\xi_{1,t},\xi_{2,t},\xi_{3,t}) + C_{t-1}(\theta_{1,t}+\sqrt{h_{1,t}},\theta_{2,t},\xi_{1,t},\xi_{2,t},\xi_{3,t}) = 0, \tag{12}$$

$$\begin{aligned}
&\lambda_{21}h_{1,t}+\lambda_2 h_{2,t} - C_{t-1}(\theta_{1,t},\theta_{2,t},\xi_{1,t},\xi_{2,t},\xi_{3,t}) \\
&+ C_{t-1}\left(\theta_{1,t}+a\sqrt{h_{1,t}}-\eta_t\sqrt{h_{2,t}},\theta_{2,t}+\sqrt{h_{2,t}},\xi_{1,t},\xi_{2,t},\xi_{3,t}\right) = 0.
\end{aligned} \tag{13}$$

Electronic copy available at: https://ssrn.com/abstract=3953826

Then the Radon-Nikodym derivative process

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = \exp\left(\sum_{s=1}^{t} \theta_{1,t} z_{1,s} + \theta_{2,t} z_{2,s} + \xi_{3,t} z_{1,s} z_{2,s} + \xi_{1,t} h_{1,s+1} + \xi_{2,t} h_{2,s+1} - C_{s-1}(\theta_{1,s}, \theta_{2,s}, \xi_{1,s}, \xi_{2,s}, \xi_{3,s})\right), \tag{14}$$

defines an equivalent martingale measure $\mathbb{Q}$*.*

Proof. See A for the proof. □

Note that given the expression for C_{t-1}, one can write Equations (12) and (13) in an explicit form and solve for $\theta_{1,t}$ and $\theta_{2,t}$ while having $\xi_{1,t}$, $\xi_{2,t}$, and $\xi_{3,t}$ as free parameters.

We can calculate the explicit form of C_{t-1} in our model. To do this first observe that

$$\begin{aligned}\exp(C_{t-1}) = &\exp\left(\xi_{1,t}\omega_1 + \xi_{1,t}\left(\beta_1 + \alpha_1\gamma_1^2\right) h_{1,t} + \xi_{2,t}\omega_2 + \xi_{2,t}\left(\beta_2 + \alpha_2\gamma_2^2\right) h_{2,t}\right) \\ &\times E_{t-1}^{\mathbb{P}}\left[\exp\left(\left(\xi_{1,t}\alpha_1 + \xi_{2,t}\alpha_2\eta_t^2\right) z_{1,t}^2 + \left(\theta_{1,t} - 2\xi_{1,t}\alpha_1\gamma_1\sqrt{h_{1,t}} + 2\xi_{2,t}\alpha_2\gamma_2\eta_t\sqrt{h_{2,t}}\right) z_{1,t}\right.\right. \\ &\left.\left. + \ \xi_{2,t}\alpha_2 z_{2,t}^2 + \left(\theta_{2,t} - 2\xi_{2,t}\alpha_2\gamma_2\sqrt{h_{2,t}}\right) z_{2,t} + \left(\xi_{3,t} - 2\xi_{2,t}\alpha_2\eta_t\right) z_{1,t} z_{2,t}\right)\right].\end{aligned}$$

Next, using the bivariate relationship

$$\begin{aligned}&E[\exp(a_1 z_1^2 + b_1 z_1 + a_2 z_2^2 + b_2 z_2 + c z_1 z_2)] \\ &\quad = \exp\left(-\frac{1}{2}\log\left((1-2a_1)(1-2a_2) - c^2\right) + \frac{(1-2a_2)b_1^2 + (1-2a_1)b_2^2 + 2cb_1b_2}{2\left((1-2a_1)(1-2a_2) - c^2\right)}\right),\end{aligned} \tag{15}$$

which follows from the univariate version $E[\exp(az^2 + bz)] = \exp\left(-\frac{1}{2}\log(1-2a) + \frac{b^2}{2(1-2a)}\right)$ by conditioning, we get

$$\begin{aligned}C_{t-1} = &\xi_{1,t}\omega_1 + \xi_{1,t}\left(\beta_1 + \alpha_1\gamma_1^2\right) h_{1,t} + \xi_{2,t}\omega_2 + \xi_{2,t}\left(\beta_2 + \alpha_2\gamma_2^2\right) h_{2,t} - \frac{1}{2}\log K_t \\ &+ \frac{1}{2K_t}\left[\left(1 - 2\xi_{2,t}\alpha_2\right)\left(\theta_{1,t} + B_{1,t}\right)^2 + \left(1 - 2\xi_{1,t}\alpha_1 - 2\xi_{2,t}\alpha_2\eta_t^2\right)\left(\theta_{2,t} + B_{2,t}\right)^2\right. \\ &\left. + 2\left(\xi_{3,t} - 2\xi_{2,t}\alpha_2\eta\right)\left(\theta_{1,t} + B_{1,t}\right)\left(\theta_{2,t} + B_{2,t}\right)\right],\end{aligned} \tag{16}$$

with

$$B_{1,t} = -2\xi_{1,t}\alpha_1\gamma_1\sqrt{h_{1,t}} + 2\xi_{2,t}\alpha_2\gamma_2\eta_t\sqrt{h_{2,t}} \tag{17}$$

$$B_{2,t} = -2\xi_{2,t}\alpha_2\gamma_2\sqrt{h_{2,t}} \tag{18}$$

$$K_t = (1 - 2\xi_{1,t}\alpha_1 - 2\xi_{2,t}\alpha_2\eta_t^2)(1 - 2\xi_{2,t}\alpha_2) - \left(\xi_{3,t} - 2\xi_{2,t}\alpha_2\eta_t\right)^2. \tag{19}$$

Electronic copy available at: https://ssrn.com/abstract=3953826

Using the above calculations, we can now obtain the conditional distribution of the innovation term $(z_{1,t}, z_{2,t})$ under our change of measure.

Proposition 2.2. *Under the change of measure in Proposition 2.1 the risk-neutral distribution of the innovation term $(z_{1,t}, z_{2,t})$ is bivariate normal described by*

$$z_{1,t} = \sqrt{d_{1,t}} z^*_{1,t} + b_{1,t}\sqrt{h_{1,t}} \tag{20}$$

$$z_{2,t} = \sqrt{d_{1,t}}\left(\frac{d_{3,t}}{d_{1,t}} + \eta_t\right) z^*_{1,t} + b_{1,t}\eta_t\sqrt{h_{1,t}} + \sqrt{d_{2,t}}\sqrt{1 - \left(\frac{d^2_{3,t}}{d_{1,t}d_{2,t}} + \frac{2\eta_t d_{3,t}}{d_{2,t}}\right)} z^*_{2,t} + b_{2,t}\sqrt{h_{2,t}} \tag{21}$$

where $z^*_{1,t}$ and $z^*_{2,t}$ are independent standard normal random variables, and where the parameters $d_{1,t}$, $d_{2,t}$, $d_{3,t}$, $b_{1,t}$, and $b_{2,t}$ are mapped to $\xi_{1,t}$, $\xi_{2,t}$, $\xi_{3,t}$, $\theta_{1,t}$, and $\theta_{2,t}$ via Equations (80)-(84) *in the Appendix.*

Proof. See A for the proof. □

The terms $d_{1,t}$ and $b_{1,t}$ are pricing kernel generalizations of the one-dimensional constant in (Christoffersen, Heston, and Jacobs 2013). Parameters $d_{2,t}$ and $b_{2,t}$ similarly extend the constant counterparts in (Escobar-Anel, Rastegari, and Stentoft 2020). An additional flexibility in terms of pricing kernel is captured by $d_{3,t}$, while η_t is the new modeling parameter. Substituting the innovation term $(z_{1,t}, z_{2,t})$ in the data generating process in Equations (1)-(4) gives the risk-neutral dynamics as

$$R_{1,t} = r + (\lambda_{1,t} + b_{1,t})\, h_{1,t} + \sqrt{d_{1,t}h_{1,t}} z^*_{1,t} \tag{22}$$

$$\begin{aligned} R_{2,t} = {} & r + (\lambda_{21,t} + a b_{1,t})\, h_{1,t} + (\lambda_2 + b_{2,t})\, h_{2,t} + \sqrt{d_{1,t}}\left(a\sqrt{h_{1,t}} + \frac{d_{3,t}}{d_{1,t}}\sqrt{h_{2,t}}\right) z^*_{1,t} \\ & + \sqrt{d_{2,t}}\sqrt{1 - \left(\frac{d^2_{3,t}}{d_{1,t}d_{2,t}} + \frac{2\eta_t d_{3,t}}{d_{2,t}}\right)} z^*_{2,t} \end{aligned} \tag{23}$$

$$h_{1,t} = \omega_1 + \beta_1 h_{1,t-1} + d_{1,t-1}\alpha_1 \left(z^*_{1,t-1} - (\gamma_1 - b_{1,t-1})\sqrt{\frac{h_{1,t-1}}{d_{1,t-1}}}\right)^2 \tag{24}$$

$$h_{2,t} = \omega_2 + \beta_2 h_{2,t-1} + \alpha_2 \left(\sqrt{d_{2,t}}\sqrt{1 - \left(\frac{d^2_{3,t}}{d_{1,t}d_{2,t}} + \frac{2\eta_t d_{3,t}}{d_{2,t}}\right)} z^*_{2,t-1} + \frac{d_{3,t}}{\sqrt{d_{1,t}}} z^*_{1,t-1} - (\gamma_2 - b_{2,t-1})\sqrt{h_{2,t-1}}\right)^2, \tag{25}$$

with certain relations between risk premia parameters $\lambda_{1,t}$, $\lambda_{21,t}$, and $\lambda_{2,t}$ and the parameters $b_{1,t}$, $b_{2,t}$, $b_{3,t}$, $d_{1,t}$, and $d_{2,t}$ that ensure the martingale property. These relations can be

Electronic copy available at: https://ssrn.com/abstract=3953826

obtained from Equations (12) and (13), or equivalently from calculation of the expected conditional discounted price in Equations (22)-(25).

Notice that the model is non-affine under the equivalent martingale measure in general. However, a simple choice of $d_{3,t}$ (i.e. $\xi_{3,t}$) leads to an affine model under $\mathbb{Q}$ as demonstrated in the following Corollary.

