Article

Mean-Reverting 4/2 Principal Components Model. Financial Applications

Marcos Escobar-Anel *,† [ID] **and Zhenxian Gong** †

Department of Statistical and Actuarial Sciences, University of Western Ontario, London, ON N6A5B7, Canada; zgong26@uwo.ca
* Correspondence: marcos.escobar@uwo.ca
† These authors contributed equally to this work.

Abstract: In this paper, we propose a new multivariate mean-reverting model incorporating state-of-the art 4/2 stochastic volatility and a convenient principal component stochastic volatility (PCSV) decomposition for the stochastic covariance. We find a quasi closed-form characteristic function and propose analytic approximations, which aid in the pricing of derivatives and calculation of risk measures. Parameters are estimated on three bivariate series, using a two-stage methodology involving method of moments and least squares. Moreover, a scaling factor is added for extra degrees of freedom to match data features. As an application, we consider investment strategies for a portfolio with two risky assets and a risk-free cash account. We calculate value-at-risk (VaR) values at a 95% risk level using both simulation-based and distribution-based methods. A comparison of these VaR values supports the effectiveness of our approximations and the potential for higher dimensions.

Keywords: principal component analysis; 4/2 stochastic volatility model; moment-generating function; risk management calculations

1. Introduction

In mathematical finance, principal component analysis (PCA) is used to reduce dimensionality in the explanation of a vector of asset returns; see, for instance, Alexander (2001) for discrete-time model applications. The methodology has also been used in continuous-time stochastic processes for financial applications; see Escobar et al. (2010) and Escobar and Olivares (2013) for its usage in collateralized debt obligations (CDO) and exotic financial derivatives pricing, as well as Escobar and Gschnaidtner (2018) and more generally De Col et al. (2013) for factor and PC analyses applications to foreign exchange data.

When modeling financial or any complex data, one can focus on capturing the stylized facts reported in the literature. The best-known features of financial data are as follows: fat tails, changing volatilities and correlations, the leverage effect and co-volatility movements. Examples of a refined fact of financial data are the smiles and smirks of the implied volatility surface. To capture them, in Christoffersen et al. (2009), the authors proposed a PCA-inspired stochastic covariance (SC) model using the popular Heston stochastic volatility model Heston (1993) as the underlying component.

These features can be captured by rich SC models, with proper marginal structures. SC models have received significant attention in the literature; the best-known representatives are the stochastic Wishart family, see Da Fonseca et al. (2007); Gouriéroux (2006), and the Ornstein–Uhlenbeck (OU) family, see Muhle-Karbe et al. (2012), of models, as well as general linear-quadratic jump-diffusions, see Cheng and Scaillet (2007). Even though these models show realistic advantages over the classical Black–Scholes model, they lose their tractability as a result of an increase in model dimensions (i.e., an increase in the number of parameters and simulation paths). This is commonly known as the curse of dimensionality, and PCA is a viable method to control the problem with dimensionality. Inspired by this, principal component stochastic volatility (PCSV) models are built from

a linear combination of tractable one-dimensional counterparts. Their applications have been studied in a series of papers since 2010; see, for example, De Col et al. (2013); Escobar (2018); Escobar et al. (2010); Escobar-Anel and Moreno-Franco (2019).

Current PCSV models rely on Heston SV for the components, also known as the 1/2 model. A new model for single assets, namely, the 4/2 volatility process, was masterfully presented in Grasselli (2016). Notably, the Heston model (the 1/2 model) predicts that the implied volatility skew will flatten when the instantaneous volatility increases (e.g., financial crises), while another embedded structure, the 3/2 model see Platen (1997) predicts steeper skews. The author argues that the two processes complement each other in better explaining the implied volatility surface. There are many interesting, recent generalizations of the 4/2 model, see, for example, Cui et al. (2021) and Kirkby and Nguyen (2020) and the literature therein.

The works presented above mostly target the equity market and are built based on geometric Brownian motion (GBM)-type processes, and they are hence not suitable for commodities and volatility indexes. These asset classes display mean-reverting and spillover effects, both of which are stylized facts not seen in equities. Mean-reverting effects capture the stationary behavior of prices, which tend to go back to a long-term mean. On the other hand, spillover refers to the impact of one asset on the trends (drift) of other assets (i.e., the impact of one asset on the long term average "stationary price" of a second asset). Our modeling in this paper will ensure that these two facts are captured.