Corollary 2.3. *Assume $d_{3,t} = 0$ (i.e. $\xi_{3,t} = \eta_t$) with $\eta_t = g(h_{1,t}, h_{2,t})$ and let d_1, d_2, b_1, and b_2 be constant parameters. Under the change of measure in Proposition 2.1, the risk-neutral dynamics reduces to the affine ERS-GARCH process described by*

$$R_{1,t} = r - \frac{1}{2}h^*_{1,t} + \sqrt{h^*_{1,t}}z^*_{1,t} \tag{26}$$

$$R_{2,t} = r - \frac{1}{2}a^2 h^*_{1,t} - \frac{1}{2}h^*_{2,t} + a\sqrt{h^*_{1,t}}z^*_{1,t} + \sqrt{h^*_{2,t}}z^*_{2,t} \tag{27}$$

$$h^*_{i,t} = \omega^*_i + \beta_i h^*_{i,t-1} + \alpha^*_i \left(z^*_{i,t-1} - \gamma^*_i \sqrt{h^*_{i,t-1}}\right)^2, \quad i = 1, 2, \tag{28}$$

where $z^*_{1,t}$ and $z^*_{2,t}$ are independent standard normal random variables, and the risk-neutral variables and parameters are defined as

$$h^*_{j,t} = d_j h_{j,t}, \quad \alpha^*_j = d_j^2 \alpha_j, \quad \omega^*_j = d_j \omega_j, \quad \gamma^*_j = \frac{\gamma_j - b_j}{d_j}, \tag{29}$$

with the following relations between equity risk premium parameters in the physical model

$$b_1 = -\lambda_1 - \frac{1}{2}d_1, \quad \lambda_{21} = -\frac{1}{2}a^2 d_1 - a b_1, \quad b_2 = -\lambda_2 - \frac{1}{2}d_2. \tag{30}$$

Proof. See A for the proof. □

Remark 2.4. Notice that the risk-neutral process in the above corollary can be written in matrix form as

$$R_t = r\mathbf{1} + \Lambda^* \operatorname{diag}(h^*_t)\mathbf{1} + A^*_t z^*_t, \tag{31}$$

where

$$\Lambda^* = \begin{bmatrix} -\frac{1}{2} & 0 \\ -\frac{1}{2}a^2 & -\frac{1}{2} \end{bmatrix}, \quad \operatorname{diag}(h^*_t) = \begin{bmatrix} h^*_{1,t} & 0 \\ 0 & h^*_{2,t} \end{bmatrix}, \quad A^*_t = \begin{bmatrix} \sqrt{h^*_{1,t}} & 0 \\ a\sqrt{h^*_{1,t}} & \sqrt{h^*_{2,t}} \end{bmatrix},$$

and where $z^*_t | F_{t-1}$ is an independent sequence of bivariate standard normal random variables. Moreover, we can readily write the joint moment generating function of log-price under $\mathbb{Q}$

Electronic copy available at: https://ssrn.com/abstract=3953826

as

$$m^{\mathbb{Q}}(u_1,u_2)=E_t^{\mathbb{Q}}[e^{u_1x_{1,T}+u_2x_{2,T}}]=\exp\left(u_1x_{1,t}+u_2x_{2,t}+J_{T-t}+K_{1,T-t}h_{1,t+1}+K_{2,T-t}h_{2,t+1}\right), \tag{32}$$

where J_{T-t}, $K_{1,T-t}$ and $K_{2,T-t}$ are given recursively by

$$J_k=J_{k-1}+r\sum_{i=1}^{2}u_i+\sum_{j=1}^{2}\omega_jK_{j,k-1}-\frac{1}{2}\sum_{j=1}^{2}\log(1-2K_{j,k-1}\alpha_j) \tag{33}$$

$$K_{1,k}=-\frac{1}{2}u_1-\frac{1}{2}a^2u_2+K_{1,k-1}(\beta_1+\alpha_1\gamma_1^2)+\frac{(u_1+au_2-2K_{1,k-1}\alpha_1\gamma_1)^2}{2(1-2K_{1,k-1}\alpha_1)} \tag{34}$$

$$K_{2,k}=-\frac{1}{2}u_2+K_{2,k-1}(\beta_2+\alpha_2\gamma_2^2)+\frac{(u_2-2K_{2,k-1}\alpha_2\gamma_2)^2}{2(1-2K_{j,k-1}\alpha_2)}, \tag{35}$$

for $k=0,1,\ldots,T-t$ with initial values $J_0=K_{1,0}=K_{2,0}=0$.

It is important to observe the additional flexibility of our model (due to presence of η_t) compared to the ERS-GARCH model. To demonstrate this, note that under the change of measure in Proposition 2.1 and Corollary 2.3 the risk-neutral conditional distribution of the innovation term $(z_{1,t},z_{2,t})$ is represented by

$$z_{1,t}=\sqrt{d_1}z_{1,t}^*+b_1\sqrt{h_{1,t}} \tag{36}$$

$$z_{2,t}=\eta_t\left(\sqrt{d_1}z_{1,t}^*+b_1\sqrt{h_{1,t}}\right)+\sqrt{d_2}z_{2,t}^*+b_2\sqrt{h_{2,t}} \tag{37}$$

where $z_{1,t}^*$ and $z_{2,t}^*$ are independent standard normal random variables.[1] When $\eta_t=0$ the model is reduced to the ERS-GARCH model under both physical and risk-neutral measures. In this case the change of measure imposes a shift in the conditional mean and variance of innovations $z_{1,t}$ and $z_{2,t}$ while the conditional correlation between innovations remains the same. However, when $\eta_t\neq 0$, we also get a change in correlation of shocks from $\mathbb{P}$ to $\mathbb{Q}$. We will further analyze the effect of η_t on the conditional correlation in the following section.

2.3 Impact on conditional covariance matrix

We finish this section by studying the impact of the new parameter, η_t, on the pricing kernel under the prescribed assumption $\xi_{3,t}=\eta_t$. Specifically, we are interested in the changes it imposes on variance and correlation when shifting from the $\mathbb{P}$ to the $\mathbb{Q}$ measure. The $\mathcal{F}_{t-1}$

[1] This is straightforward to obtain by substituting $d_{3,t}=0$ in Equations (20) and (21).

Electronic copy available at: https://ssrn.com/abstract=3953826

conditional covariance matrices under $\mathbb{P}$ and $\mathbb{Q}$ are given by

$$\begin{aligned} Cov^P_{t-1}(R_t) &= \begin{bmatrix} h_{1,t} & ah_{1,t} - \eta_t\sqrt{h_{1,t}h_{2,t}} \\ ah_{1,t} - \eta_t\sqrt{h_{1,t}h_{2,t}} & a^2h_{1,t} + (1+\eta_t^2)h_{2,t} - 2a\eta_t\sqrt{h_{1,t}h_{2,t}} \end{bmatrix} \\ &= \begin{bmatrix} h_{1,t} & ah_{1,t} - \eta_t\sqrt{h_{1,t}h_{2,t}} \\ ah_{1,t} - \eta_t\sqrt{h_{1,t}h_{2,t}} & \left(a\sqrt{h_{1,t}} - \eta_t\sqrt{h_{2,t}}\right)^2 + h_{2,t} \end{bmatrix} \text{ and} \end{aligned} \tag{38}$$

$$Cov^Q_{t-1}(R_t) = \begin{bmatrix} d_1h_{1,t} & ad_1h_{1,t} \\ ad_1h_{1,t} & a^2d_1h_{1,t} + d_2h_{2,t} \end{bmatrix}, \tag{39}$$

respectively. Hence, the change in the conditional variance of each asset and the change in correlation from $\mathbb{P}$ to $\mathbb{Q}$ are given by

$$\Delta Var_{t-1}(R_{1,t}) = Var^Q_{t-1}(R_{1,t}) - Var^P_{t-1}(R_{1,t}) = (d_1 - 1)h_{1,t} \tag{40}$$

$$\Delta Var_{t-1}(R_{2,t}) = Var^Q_{t-1}(R_{2,t}) - Var^P_{t-1}(R_{2,t}) = a^2(d_1-1)h_{1,t} + (d_2 - 1 - \eta_t^2)h_{2,t} + 2a\eta_t\sqrt{h_{1,t}h_{2,t}} \tag{41}$$

$$\Delta Corr_{t-1}(R_t) = Corr^Q_{t-1}(R_t) - Corr^P_{t-1}(R_t) = \frac{a}{\sqrt{a^2 + \frac{d_2}{d_1}\frac{h_{2,t}}{h_{1,t}}}} - \frac{a - \eta_t\sqrt{\frac{h_{2,t}}{h_{1,t}}}}{\sqrt{\left(a - \eta_t\sqrt{\frac{h_{2,t}}{h_{1,t}}}\right)^2 + \frac{h_{2,t}}{h_{1,t}}}}. \tag{42}$$

Let us interpret the parameters in the previous equations to provide context on existing results and the contributions of our work.

First, the parameter d_1 impacts Equations (40), (41) and (42), but the clearer interpretation comes from Equation (40). This equation highlights the main innovation in (Christoffersen, Heston, and Jacobs 2013). A parameter $d_1 \neq 1$ makes the pricing kernel dependent on the variance of the first asset, h_1 (i.e. $\xi_1 \neq 0$ in Proposition 2.1), and it also generates a variance premium on the first asset (i.e. a change in variance from $\mathbb{P}$ to $\mathbb{Q}$). Hence, d_1 serves as a degree of freedom to control the variance of the first asset and its changes from $\mathbb{P}$ to $\mathbb{Q}$.

Similarly, d_2 influences Equations (41) and (42), so in order to facilitate its interpretation, we study d_2 as a stand alone parameter, i.e. we set $d_1 = 1$ and $\eta_t = 0$. In this case we get

$$\Delta Var_{t-1}(R_{2,t}) = Var^Q_{t-1}(R_{2,t}) - Var^P_{t-1}(R_{2,t}) = (d_2 - 1)h_{2,t} \tag{43}$$

$$\Delta Corr_{t-1}(R_t) = Corr^Q_{t-1}(R_t) - Corr^P_{t-1}(R_t) = \frac{a}{\sqrt{a^2 + d_2\frac{h_{2,t}}{h_{1,t}}}} - \frac{a}{\sqrt{a^2 + \frac{h_{2,t}}{h_{1,t}}}}. \tag{44}$$

The d_2 parameter is the main innovation in (Escobar-Anel, Rastegari, and Stentoft 2020).

Electronic copy available at: https://ssrn.com/abstract=3953826

A $d_2 \neq 1$ makes the pricing kernel depend on the idiosyncratic variance of the second asset, h_2 (i.e. $\xi_2 \neq 0$ in Proposition 2.1), and it also creates a variance premium on the second asset (i.e. a change in variance from $\mathbb{P}$ to $\mathbb{Q}$). However, as Equation (44) shows d_2 also generates a premium on correlation making the pricing kernel depend on the covariance (i.e. $\eta_3 \neq 0$ in Proposition 2.1). Unfortunately, d_2 is only a single second degree of freedom to control changes from $\mathbb{P}$ to $\mathbb{Q}$ on the two targets: the variance of second asset and the correlation. In other words, we need a third degree of freedom to separate variance premium from correlation premium.

Now we are ready to understand the impact of the new parameter, η_t, and its role as a third degree of freedom designed to control, mainly, changes in correlation. Setting $d_1 = d_2 = 1$, we obtain

$$\Delta Var_{t-1}\left(R_{2,t}\right) = Var_{t-1}^{Q}\left(R_{2,t}\right) - Var_{t-1}^{P}\left(R_{2,t}\right) = -\eta_t^2 h_{2,t} + 2a\eta_t\sqrt{h_{1,t}h_{2,t}} \tag{45}$$

$$\Delta Corr_{t-1}\left(R_t\right) = Corr_{t-1}^{Q}\left(R_t\right) - Corr_{t-1}^{P}\left(R_t\right) = \frac{a}{\sqrt{a^2 + \frac{h_{2,t}}{h_{1,t}}}} - \frac{a - \eta_t\sqrt{\frac{h_{2,t}}{h_{1,t}}}}{\sqrt{\left(a - \eta_t\sqrt{\frac{h_{2,t}}{h_{1,t}}}\right)^2 + \frac{h_{2,t}}{h_{1,t}}}}. \tag{46}$$

This shows that $\eta_t \neq 0$ also generates a variance premium on the second asset and a premium on correlation. Note that the changes in the variance and correlation resulting from having $\eta_t \neq 0$ are decidedly different from that of having $d_2 \neq 1$. Thus, this new parameter indeed constitutes a third degree of freedom to control changes from $\mathbb{P}$ to $\mathbb{Q}$ that liberates, for the first time, marginal (variance) premiums from joint (correlation) premium.

In our empirical application, we use a two-parameter specification for η_t in which we let $\eta_t = \eta_1 + \eta_2\sqrt{\frac{h_{1,t}}{h_{2,t}}}$, where η_1 and η_2 are constants. With this specification the dynamics of the second asset in Equations (2) and (4) has the explicit form

$$R_{2,t} = r + \lambda_{21}h_{1,t} + \left((a - \eta_2)\sqrt{h_{1,t}} - \eta_1\sqrt{h_{2,t}}\right)z_{1,t} + \lambda_2 h_{2,t} + \sqrt{h_{2,t}}z_{2,t} \tag{47}$$

$$h_{2,t} = \omega_2 + \beta_2 h_{2,t-1} + \alpha_2\left(z_{2,t-1} - \left(\eta_1 + \eta_2\sqrt{\frac{h_{1,t}}{h_{2,t}}}\right)z_{1,t-1} - \gamma_2\sqrt{h_{2,t-1}}\right)^2. \tag{48}$$

Moreover, with this specification the $\mathcal{F}_{t-1}$ conditional covariance matrices under $\mathbb{P}$ and $\mathbb{Q}$

Electronic copy available at: https://ssrn.com/abstract=3953826

are given by

$$Cov_{t-1}^{P}\left(R_t\right) = \begin{bmatrix} h_{1,t} & \left(a-\eta_2\right)h_{1,t} - \eta_1\sqrt{h_{1,t}h_{2,t}} \\ \left(a-\eta_2\right)h_{1,t} - \eta_1\sqrt{h_{1,t}h_{2,t}} & \left(\left(a-\eta_2\right)\sqrt{h_{1,t}} - \eta_1\sqrt{h_{2,t}}\right)^2 + h_{2,t} \end{bmatrix} \tag{49}$$

$$Cov_{t-1}^{Q}\left(R_t\right) = \begin{bmatrix} d_1 h_{1,t} & a d_1 h_{1,t} \\ a d_1 h_{1,t} & a^2 d_1 h_{1,t} + d_2 h_{2,t} \end{bmatrix}, \tag{50}$$

and the change in the conditional variance of each asset and the change in correlation (from $\mathbb{P}$ to $\mathbb{Q}$) is given by

$$\Delta Var_{t-1}\left(R_{1,t}\right) = (d_1 - 1)h_{1,t} \tag{51}$$

$$\Delta Var_{t-1}\left(R_{2,t}\right) = \left(a^2(d_1-1) + \left(2a-\eta_2\right)\eta_2\right)h_{1,t} + \left(d_2 - 1 - \eta_1^2\right)h_{2,t} + 2\left(a-\eta_2\right)\eta_1\sqrt{h_{1,t}h_{2,t}} \tag{52}$$

$$\Delta Corr_{t-1}\left(R_t\right) = \frac{a}{\sqrt{a^2 + \frac{d_2}{d_1}\frac{h_{2,t}}{h_{1,t}}}} - \frac{a - \left(\eta_1\sqrt{\frac{h_{2,t}}{h_{1,t}}} + \eta_2\right)}{\sqrt{\left(a - \left(\eta_1\sqrt{\frac{h_{2,t}}{h_{1,t}}} + \eta_2\right)\right)^2 + \frac{h_{2,t}}{h_{1,t}}}}, \tag{53}$$

respectively.

Our particular parametrization is driven by two considerations. First, the special case of a constant η_t occurs when $\eta_2 = 0$. Second, and more interestingly, when $\eta_1 = 0$, and letting $\widehat{a} = a - \eta_2$, the dynamics of the second asset under the historical measure are given by

$$R_{2,t} = r + \lambda_{21}h_{1,t} + \widehat{a}\sqrt{h_{1,t}}z_{1,t} + \lambda_2 h_{2,t} + \sqrt{h_{2,t}}z_{2,t} \tag{54}$$

$$h_{2,t} = \omega_2 + \beta_2 h_{2,t-1} + \alpha_2\left(z_{2,t-1} - \eta_2\sqrt{\frac{h_{1,t}}{h_{2,t}}}z_{1,t-1} - \gamma_2\sqrt{h_{2,t-1}}\right)^2. \tag{55}$$

The risk neutral dynamics are given by

$$R_{2,t} = r - \frac{1}{2}a^{*2}h_{1,t}^{*} - \frac{1}{2}h_{2,t}^{*} + a^{*}\sqrt{h_{1,t}^{*}}z_{1,t}^{*} + \sqrt{h_{2,t}^{*}}z_{2,t}^{*} \tag{56}$$

$$h_{2,t}^{*} = \omega_2^{*} + \beta_2 h_{2,t-1}^{*} + \alpha_2^{*}\left(z_{2,t-1}^{*} - \gamma_2^{*}\sqrt{h_{2,t-1}^{*}}\right)^2, \tag{57}$$

where risk-neutral variables and parameters are defined as

$$h_{1,t}^{*} = d_1 h_{1,t}, \quad h_{2,t}^{*} = d_2 h_{2,t}, \quad \alpha_2^{*} = d_2^2\alpha_2, \quad \omega_2^{*} = d_2\omega_2, \quad \gamma_2^{*} = \frac{\gamma_2 + \lambda_2}{d_2} + \frac{1}{2}, \quad a^{*} = \widehat{a} + \eta_2. \tag{58}$$

This shows quite clearly that η_2 indeed controls the changes in correlation between $\mathbb{P}$ and

Electronic copy available at: https://ssrn.com/abstract=3953826

$\mathbb{Q}$. The fact that this specification is supported empirically providing a better fit to the data than either setting $\eta_1 = 0$ or $\eta_2 = 0$ lends further credibility to this choice.

3 Empirical analysis

In this section we consider an empirical application of our model with the two-parameter specification for η_t. We first demonstrate how the parameters of the model can be estimated by maximizing the likelihood of returns and volatility indices. We then examine the effect that the new parameter η_t has on the marginal pricing kernel and illustrate the improved performance in terms of pricing options on one asset. Finally, we examine the effect on the bivariate pricing kernel and document the impact of having a covariance dependent pricing kernel on (bivariate) correlation options.

3.1 Parameter estimation

To estimate the model parameters, we use daily returns of S&P 500 Index (SPX) as the first asset, and the daily returns of the Dow Jones Industrial Average (DJX), NASDAQ 100 (NDX) and Russell 2000 (RUT), respectively, as the second asset. To incorporate options data into our estimation, we use daily values of the volatility index for each of these indices. Historical values of these volatility indices with symbols VIX, VXD, VXN and RVX, respectively, are publicly available on the Chicago Board Options Exchange (CBOE) website. The volatility index data covers the period from June 2010 to December 2020. Following (Christoffersen, Heston, and Jacobs 2013) and (Escobar-Anel, Rastegari, and Stentoft 2020) we use a longer period for the returns and use data from January 2000 to December 2020. This enhances the stability of the optimization and helps in proper identification of equity premium parameters.

To obtain the return likelihood given the series of daily returns for the two assets, $\{R_t : t = 1, \ldots, T\}$, observe that $R_t|_{\mathcal{F}_{t-1}} \sim N(r\mathbf{1}+\Lambda \operatorname{diag}(h_t)\mathbf{1}, A_t A_t')$. Hence, the return likelihood is

$$\begin{aligned} L_R(\Theta) &= -\frac{1}{2}\sum_{t=1}^{T}\left(2\log(2\pi) + \log(\det(A_t A_t')) + z_t' z_t\right) \\ &= -\frac{1}{2}\sum_{t=1}^{T}\left(2\log(2\pi) + \log(h_{1,t}) + \log(h_{2,t}) + z_{1,t}^2 + z_{2,t}^2\right), \end{aligned} \tag{59}$$

Electronic copy available at: https://ssrn.com/abstract=3953826

where the vector of residuals is given by

$$z_t = A_t^{-1}(R_t - r\mathbf{1} - \Lambda \operatorname{diag}(h_t)\mathbf{1}), \quad A_t^{-1} = \begin{bmatrix} \frac{1}{\sqrt{h_{1,t}}} & 0 \\ -\frac{a}{\sqrt{h_{2,t}}} + \frac{\eta_t}{\sqrt{h_{1,t}}} & \frac{1}{\sqrt{h_{2,t}}} \end{bmatrix}, \tag{60}$$

or equivalently

$$\begin{aligned} z_{1,t} &= \frac{R_{1,t} - r - \lambda_{11} h_{1,t}}{\sqrt{h_{1,t}}} \\ z_{2,t} &= \frac{R_{2,t} - r - \lambda_{21} h_{1,t} - \lambda_{22} h_{2,t} - \left(a\sqrt{h_{1,t}} - \eta_t \sqrt{h_{2,t}}\right) z_{1,t}}{\sqrt{h_{2,t}}}, \end{aligned}$$

and where the entries of h_t are updated from

$$h_{1,t} = \omega_1 + \beta_1 h_{1,t-1} + \alpha_1 \left(z_{1,t-1} - \gamma_1 \sqrt{h_{1,t-1}}\right)^2 \tag{61}$$

$$h_{2,t} = \omega_2 + \beta_2 h_{2,t-1} + \alpha_2 \left(z_{2,t-1} - \left(\eta_1 + \eta_2 \sqrt{\frac{h_{1,t}}{h_{2,t}}}\right) z_{1,t-1} - \gamma_2 \sqrt{h_{2,t-1}}\right)^2. \tag{62}$$

We set the starting values of $h_{j,t}$ to be their long-run values when $\eta_t = 0$ given by

$$h_1 = \frac{\omega_1 + \alpha_1}{1 - \beta_1 - \alpha_1 \gamma_1^2}, \text{ and } \quad h_2 = \frac{\omega_2 + \alpha_2}{1 - \beta_2 - \alpha_2 \gamma_2^2},$$

respectively.