Our modeling is inspired by a recent paper by Cheng et al. (2019) that introduces a generalized multivariate mean-reverting 4/2 factor analysis (FA) model. The model uses the one-dimensional mean-reverting 4/2 stochastic volatility proposed by Escobar-Anel and Gong (2020) as the underlying model. They obtained an analytical representation of the characteristic function (c.f.) of a vector of asset prices as well as a second conditional c.f. for non-mean-reverting nested cases. Thus, the FFT-based option pricing method, for example Carr and Madan (1999), can be used, and exact simulation is possible. The authors further identified a set of conditions that not only produces well-defined changes of measure, but also avoids local martingales for risk-neutral pricing purposes.

In this paper, we make several contributions to the literature:

- We studied in detail a multivariate mean-reverting 4/2 stochastic volatility model based on PCA, which is inspired in the general framework of Cheng et al. (2019). The SC in the new model is decomposed into constant eigenvectors that capture the correlation among assets and a diagonal eigenvector matrix whose entries are modeled by the 4/2 process.
- The PCA structure allows us to find a semi-closed-form c.f. for the vector of returns. It permits the extension to multidimensions of simple but accurate approximation approaches, first introduced in Escobar-Anel and Gong (2020) for one dimension, to find closed-form approximations to the c.f., which are proven to be accurate for realistic parameter settings.
- We use the estimation approaches developed in Escobar-Anel and Gong (2020) to estimate the parameters for special cases of the proposed model. Here, we use two pairs of bivariate time series capturing both the asset and its variance. Estimation of multidimensional processes is rare in the literature, and our work demonstrates that many, but not all, of the parameters are statistically significant, confirming stylized facts of commodity prices and volatility indexes such as stochastic correlation and spill-over effects.
- A risk management application, based on a constant proportion strategies portfolio, for example Merton (1975) and DeMiguel et al. (2009), demonstrates the accuracy of the approximation.

The rest of the paper is organized as follows: in Section 2, we define the model and derive two sub-models for two parametric constructions. Then, in Section 3, we expand the theoretical results for the c.f.s obtained in Cheng et al. (2019) with approximations. In Section 4, we focus on estimation for the multivariate mean-reverting 4/2 stochastic

volatility model. The estimation method considered is based on the method developed in
Escobar-Anel and Gong (2020) with the introduction of a new scaling factor. Thereafter, in
Section 5, we construct a portfolio with two risky assets and one risk-free cash account, and
we subsequently calculate the value-at-risk (VaR) at a 95% level using various techniques.
Finally, Section 6 concludes the paper.

2. Model Definition

In this section, we define the general model. We first introduce a model with spillover
effects, and we later cover models with separable spillover effects and no spillover effects
as special cases. We also study the implications of the model on the covariance process.

2.1. General Model Setup

Suppose $\mathbf{X_t} = (X_1(t), \ldots, X_n(t))'$ is a vector of assets. The dynamics for each asset
$X_i(t)$ under the so-called historical measure $\mathbb{P}$ is defined as

$$
\begin{cases}
\frac{dX_i(t)}{X_i(t)} &= \left[L_i + \sum_{j=1}^{n} c_{ij}\left(\sqrt{\nu_j(t)} + \frac{b_j}{\sqrt{\nu_j(t)}}\right)^2 - \sum_{j=1}^{n} \beta_{ij} ln(X_j(t)) \right] dt + \sum_{j=1}^{n} a_{ij}\left(\sqrt{\nu_j(t)} + \frac{b_j}{\sqrt{\nu_j(t)}}\right) dW_j^{\mathbb{P}}(t), \\
d\nu_j(t) &= \alpha_j(\theta_j - \nu_j(t))dt + \xi_j\sqrt{\nu_j(t)}dB_j^{\mathbb{P}}(t), \quad j = 1, \ldots, n
\end{cases}
\tag{1}
$$