For the VIX likelihood, we use the bivariate specification from (Escobar-Anel, Rastegari, and Stentoft 2020) which extends the univariate formulation in (Hao and Zhang 2013), (Kanniainen, Lin, and Yang 2014) and (Badescu, Chen, Couch, and Cui 2019). More precisely, 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 model-implied VIX values, then we define the vector of VIX errors as

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

Since the dynamics under $\mathbb{Q}$ in our model correspond to those from the ESR-GARCH model in (Escobar-Anel, Rastegari, and Stentoft 2020) the closed form formula for the volatility index (VIX) of the two assets is given by

$$VIX_{j,t} = 100\sqrt{252\overline{H_{j,t}^*}}, \tag{63}$$

Electronic copy available at: https://ssrn.com/abstract=3953826

where $\overline{H^*_{1,t}}$ and $\overline{H^*_{2,t}}$ can be calculated from

$$\overline{H^*_{1,t}} = \mu_{1,h} + \frac{1-p_1^\tau}{\tau(1-p_1)}(h^*_{1,t+1} - \mu_{1,h}) \tag{64}$$

$$\overline{H^*_{2,t}} = a^2\mu_{1,h} + a^2\frac{1-p_1^\tau}{\tau(1-p_1)}(h^*_{1,t+1} - \mu_{1,h}) + \mu_{2,h} + \frac{1-p_2^\tau}{\tau(1-p_2)}(h^*_{2,t+1} - \mu_{2,h}), \tag{65}$$

where p_j and $\mu_{j,h}$ are the risk-neutral persistence and long-run value of each factor variance $h^*_{j,t}$ given by

$$p_j = \beta^*_j + \alpha^*_j\gamma^{*2}_j, \text{ and } \quad \mu_{j,h} = \frac{\omega_j + \alpha_i}{1-p_j}, \text{ for } \quad j = 1, 2. \tag{66}$$

We allow the individual VIX errors to be autocorrelated of first order with coefficient vector ρ. We assume the individual error components are independent and normally distributed with zero mean such that $\varepsilon^V_t \sim N(0, \Sigma_V)$, where $\Sigma_V = diag(\sigma_1^2, \sigma_2^2)$ is a diagonal matrix of error variances. Then the VIX likelihood function is given by

$$\begin{aligned} L_V(\Theta^*) = &-\frac{1}{2}\sum_{i=1}^{2}\left(\log(2\pi) + \log(\sigma_i^2) + \frac{\left(\varepsilon^V_{i,1}\right)^2}{\sigma_i^2}\right) \\ &-\frac{1}{2}\sum_{i=1}^{2}\sum_{t=2}^{T}\left(\log(2\pi) + \log(\sigma_i^2(1-\rho_i^2)) + \frac{(\varepsilon^V_{i,t} - \rho_i\varepsilon^V_{i,t-1})^2}{\sigma_i^2(1-\rho_i^2)}\right), \end{aligned} \tag{67}$$

where the first term allows for some effect of the initial observations, see (Kanniainen, Lin, and Yang 2014) and references therein for details. Here $\Theta^* = (\omega_1^*, \omega_2^*, \alpha_2^*, \alpha_2^*, \beta_1^*, \beta_2^*, \gamma_1^*, \gamma_2^*)$ denotes the risk-neutral parameters which are mapped from the physical parameters by Equation (29). For convenience, and since we do not need these parameters for option pricing, we omit the autoregressive parameters from Θ^* and replace the variance of the VIX errors by the sample variance $\sigma_i^2 \approx s_i^2 = Var\left[\varepsilon^V_{i,t}\right]$.

There are 15 parameters in our model: $\Theta = (\omega_1, \omega_2, \alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2, \lambda_1, \lambda_2, d_1, d_2, a, \eta_1, \eta_2)$. To estimate the parameters we maximize the total likelihood which is the sum of the returns likelihood and the volatility index likelihood. This is given by

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

and allows us to estimate all the model parameters with the available return and VIX data. To ensure that the global maximum of the likelihood function is attained, we use a large pool of initial values in the optimization algorithm and report results for the (global) optimum across this sample of starting values.

Electronic copy available at: https://ssrn.com/abstract=3953826

Table 1: Estimation results for the HN-GARCH model fitted to S&P 500 Index data

$\omega_1(\times$E-6)	$\alpha_1(\times$E-6)	β_1	$\gamma_1(\times$E+2)	λ_1	d_1	Like
2.1295	5.3265	0.6301	2.4148	1.6401	1.6141	32712.6
(0.100)	(0.212)	(0.009)	(0.070)	(1.116)	(0.041)	

This table shows the parameter estimates (and standard errors in parenthesis) for the univariate model fitted on the time series of daily S&P 500 Index (SPX) returns from 2000 through 2020 and the corresponding implied volatility values from June 2010 through 2020.

It should be noted that the structure of the model allows for sequential estimation of parameters. Thus, in a first step we estimate the parameters $(\omega_1, \alpha_1, \beta_1, \gamma_1, \lambda_1, d_1)$ in the univariate HN-GARCH model given by Equations (1), (3), (26) and (28), by fitting the model to S&P 500 returns and VIX data. This also ensures that the same estimated values are used for the first asset for all bivariate models. Table 1 shows the results of this first step. From the table we observe that all parameters are significant, except for the equity risk premium parameter λ_1. The significant estimate of d_1 indicates that the marginal pricing kernel for SPX is indeed non-linear as suggested by (Christoffersen, Heston, and Jacobs 2013).

In the second step, we keep the estimated parameters for SPX fixed, and estimate the rest of the parameters using both assets' data. Table 2 contains the estimation results for the second asset. The model with $\eta_t = 0$ corresponds to the bivariate ERS-GARCH model, and the one with $\eta_t \neq 0$ is the new model. We observe that the additional parameters η_1 and η_2 are all significantly different from zero, except for η_1 with DJX data. In line with this observation, the likelihood values are much higher for our new model demonstrating the improved fit to index return and volatility index data.

3.2 Marginal pricing kernels and univariate option pricing

We now consider the implications of our proposed model on the marginal pricing kernel of the second asset and the effect such differences has for the pricing of options on this asset. To illustrate the significance of our model, we compare it to two alternative and restricted models. The first model is the univariate HN-GARCH model with variance-dependent pricing kernel as in (Christoffersen, Heston, and Jacobs 2013). This is a particular case of our new model where both potential drivers of correlation, i.e. a and η_t, are zero. The second model is the bivariate ERS-GARCH model of (Escobar-Anel, Rastegari, and Stentoft 2020), which is a particular case of our new model with $\eta_t = 0$. Our new model, on the other hand, incorporates both a non-zero value for a and the additional parameters in η_t.

Let $M_{j,t} = \frac{f_j^{\mathbb{Q}}(S_{j,t})}{f_j^{\mathbb{P}}(S_{j,t})}$ be the $\mathcal{F}_{t-1}$-conditional marginal pricing kernel of the j-th asset

Electronic copy available at: https://ssrn.com/abstract=3953826

Table 2: Estimation results for the model fitted to individual stock returns and VIX

Model	ω_2	α_2 (×E-6)	β_2	γ_2 (×100)	λ_2	d_2	a	η_1 (×0.1)	η_2 (×0.1)	Like
					Panel A: DJX					
$\eta_t = 0$	0	0.6128	0.9153	0.0899	-8.0448	3.0833	0.9187	0	0	40120.6
	-	(0.039)	(0.005)	(0.170)	(4.692)	(0.184)	(0.003)	-	-	
$\eta_t \neq 0$	0	0.5509	0.9194	0.2664	-7.3321	2.5657	0.8690	0.4626	-0.6385	40142.2
	-	(0.036)	(0.005)	(0.182)	(4.703)	(0.278)	(0.007)	(0.317)	(0.101)	
					Panel B: NDX					
$\eta_t = 0$	0	2.9296	0.9143	1.0653	-0.8244	1.1910	1.0657	0	0	34896.6
	-	(0.202)	(0.005)	(0.069)	(1.457)	(0.078)	(0.005)	-	-	
$\eta_t \neq 0$	0	1.6764	0.8909	2.0309	0.5436	1.1962	0.8889	-7.5821	1.6827	35165.3
	-	(0.105)	(0.005)	(0.107)	(1.596)	(0.103)	(0.006)	(0.284)	(0.174)	
					Panel C: RUT					
$\eta_t = 0$	0	2.2905	0.9335	0.7235	-0.2734	1.9563	1.1168	0	0	34997.7
	-	(0.171)	(0.004)	(0.080)	(2.026)	(0.083)	(0.005)	-	-	
$\eta_t \neq 0$	0	2.0591	0.9242	1.0541	0.2184	1.5961	1.0000	-4.7739	1.4036	35057.9
	-	(0.154)	(0.005)	(0.089)	(2.029)	(0.113)	(0.009)	(0.339)	(0.210)	

This table shows the parameter estimates (and standard errors in parenthesis) for the second asset in the bivariate models fitted on the time series of daily stock returns from 2000 through 2020 and the corresponding implied volatility values from June 2010 through 2020 keeping the parameters for the first index fixed at their respective values from Table 1.