where $W_i^{\mathbb{P}}(t)$ and $B_j^{\mathbb{P}}(t)$ are independent Brownian motions if $i \neq j$, and they are correlated
if $i = j$, that is, the quadratic variation $\langle W_i^{\mathbb{P}}(t), B_i^{\mathbb{P}}(t) \rangle = \rho_i t$, where ρ_i is constant. The
parameters for each $\nu_j(t)$ process are positive and satisfy the Feller condition: $\alpha_j\theta_j > \frac{\xi_j^2}{2}$.
Moreover, we assume that the mean-reverting level of $\nu_j(t)$ decreases as j increases; that is,
$0 < \theta_j < \theta_{j-1}$, for $j = 2, \ldots, n$. This last feature is intended to sort the eigenvalues in order
of importance, and b_j's measure the "weight" of the 3/2 components.

This model is unlike a traditional mean-reverting model, as it takes into account the
spillover effects in the drift, which appear in the form of $\beta_{ij}, i \neq j$. We mentioned above
that the spillover effects show the impacts of one asset on others; these impacts shall not
be confused with correlations. The correlations are reflected in the price trend of both
assets, capturing co-movements between assets. Spillover effects describe the impact on
the mean-reverting level of one asset by others (i.e., the shift in the long term mean due to
the movements of other assets). The concept of spillover effects can be understood as how
much, for example, a demand curve of one good shifts according to the change in factors
of other goods. In addition, although this paper does not dwell into risk-neutral pricing,
we should highlight that changes of measure are feasible on each principal component
along the lines of Proposition 4 in Escobar-Anel and Gong (2020), this is changes of the
type: $dW_j^{\mathbb{P}} = \lambda_j\left(\sqrt{\nu_j(t)} + \frac{b_j}{\sqrt{\nu_j(t)}}\right)dt + dW_j^{\mathbb{Q}}, dB_j^{\mathbb{P}} = \lambda_{n\nu_j}\sqrt{\nu_j}dt + dB_j^{\mathbb{Q}}$.

Equation (1) can be written in matrix form as follows:

$$
\mathbf{dX_t} = diag(\mathbf{X_t})\left[(\mathbf{L} + \mathbf{CV_t} - \mathbf{B} \cdot \mathbf{ln(X_t)})dt + \mathbf{\Sigma_t^{\frac{1}{2}}} \mathbf{dW_t^{\mathbb{P}}} \right].
\tag{2}
$$

where $\mathbf{W_t} = (W_1^{\mathbb{P}}(t), \ldots, W_n^{\mathbb{P}}(t))'$ is a vector of independent standard Brownian motions;

$$
\mathbf{B} = \begin{pmatrix} \beta_{11} & \cdots & \beta_{1n} \\ \vdots & \ddots & \vdots \\ \beta_{n1} & \cdots & \beta_{nn} \end{pmatrix}, \mathbf{L} = (L_1, \ldots, L_n)', \mathbf{C} = \begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \cdots & c_{nn} \end{pmatrix}, \mathbf{V_t} = \left(\left(\sqrt{\nu_1(t)} + \frac{b_1}{\sqrt{\nu_1(t)}}\right)^2, \ldots, \left(\sqrt{\nu_n(t)} + \frac{b_n}{\sqrt{\nu_n(t)}}\right)^2 \right)'
$$

and $\mathbf{ln(X_t)} = (ln(X_1(t)), \ldots, ln(X_n(t)))'$.

We first assume that the eigenvalues of the matrix $-\mathbf{B} = (-\beta_{ij})$ are all negative; this
is similar to the literature, see Langetieg (1980) and Larsen (2010). This assumption will
be used to explain some of the estimation results. Here, $\mathbf{B}$ captures the spillover effects,
while $\mathbf{C}$ contains risk premiums associated with the assets, and the long-term average for
the assets is determined by $\mathbb{E}\left[\mathbf{B}^{-1}(\mathbf{L} + \mathbf{CV_t}) \right]$.