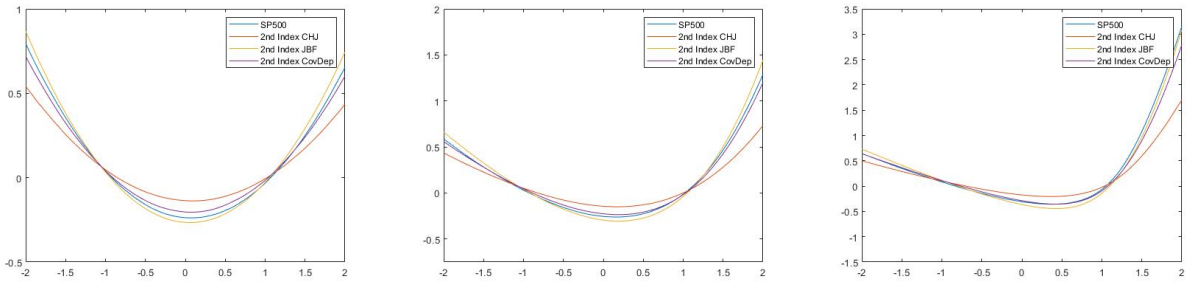
over one day with $j = 1, 2$. The physical and risk-neutral dynamics of the two assets in Equations (1)-(4) and (26)-(28) show that each asset price, $S_{j,t}$, has a log-normal $\mathcal{F}_{t-1}$-conditional marginal distribution under both measures. Let $S_{j,t} \overset{\mathbb{P}}{\sim} \log N(\mu_{j,t}, \sigma_{j,t}^2)$ and $S_{j,t} \overset{\mathbb{Q}}{\sim} \log N(\mu_{j,t}^*, {\sigma_{j,t}^*}^2)$ for $j = 1, 2$. As the log-normal density is given by $f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(\frac{(x-\mu)^2}{2\sigma^2}\right)$, the conditional 1-day log-kernel is obtained as

$$\begin{aligned} \log M_{j,t} &= \log(\sigma_{j,t}) - \log\left(\sigma_{j,t}^*\right) + \frac{1}{2}\left(\frac{\log S_{j,t} - \mu_{j,t}}{\sigma_{j,t}}\right)^2 - \frac{1}{2}\left(\frac{\log S_{j,t} - \mu_{j,t}^*}{\sigma_{j,t}^*}\right)^2 \\ &= \log(\sigma_{j,t}) - \log\left(\sigma_{j,t}^*\right) + \frac{1}{2\sigma_{j,t}^2 {\sigma_{j,t}^*}^2}\left({\sigma_{j,t}^*}^2(\log S_{j,t} - \mu_{j,t})^2 - \sigma_{j,t}^2(\log S_{j,t} - \mu_{j,t}^*)^2\right), \end{aligned} \tag{68}$$

which is a quadratic function of $\log S_{j,t}$ and hence has a (potentially inverted) U-shaped plot. The concavity or convexity of the marginal log-kernel is determined by the sign of $({\sigma_{j,t}^*}^2 - \sigma_{j,t}^2)/(2\sigma_{j,t}^2{\sigma_{j,t}^*}^2)$ which is the coefficient of the square term.

The marginal log-kernel of the first asset corresponds to $j = 1$ in Equation (68) with

Electronic copy available at: https://ssrn.com/abstract=3953826



(a) 1 day (b) 1 week (c) 1 month

Figure 1: Logarithm of marginal pricing kernel for models fitted to DJX

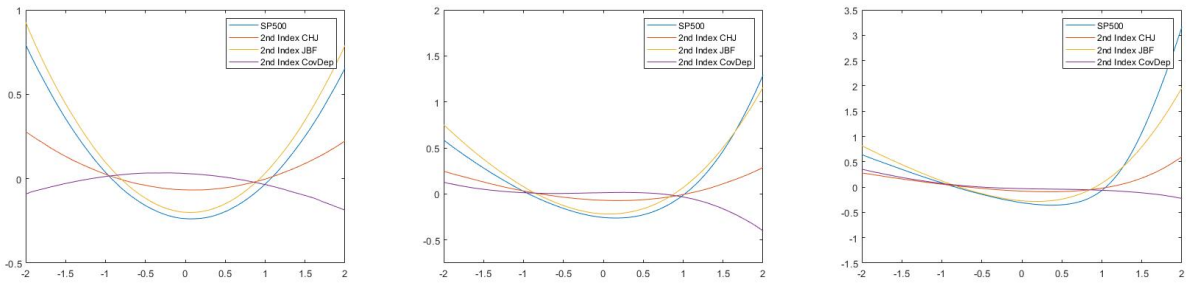
physical parameters $\mu_{1,t} = \log S_{1,t-1} + r + \lambda_1 h_{1,t}$ and $\sigma_{1,t}^2 = h_{1,t}$, and risk-neutral parameters $\mu_{1,t}^* = \log S_{1,t-1} + r - \frac{1}{2} d_1 h_{1,t}$ and ${\sigma_{1,t}^*}^2 = d_1 h_{1,t}$. The sign of the coefficient of the square term is determined by ${\sigma_{1,t}^*}^2 - \sigma_{1,t}^2 = (d_1 - 1) h_{1,t}$. Hence, as noted in (Christoffersen, Heston, and Jacobs 2013), when $d_1 > 1$ the log-kernel is indeed concave. When $d_1 = 1$ there is no variance premium and the log-kernel is linear as in the original HN-GARCH model.

As for the second asset, Equation (68) with $j = 2$ gives the marginal log-kernel. Using independence of innovation terms, we have $\mu_{2,t} = \log S_{2,t-1} + r + \lambda_{21} h_{1,t} + \lambda_2 h_{2,t}$ and ${\sigma_{2,t}}^2 = a^2 h_{1,t} + (1 + \eta_t^2) h_{2,t} - 2a\eta_t \sqrt{h_{1,t} h_{2,t}}$ under the physical measure, and $\mu_{2,t}^* = \log S_{2,t-1} + r - \frac{1}{2} a^2 d_1 h_{1,t} - \frac{1}{2} d_2 h_{2,t}$ and ${\sigma_{2,t}^*}^2 = a^2 d_1 h_{1,t} + d_2 h_{2,t}$ under the equivalent martingale measure. Here, the coefficient of the square term is a function of the parameters a, d_1, d_2, and η_t. In the case of independent assets where $\eta_t = a = 0$, we have ${\sigma_{2,t}^*}^2 - \sigma_{2,t}^2 = (d_2 - 1) h_{2,t}$ which corresponds to a univariate HN-GARCH model for the second asset with a variance-dependent kernel. The case $\eta_t = 0$ with $a \neq 0$ corresponds to the bivariate ESR-GARCH model and gives ${\sigma_{2,t}^*}^2 - \sigma_{2,t}^2 = a^2 (d_1 - 1) h_{1,t} + (d_2 - 1) h_{2,t}$. In this case, if $d_1, d_2 > 1$ the marginal log-kernel is again concave. In our model where $a, \eta_t \neq 0$ the presence of η_t may however result in an inverted U-shaped kernel even when $d_1, d_2 > 1$.

Figures 1, 2 and 3 show the logarithm of the marginal pricing kernel over periods of 1 day, 1 week and 1 month for the three choices of the second asset in our model, namely, DJX, NDX and RUT. Plots are based on our estimates of model parameters in Tables 1 and 2, and the conditional variance components $h_{j,t}$ are set to their unconditional values. Each plot shows three curves corresponding to the three different models. As a benchmark the figures also plot the logarithm of the marginal pricing kernel for SPX.

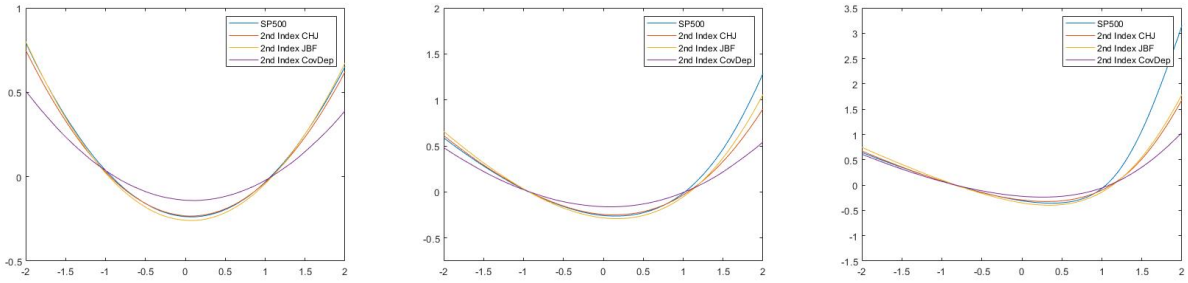
In all figures, the univariate HN-GARCH pricing kernel is U-shaped similar to those in (Christoffersen, Heston, and Jacobs 2013). The non-linear kernel is a result of a variance-dependent kernel which corresponds to $d_2 \neq 1$. The marginal log kernel implied by the ERS-

Electronic copy available at: https://ssrn.com/abstract=3953826



(a) 1 day (b) 1 week (c) 1 month

Figure 2: Logarithm of marginal pricing kernel for models fitted to NDX



(a) 1 day (b) 1 week (c) 1 month

Figure 3: Logarithm of marginal pricing kernel for models fitted to RUT

Electronic copy available at: https://ssrn.com/abstract=3953826

GARCH model is also non-linear, but different from the univariate model due to presence of correlation under both the physical and risk-neutral models. For DJX and RUT the kernels are quite close, in particular for shorter maturities, and the results for DJX are very close to that of SPX.

In all the figures, however, we can see the effect of the additional parameters η_1 and η_2 on the pricing kernel when comparing our new model to the restricted model with $\eta_1 = \eta_2 = 0$. Figure 3, for example, shows that the marginal log kernel in our model is clearly different from the restricted model in particular towards the tail of the distribution for RUT. When fitted to NDX data, the change is dramatic as seen in Figure 2. This is consistent with the relative magnitude of η_1 and η_2 which are larger for NDX estimates compared to those for RUT. In the case of DJX data, η_1 is not significantly different from zero, and η_2 is relatively small. Hence, in Figure 1 we see much smaller differences between marginal log kernels of the models.

We now examine the importance of the differences in pricing kernel for (univariate) option pricing. To price options we need the marginal moment generating functions, $m_1^{\mathbb{Q}}(u_1)$ and $m_2^{\mathbb{Q}}(u_2)$, which are obtained by setting $u_2 = 0$ and $u_1 = 0$, respectively, in the joint mgf given by Equation (32). With this, the value of a European call option on the j'th asset with strike price K and time to maturity T, can be calculated in closed form as

$$C = \frac{S_{j,t} - Ke^{-r(T-t)}}{2} + \frac{e^{-r(T-t)}}{\pi} \int_0^\infty \text{Re}\left[\frac{K^{-iu_j}(m_j^{\mathbb{Q}}(iu_j + 1) - Km_j{}^{\mathbb{Q}}(iu_j))}{iu_j}\right] du_j, \quad (69)$$

where Re[.] denotes the real part of a complex number. The price of a European put option can be obtained from the put-call parity.

Table 3 shows the option pricing performance of our model ($\eta_t \neq 0$) and the restricted model ($\eta_t = 0$) using three metrics: 1) implied VIX likelihood from the estimation, 2) call options implied volatility root mean squared error (IVRMSE), and 3) put options IVRMSE. We first note that the results for the VIX likelihood in Panel A complement the total likelihoods reported in Table 2 and demonstrate the significantly improved fit to the data of the model with covariance dependent pricing kernel. As it was the case with the total likelihood, our new model has much higher VIX likelihood values for all three bivariate pairs of index data.[2] The largest improvements in VIX likelihood are found for NDX, which is also where the estimated values for new parameters are the largest.

Panels B and C of Table 3 show that our model also outperforms the restricted model when it comes to pricing actual options and this holds for both call and put options. In

[2]It also follows that our model improves by a large amount on the return likelihood for all three assets.

Electronic copy available at: https://ssrn.com/abstract=3953826

Table 3: Pricing results

Panel A: VIX Likelihood

Model	DJX	NDX	RUT
$\eta = 0$	16001.0	15745.7	15618.6
$\eta \neq 0$	16019.9	15775.1	15634.3

Panel B: Call Option IVRMSE

Model	DJX	NDX	RUT
$\eta = 0$	0.1094	0.1031	0.1267
$\eta \neq 0$	0.0926	0.0664	0.0929

Panel C: Put Option IVRMSE

Model	DJX	NDX	RUT
$\eta = 0$	0.0877	0.0848	0.1078
$\eta \neq 0$	0.0737	0.0523	0.0771

This table shows the empirical performance of our proposed model using three metrics. The VIX likelihood in Panel A correspond to those implied by the corresponding parameter estimates. The call and put option IVRMSEs in Panels B and C, respectively, correspond to the root mean squared error of the implied Black Scholes volatilities from actual option prices and our model implied option prices.

particular, the new model we propose has at least 15% smaller errors, for call options on DJX the relative error is $(0.092 - 0.1094)/0.1094 = -15.36\%$, and could lead to improvements of up to 38% in option pricing for NDX. Put differently, the pricing error incurred when using the restricted model with $\eta_t = 0$ are $(0.1031 - 0.0664)/0.0664 = 55\%$ and $(0.0848 - 0.0523)/0.0523 = 62\%$ larger than with our model for call and put options, respectively, for NDX. Consistent with the new parameter estimates being smallest for DJX the difference in pricing errors are also smallest for this index. The differences though are still economically significant and larger than 18% for both types of options demonstrating the important improvement our proposed model offers on the ERS-GARCH model.

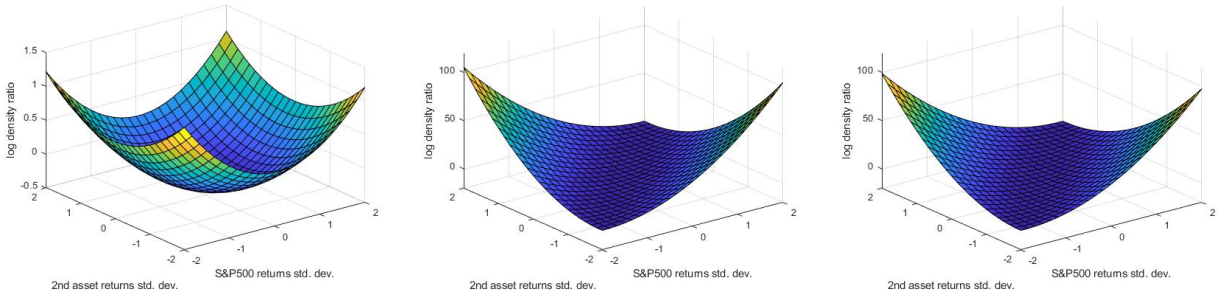
3.3 Joint pricing kernels and the impact on two-asset option prices

We now consider the joint pricing kernel in our bivariate model. The joint log-normal density of a bivariate random variable Y is given as

$$f_Y(Y) = \frac{1}{2\pi y_1 y_2 \sqrt{\det \Sigma}} \exp\left(-\frac{1}{2}\left(\log Y - \mu\right)' \Sigma^{-1} \left(\log Y - \mu\right)\right). \tag{70}$$

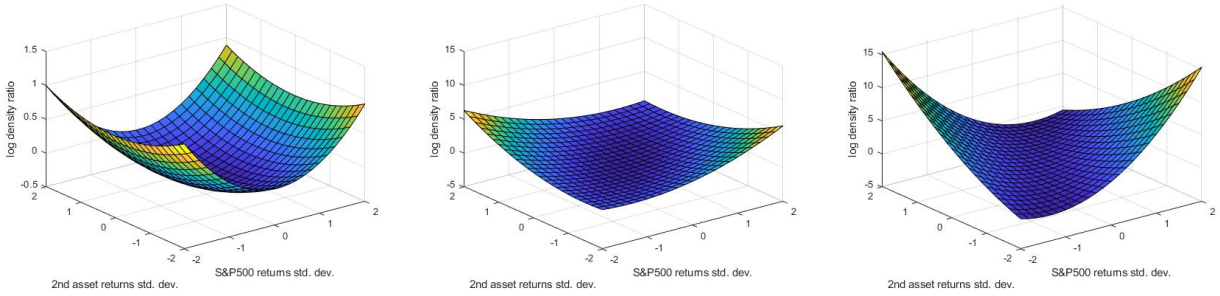
Using the matrix form of our model in Equation (5), we have the conditional joint distribution of log-price as $\log S_t \overset{\mathbb{P}}{\sim} N(\mu_t, \Sigma_t)$ with $\mu_t = \log S_{t-1} + r\mathbf{1} + \Lambda \operatorname{diag}(h_t)\mathbf{1}$ and $\Sigma_t = A_t A_t'$.

Electronic copy available at: https://ssrn.com/abstract=3953826



(a) Independent Heston-Nandi (b) ERS-GARCH (c) Covariance-dependent

Figure 4: Logarithm of 1-day joint pricing kernel for models fitted to DJX data



(a) Independent Heston-Nandi (b) ERS-GARCH (c) Covariance-dependent

Figure 5: Logarithm of 1-day joint pricing kernel for models fitted to NDX data

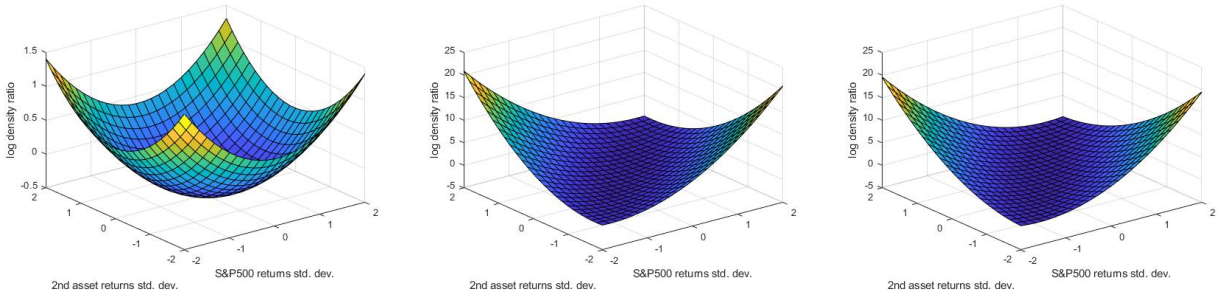
Under the equivalent martingale measure Equation (31) gives $\log S_t \overset{\mathbb{Q}}{\sim} N(\mu_t^*, \Sigma_t^*)$ with $\mu_t^* = \log S_{t-1} + r\mathbf{1} + \Lambda^* \operatorname{diag}(h_t^*)\mathbf{1}$ and $\Sigma_t^* = A_t^* A_t^{*\prime}$. Therefore, the conditional 1-day log-kernel is obtained as

$$\log M_t = -\frac{1}{2}\log(d_1 d_2) - \frac{1}{2}\left(\log S_t - \mu_t^*\right)'\left(\Sigma_t^*\right)^{-1}\left(\log S_t - \mu_t^*\right) - \frac{1}{2}\left(\log S_t - \mu_t\right)'\left(\Sigma_t\right)^{-1}\left(\log S_t - \mu_t\right). \tag{71}$$

In Figures 4, 5 and 6 we plot the 1-day bivariate kernel for the three models under consideration for the three bivariate pairs of indices, respectively. These figures show a clear difference between the case without correlation and with correlation. Moreover, while the differences in the 1-day bivariate kernels with correlation appear small for RUT and DJX, Figure 5 shows that these may be large when the estimated values for η_1 and η_2 are large as is the case for NDX.

To illustrate the effect of correlation risk premia on option prices, we now consider the pricing of correlation options, which our model allows for in closed form. To be specific, the

Electronic copy available at: https://ssrn.com/abstract=3953826



(a) Independent Heston-Nandi (b) ERS-GARCH (c) Covariance-dependent

Figure 6: Logarithm of 1-day joint pricing kernel for models fitted to RUT data

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

$$C = e^{-r(T-t)} E_t^{\mathbb{Q}}[\max(e^{x_{1,T}} - K_1, 0) \max(e^{x_{2,T}} - K_2, 0)] \tag{72}$$

$$= e^{-r(T-t)} \left(\int_{\log K_1}^{+\infty} \int_{\log K_2}^{+\infty} \max((e^{x_{1,T}} - K_1)\,(e^{x_{2,T}} - K_2)\,, 0) f^{\mathbb{Q}}(x_{1,T}, x_{2,T}) dx_{1,T} dx_{2,T} \right), \tag{73}$$

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

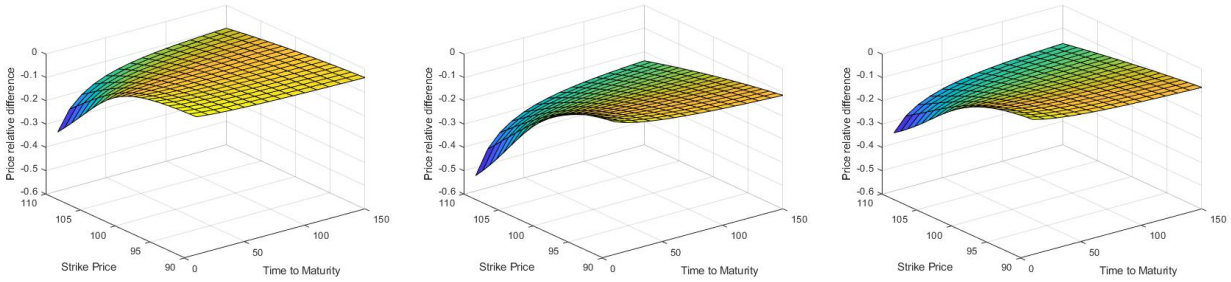
$$C = \frac{e^{-\alpha_1 \log K_1 - \alpha_2 \log K_2}}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-i(u_1 \log K_1 + u_2 \log K_2)} \Psi(u_1, u_2) du_1 du_2, \text{where}$$

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

Since the joint mgf is available in our model and given by Equation (32), closed form option pricing is feasible.