We next assume a principal component decomposition on the instantaneous covariance matrix Σ_t: $\Sigma_t dt = \mathbf{A} diag(\mathbf{V_t})\mathbf{A}'dt$, where $\mathbf{A} = (a_{ij})_{n \times n}$ is an orthogonal matrix with constant entries, and it captures the correlations among assets. We craft the matrix $\mathbf{C}$ in such a way that allows for c.f. analytical approximations; that is $\mathbf{C} = \mathbf{A}\tilde{\mathbf{C}} + \frac{1}{2}(\mathbf{A} \circ \mathbf{A})$, where $\tilde{\mathbf{C}} = diag(c_1, \ldots, c_n)$ and $(\mathbf{A} \circ \mathbf{A})$ denotes the Hadamard product of $\mathbf{A}$. The dynamics of log price $Y_i(t) = ln(X_i(t))$ is as follows:

$$\begin{cases} dY_i(t) & = \left[L_i + \sum\limits_{j=1}^{n} a_{ij}c_j\left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}}\right)^2 - \sum\limits_{j=1}^{n} \beta_{ij}Y_j(t) \right]dt + \sum_{j=1}^{n} a_{ij}\left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}}\right)dW_j^{\mathbb{P}}(t), \\ dv_j(t) & = \alpha_j(\theta_j - v_j(t))dt + \xi_j\sqrt{v_j(t)}dB_j^{\mathbb{P}}(t), \; j = 1, \ldots, n. \end{cases} \quad (3)$$

Based on the applications, our model can be reduced to three subcases for which we are able to approximate the c.f. with analytic functions:

- $b_j = 0$: This is a generalization of Escobar et al. (2010) to multivariate mean-reverting asset classes. If $n = 1$, we get the model considered in Benth (2011).
- $\rho_j = 0$: This case applies to the assets whose price series demonstrates an abnormal increase or decrease, but no leverage effect is observed for the assets of interest. The term "leverage effect" was first defined and studied in Black (1976). It describes the negative correlation between an asset's volatility and its return.
- $b_j = 0, \rho_j = 0$: This case can be generated by either of the two previous cases. It applies better to assets that exhibit mild behavior in their price series; at the same time, no leverage effect is identified.

We explain how to approximate the c.f. with analytic functions for these three cases in Section 3.

2.1.1. Separable Spillover Effect

In this section, we assume a convenient structure in the spillover matrix $\mathbf{B}$ to obtain the c.f., and by doing so, we obtain another solvable case. Here, we further simplify the model by rewriting it in terms of n independent processes. We demonstrate this procedure by first writing Equation (3) in matrix form:

$$\mathbf{dY_t} = (\mathbf{L} + \mathbf{A}\tilde{\mathbf{C}}\mathbf{V_t} - \mathbf{B}\mathbf{Y_t})dt + \mathbf{A} diag(\mathbf{V_t})^{\frac{1}{2}}\mathbf{dW_t^{\mathbb{P}}}. \quad (4)$$

Multiplying both sides of Equation (4) by $\mathbf{A}^{-1}$, we get

$$\mathbf{A}^{-1}\mathbf{dY_t} = (\mathbf{A}^{-1}\mathbf{L} + \mathbf{A}^{-1}\mathbf{A}\tilde{\mathbf{C}}\mathbf{V_t} - \mathbf{A}^{-1}\mathbf{B}\mathbf{Y_t})dt + diag(\mathbf{V_t})^{\frac{1}{2}}\mathbf{dW_t^{\mathbb{P}}}. \quad (5)$$

Suppose the matrix $\mathbf{B}$ can be written as follows: $\mathbf{B} = \mathbf{A}\tilde{\mathbf{B}}\mathbf{A}^{-1}$, where $\tilde{\mathbf{B}} = diag(\tilde{\beta}_1, \ldots, \tilde{\beta}_n)$ is a diagonal matrix (i.e., whose entries are eigenvalues of $\mathbf{B}$). Using this result, and applying a simple transformation $\mathbf{M_t} = \mathbf{A}^{-1}\mathbf{Y_t}$, we arrive at a new mean-reverting process with diagonal matrix $\tilde{\mathbf{B}}$:

$$\mathbf{dM_t} = (\mathbf{A}^{-1}\mathbf{L} + \tilde{\mathbf{C}}\mathbf{V_t} - \tilde{\mathbf{B}}\mathbf{M_t})dt + diag(\mathbf{V_t})^{\frac{1}{2}}\mathbf{dW_t^{\mathbb{P}}}.$$