In Figure 7 we plot the relative differences in the price of correlation call options when considering the correlation-dependent pricing kernel proposed here and the restricted version of this corresponding to the ERS-GARCH model. From this figure we observe that the model with a variance-dependent kernel overprices correlation options compared to the model with correlation-dependent kernel, especially for deep out-of-the-money options with very short maturity. The difference in pricing is more pronounced when using NDX estimates which can

Electronic copy available at: https://ssrn.com/abstract=3953826



(a) DJX (b) NDX (c) RUT

Figure 7: Relative differences in prices for two-asset correlation options

This figure plots the differences in prices for two-asset correlation options priced with a variance dependent pricing kernel and with a covariance dependent kernel relative to the former, i.e. divided by the price obtained with a variance dependent pricing kernel.

be attributed to relatively large magnitudes of correlation premium parameters η_1 and η_2. The pricing difference is smaller in the case of DJX, which is consistent with small estimated values for η_1 and η_2. Across all three cases, though, the differences are large and range from 35% to 53%.

4 Conclusions

This paper introduces a class of multivariate GARCH models which admits both a variance and a correlation dependent kernel. In particular, the pricing kernel we propose is driven by, at least, three free parameters accounting for shifts from the historical ($\mathbb{P}$) measure to the risk-neutral ($\mathbb{Q}$) measure in the conditional covariance matrix. The model is an extension of the ERS-GARCH model of (Escobar-Anel, Rastegari, and Stentoft 2020) and extends the existing literature by explicitly modelling correlation dependent pricing kernels.

Although the general model is potentially non-affine under the historical measure, we establish simple conditions under which affine dynamics under the risk-neutral measure are ensured. This captures a key contribution of our paper; proposing a model which is not necessarily affine under the historical measure, and thus more flexible and allowing for a better fit to empirical data, but can be made affine under the risk-neutral measure, and therefore continues to allow for efficient pricing of derivatives.

We perform a full calibration of our proposed model to three bivariate series fitting both index returns and their corresponding volatility indexes in a joint maxmimum likelihood estimation. The superiority of our model is capture by improvements in both the overall

Electronic copy available at: https://ssrn.com/abstract=3953826

likelihood and the VIX-implied likelihoods, as well a better fitting of one dimensional (or marginal) option prices. The new degree freedom in the covariance-dependent kernel is also shown to significantly change prices of two-asset correlation options.

References

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 (2019): "Closed-form Variance Swap Prices under General Affine GARCH Models and Their Continuous-time Limits," *Annals of Operations Research*, 282, 27–57.

Badescu, A., R. J. Elliott, and J.-P. Ortega (2015): "Non-Gaussian GARCH option pricing models and their diffusion limits," *European Journal of Operational Research*, 247(3), 820–830.

Bakshi, G., and D. Madan (2000): "Spanning and Derivative-Security Valuation," *Journal of Financial Economics*, 55(2), 205–238.

Bégin, J.-F., C. Dorion, and G. Gauthier (2020): "Idiosyncratic Jump Risk Matters: Evidence from Equity Returns and Options," *The Review of Financial Studies*, 33(1), 155–211.

Bollerslev, T. (1986): "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31(3), 307–327.

Christoffersen, P., S. Heston, and K. Jacobs (2006): "Option Valuation with Conditional Skewness," *Journal of Econometrics*, 131(1-2), 253–284.

——— (2013): "Capturing Option Anomalies with a Variance-Dependent Pricing Kernel," *The Review of Financial Studies*, 26(8), 1962–2006.

Date, P., and S. Islyaev (2015): "A fast calibrating volatility model for option pricing," *European Journal of Operational Research*, 243(2), 599–606.

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.

Electronic copy available at: https://ssrn.com/abstract=3953826

Engle, R. F. (1982): "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, pp. 987–1007.

Escobar-Anel, M., J. Rastegari, and L. Stentoft (2020): "Affine multivariate GARCH models," *Journal of Banking & Finance*, 118, 105895.

——— (2021): "Option pricing with conditional GARCH models," *European Journal of Operational Research*, 289(1), 350–363.

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., 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.

Menn, C., and S. T. Rachev (2005): "A GARCH option pricing model with α-stable innovations," *European Journal of Operational Research*, 163(1), 201–209.

Ornthanalai, C. (2014): "Levy Jump Risk: Evidence from Options and Returns," *Journal of Financial Economics*, 112(1), 69–90.

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.

Electronic copy available at: https://ssrn.com/abstract=3953826

A Proofs of main results

Proof of Proposition 2.1. We first note that the Radon-Nikodym derivative is a positive function.

Next, using the definition of C_{t-1} we have

$$\begin{aligned}
E_{t-1}^{\mathbb{P}}\left[\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t}\right] &= E_{t-1}^{\mathbb{P}}\left[\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_{t-1}} \exp\left(\theta_{1,t}z_{1,t}+\theta_{2,t}z_{2,t}+\xi_{3,t}z_{1,t}z_{2,t}+\xi_{1,t}h_{1,t+1}+\xi_{2,t}h_{2,t+1}-C_{t-1}\right)\right] \\
&= \left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_{t-1}} \exp\left(-C_{t-1}\right) E_{t-1}^{\mathbb{P}}\left[\exp\left(\theta_{1,t}z_{1,t}+\theta_{2,t}z_{2,t}+\xi_{3,t}z_{1,t}z_{2,t}+\xi_{1,t}h_{1,t+1}+\xi_{2,t}h_{2,t+1}\right)\right] \\
&= \left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_{t-1}} \exp\left(-C_{t-1}\right)\exp\left(C_{t-1}\right) \\
&= \left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_{t-1}}.
\end{aligned}$$

This shows the Radon-Nikodym derivative is a $\mathbb{P}$-martingale. Moreover, by continuing recursively we obtain $E_0^{\mathbb{P}}\left[\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t}\right] = 1$. Therefore, Equation (14) defines a valid Radon-Nikodym derivative.

Next we show that discounted price of each asset is a $\mathbb{Q}$-martingale. To establish this, it is sufficient to show that $E_{t-1}^{\mathbb{Q}}\left[\frac{S_{i,t}}{S_{i,t-1}}\right] = e^r$ for the two assets. The first asset gives

$$\begin{aligned}
E_{t-1}^{\mathbb{Q}}\left[\frac{S_{1,t}}{S_{1,t-1}}\right] =& E_{t-1}^{\mathbb{P}}\left[\frac{S_{1,t}}{S_{1,t-1}}\frac{d\mathbb{Q}/d\mathbb{P}\big|\,\mathcal{F}_t}{d\mathbb{Q}/d\mathbb{P}\big|\,\mathcal{F}_{t-1}}\right] \\
=& \exp\left(r+\lambda_1 h_{1,t}-C_{t-1}(\theta_{1,t},\theta_{2,t},\xi_{1,t},\xi_{2,t},\xi_{3,t})\right) \\
& \times E_{t-1}^{\mathbb{P}}\left[\exp\left(\left(\sqrt{h_{1,t}}+\theta_{1,t}\right)z_{1,t}+\theta_{2,t}z_{2,t}+\xi_{3,t}z_{1,t}z_{2,t}+\xi_{1,t}h_{1,t+1}+\xi_{2,t}h_{2,t+1}\right)\right] \\
=& \exp\left(r+\lambda_1 h_{1,t}-C_{t-1}(\theta_{1,t},\theta_{2,t},\xi_{1,t},\xi_{2,t},\xi_{3,t})\right)\exp\left(C_{t-1}(\theta_{1,t}+\sqrt{h_{1,t}},\theta_{2,t},\xi_{1,t},\xi_{2,t},\xi_{3,t})\right) \\
=& e^r,
\end{aligned}$$

where the last step follows from Equation (12).

Electronic copy available at: https://ssrn.com/abstract=3953826

Finally, the risk-neutral expected return of the second asset is calculated as

$$
\begin{aligned}
E_{t-1}^{\mathbb{Q}}\left[\frac{S_{2,t}}{S_{2,t-1}}\right] &= E_{t-1}^{\mathbb{P}}\left[\frac{S_{2,t}}{S_{2,t-1}}\frac{d\mathbb{Q}/d\mathbb{P}\big|\,\mathcal{F}_t}{d\mathbb{Q}/d\mathbb{P}\big|\,\mathcal{F}_{t-1}}\right] \\
&= E_{t-1}^{\mathbb{P}}\Big[\exp\left(r+\lambda_{21}h_{1,t}+\lambda_2 h_{2,t}+\left(a\sqrt{h_{1,t}}-\eta_t\sqrt{h_{2,t}}\right)z_{1,t}+\sqrt{h_{2,t}}z_{2,t}\right) \\
&\times \exp\left(\theta_{1,t}z_{1,t}+\theta_{2,t}z_{2,t}+\xi_{3,t}z_{1,t}z_{2,t}+\xi_{1,t}h_{1,t+1}+\xi_{2,t}h_{2,t+1}-C_{t-1}(\theta_{1,t},\theta_{2,t},\xi_{1,t},\xi_{2,t},\xi_{3,t})\right)\Big] \\
&= \exp\left(r+\lambda_{21}h_{1,t}+\lambda_2 h_{2,t}-C_{t-1}(\theta_{1,t},\theta_{2,t},\xi_{1,t},\xi_{2,t},\xi_{3,t})\right) \\
&\times E_{t-1}^{\mathbb{P}}\Big[\exp\Big(\left(a\sqrt{h_{1,t}}-\eta_t\sqrt{h_{2,t}}\right)z_{1,t}+\sqrt{h_{2,t}}z_{2,t}+\theta_{1,t}z_{1,t}+\theta_{2,t}z_{2,t}+\xi_{3,t}z_{1,t}z_{2,t}+\xi_{1,t}h_{1,t+1}+\xi_{2,t}h_{2,t+1} \\
&= \exp\left(r+\lambda_{21}h_{1,t}+\lambda_2 h_{2,t}-C_{t-1}(\theta_{1,t},\theta_{2,t},\xi_{1,t},\xi_{2,t},\xi_{3,t})\right) \\
&\times \exp\left(C_{t-1}\left(\theta_{1,t}+a\sqrt{h_{1,t}}-\eta_t\sqrt{h_{2,t}},\theta_{2,t}+\sqrt{h_{2,t}},\xi_{1,t},\xi_{2,t},\xi_{3,t}\right)\right) \\
&= e^r,
\end{aligned}
$$

where the last step follows from Equation (13). This completes the proof.

□

Proof of Proposition 2.2. First observe the conditional joint cumulant generating function of $(z_{1,t}, z_{2,t})$ is given by

$$
\begin{aligned}
\log E_{t-1}^{\mathbb{Q}}\left[e^{uz_{1,t}+vz_{2,t}}\right] &= \log E_{t-1}^{\mathbb{P}}\left[e^{uz_{1,t}+vz_{2,t}}\frac{d\mathbb{Q}/d\mathbb{P}\big|\,\mathcal{F}_t}{d\mathbb{Q}/d\mathbb{P}\big|\,\mathcal{F}_{t-1}}\right] \\
&= C_{t-1}(u+\theta_{1,t},v+\theta_{2,t},\xi_{1,t},\xi_{2,t},\xi_{3,t})-C_{t-1}(\theta_{1,t},\theta_{2,t},\xi_{1,t},\xi_{2,t},\xi_{3,t}) \qquad (74)\\
&= \frac{1}{2K_t}\left[\left(2(1-2\xi_{2,t}\alpha_2)(\theta_{1,t}+B_{1,t})+2\left(\xi_{3,t}-2\xi_{2,t}\alpha_2\eta_t\right)(\theta_{2,t}+B_{2,t})\right)u\right. \qquad (75)\\
&+\left(2(1-2\xi_{1,t}\alpha_1-2\xi_{2,t}\alpha_2\eta_t^2)(\theta_{2,t}+B_{2,t})+2\left(\xi_{3,t}-2\xi_{2,t}\alpha_2\eta_t\right)(\theta_{1,t}+B_{1,t})\right)v \\
&\left.+(1-2\xi_{2,t}\alpha_2)u^2+(1-2\xi_{1,t}\alpha_1-2\xi_{2,t}\alpha_2\eta_t^2)v^2+2\left(\xi_{3,t}-2\xi_{2,t}\alpha_2\eta_t\right)uv\right], \qquad (76)
\end{aligned}
$$

where in the last step we invoked C_{t-1} from Equation (16). Comparing to the bivariate normal cgf given by

$$
\log E\left[e^{ux+vy}\right] = \mu_x u+\mu_y v+\frac{1}{2}\sigma_x^2u^2+\frac{1}{2}\sigma_y^2v^2+\sigma_{xy}uv, \qquad (77)
$$

Electronic copy available at: https://ssrn.com/abstract=3953826

we see that $(z_{1,t}, z_{2,t})'|\mathcal{F}_{t-1} \overset{\mathbb{Q}}{\sim} N(\mu_t, \Sigma_t)$ with mean

$$\mu_t = \frac{1}{K_t}\begin{bmatrix} (1-2\xi_{2,t}\alpha_2)(\theta_{1,t}+B_{1,t}) + (\xi_{3,t}-2\xi_{2,t}\alpha_2\eta_t)(\theta_{2,t}+B_{2,t}) \\ (1-2\xi_{1,t}\alpha_1-2\xi_{2,t}\alpha_2\eta_t^2)(\theta_{2,t}+B_{2,t}) + (\xi_{3,t}-2\xi_{2,t}\alpha_2\eta_t)(\theta_{1,t}+B_{1,t}) \end{bmatrix}, \tag{78}$$

and covariance matrix

$$\Sigma_t = \frac{1}{K_t}\begin{bmatrix} 1-2\xi_{2,t}\alpha_2 & \xi_{3,t}-2\xi_{2,t}\alpha_2\eta_t \\ \xi_{3,t}-2\xi_{2,t}\alpha_2\eta_t & 1-2\xi_{1,t}\alpha_1-2\xi_{2,t}\alpha_2\eta_t^2 \end{bmatrix}. \tag{79}$$

To simplify the notation, assume the following mapping between parameters $(\xi_{1,t}, \xi_{2,t}, \xi_{3,t}, \theta_{1,t}, \theta_{2,t})$ and $(d_{1,t}, d_{2,t}, d_{3,t}, b_{1,t}, b_{2,t})$

$$d_{1,t} = \frac{1-2\xi_{2,t}\alpha_2}{K_t}, \tag{80}$$

$$d_{2,t} = \frac{1-2\xi_{1,t}\alpha_1-2\xi_{2,t}\alpha_2\eta_t^2}{K_t} - \eta_t^2 d_{1,t}, \tag{81}$$

$$d_{3,t} = \frac{\xi_{3,t}-\eta_t}{K_t}, \tag{82}$$

$$b_{1,t} = \frac{1}{K_t\sqrt{h_{1,t}}}\left[(1-2\xi_{2,t}\alpha_2)(\theta_{1,t}+B_{1,t}) + (\xi_{3,t}-2\xi_{2,t}\alpha_2\eta_t)(\theta_{2,t}+B_{2,t})\right], \tag{83}$$

$$b_{2,t} = \frac{1}{K_t\sqrt{h_{2,t}}}\left[(1-2\xi_{1,t}\alpha_1-2\xi_{2,t}\alpha_2\eta_t^2)(\theta_{2,t}+B_{2,t}) + (\xi_{3,t}-2\xi_{2,t}\alpha_2\eta_t)(\theta_{1,t}+B_{1,t})\right] - \eta_t\sqrt{\frac{h_{1,t}}{h_{2,t}}}b_{1,t}. \tag{84}$$

Now, Equations (78) and (79) lead to $(z_{1,t}, z_{2,t})'|\mathcal{F}_{t-1} \overset{\mathbb{Q}}{\sim} N(\mu_t, \Sigma_t)$ with

$$\mu_t = \begin{bmatrix} b_{1,t}\sqrt{h_{1,t}} \\ b_{1,t}\eta_t\sqrt{h_{1,t}} + b_{2,t}\sqrt{h_{2,t}} \end{bmatrix}, \quad \Sigma_t = \begin{bmatrix} d_{1,t} & d_{3,t}+d_{1,t}\eta_t \\ d_{3,t}+d_{1,t}\eta_t & \eta_t^2 d_{1,t} + d_{2,t} \end{bmatrix}. \tag{85}$$

Therefore, the bivariate random variable $(z_{1,t}, z_{2,t})'$ can be represented as

$$\begin{bmatrix} z_{1,t} \\ z_{2,t} \end{bmatrix} = \begin{bmatrix} \sqrt{d_{1,t}} & 0 \\ \sqrt{d_{1,t}}\left(\frac{d_{3,t}}{d_{1,t}}+\eta_t\right) & \sqrt{d_{2,t}}\sqrt{1-\frac{d_{3,t}}{d_{2,t}}\left(\frac{d_{3,t}}{d_{1,t}}+2\eta_t\right)} \end{bmatrix}\begin{bmatrix} z_{1,t}^* \\ z_{2,t}^* \end{bmatrix} + \begin{bmatrix} b_{1,t}\sqrt{h_{1,t}} \\ b_{1,t}\eta_t\sqrt{h_{1,t}} + b_{2,t}\sqrt{h_{2,t}} \end{bmatrix}, \tag{86}$$

where $z_{1,t}^*$ and $z_{2,t}^*$ are independent standard normal random variables. The above equations give Equations (20) and (21). □

Electronic copy available at: https://ssrn.com/abstract=3953826

Proof of Corollary 2.3. Setting $d_{3,t} = 0$ ($\xi_{3,t} = \eta_t$) in Equation (19) gives $K_t = (1 - 2\xi_{2,t}\alpha_2)(1 - 2\xi_{1,t}\alpha_1 - \eta_t^2)$. Hence, Equations (80)-(81) reduce to

$$d_1 = \frac{1}{1 - 2\xi_{1,t}\alpha_1 - \eta_t^2}, \tag{87}$$

$$d_2 = \frac{1}{1 - 2\xi_{2,t}\alpha_2}. \tag{88}$$

This also implies $K_t = \frac{1}{d_1 d_2}$. Next, Equations (83)-(84) are simplified to

$$b_1 = \frac{d_1}{\sqrt{h_{1,t}}} \left[(\theta_{1,t} + B_{1,t}) + \eta_t (\theta_{2,t} + B_{2,t}) \right], \tag{89}$$

$$b_2 = \frac{d_2}{\sqrt{h_{2,t}}} (\theta_{2,t} + B_{2,t}). \tag{90}$$

Notice that given parameters d_1, d_2, b_1, b_2 the above equations can be solved for $\xi_{1,t}, \xi_{2,t}, \theta_{1,t}, \theta_{2,t}$. Hence, having d_1, d_2, b_1, b_2 constant is a plausible assumption.

Finally, applying $d_{3,t} = 0$ ($\xi_{3,t} = \eta_t$) to Equations (22)-(25) give the risk-neutral process as

$$R_{1,t} = r + (\lambda_1 + b_1)\, h_{1,t} + \sqrt{d_1 h_{1,t}} z_{1,t}^*, \tag{91}$$

$$R_{2,t} = r + (\lambda_{21} + a b_1)\, h_{1,t} + (\lambda_2 + b_2)\, h_{2,t} + \left(a\sqrt{d_1 h_{1,t}} \right) z_{1,t}^* + \sqrt{d_2 h_{2,t}} z_{2,t}^*, \tag{92}$$

$$h_{1,t} = \omega_1 + \beta_1 h_{1,t-1} + d_1 \alpha_1 \left(z_{1,t-1}^* - \left(\frac{\gamma_1 - b_1}{d_1} \right) \sqrt{d_1 h_{1,t-1}} \right)^2, \tag{93}$$

$$h_{2,t} = \omega_2 + \beta_2 h_{2,t-1} + d_2 \alpha_2 \left(z_{2,t-1}^* - \left(\frac{\gamma_2 - b_2}{d_2} \right) \sqrt{d_2 h_{2,t-1}} \right)^2. \tag{94}$$

Note that the above process is indeed risk-neutral because of relations in Equation (30). These relations combined with mappings (87) - (90) guarantee that Equations (12) and (13) hold. Alternatively, one can directly verify that given relations in Equation (30) the discounted asset prices in Equations (91) and (92) are martingales.

□

Electronic copy available at: https://ssrn.com/abstract=3953826