Each element of $\mathbf{dM_t}$ is a mean-reverting 4/2 stochastic volatility process, as in Escobar-Anel and Gong (2020). That is,

$$\begin{cases} dM_j(t) & = \left[\tilde{L}_j + \tilde{c}_j\left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}}\right)^2 - \tilde{\beta}_j M_j(t) \right]dt + \left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}}\right)dW_j^{\mathbb{P}}(t) \\ dv_j(t) & = \alpha_j(\theta_j - v_j(t))dt + \xi_j\sqrt{v_j(t)}dB_j^{\mathbb{P}}(t) \\ \langle dW_j^{\mathbb{P}}(t), dB_j^{\mathbb{P}}(t) \rangle & = \rho_j dt \end{cases} \quad (6)$$

where $\tilde{L}_j = \sum\limits_{i=1}^{n} L_i a_{ij}^*$, and a_{ij}^* are the entries of $\mathbf{A}^{-1}$. Furthermore, $\mathbf{dM_t}$ is also a vector of independent processes.

2.1.2. Model with No Spillover Effects

In this section, we assume no spillover effects among the assets (i.e., matrix **B** is diagonal). This further simplifies our model to

$$\begin{cases} \frac{dX_i(t)}{X_i(t)} &= \left[L_i + \sum_{j=1}^{n} c_{ij}\left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}}\right)^2 - \beta_i ln(X_i(t)) \right]dt + \sum_{j=1}^{n} a_{ij}\left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}}\right)dW_j^{\mathbb{P}}(t), \\ dv_j(t) &= \alpha_j(\theta_j - v_j(t))dt + \xi_j\sqrt{v_j(t)}dB_j^{\mathbb{P}}(t), \; j = 1,\ldots,n \end{cases} \quad (7)$$

The corresponding matrix representation has the same form:

$$\mathbf{dX_t} = diag(\mathbf{X_t})\left[(\mathbf{L} + \mathbf{CV_t} - \mathbf{B}\cdot\mathbf{ln(X_t)})dt + \Sigma_t^{\frac{1}{2}}\mathbf{dW_t^{\mathbb{P}}}\right], \quad (8)$$

with $\mathbf{B} = diag(\beta_1,\ldots,\beta_n)$. The dynamics of log price $Y_i(t) = ln(X_i(t))$ are then

$$\begin{cases} dY_i(t) &= \left[L_i + \sum_{j=1}^{n} a_{ij}c_j\left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}}\right)^2 - \beta_i Y_i(t) \right]dt + \sum_{j=1}^{n} a_{ij}\left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}}\right)dW_j^{\mathbb{P}}(t), \\ dv_j(t) &= \alpha_j(\theta_j - v_j(t))dt + \xi_j\sqrt{v_j(t)}dB_j^{\mathbb{P}}(t), \; j = 1,\ldots,n. \end{cases} \quad (9)$$

2.2. Properties of the Variance Vector

We devote this subsection to exploring the properties of the variance vector. This is important for understanding the instantaneous volatilities implied by our model. Recall that empirical data is related to these volatilities; therefore, one should ensure that these implied processes reflect the stylized facts of the data they cater to.

Let $\boldsymbol{\sigma}_t^2$ denote the variance vector; by definition, this is $\boldsymbol{\sigma}_t^2 = diag(\boldsymbol{\Sigma_t}) = (\mathbf{A} \circ \mathbf{A})\mathbf{V_t}$. As defined before, $\mathbf{V_t}$ is a vector of 4/2 processes (the sum of 1/2 and 3/2 processes). Therefore, $\boldsymbol{\sigma}_t^2$ can be written in terms of a linear combination of these two processes:

$$\boldsymbol{\sigma}_t^2 = \begin{pmatrix} \sum\limits_{j=1}^{n} a_{1j}^2\left(v_j(t) + \frac{b_j^2}{v_j(t)} + 2b_j\right) \\ \vdots \\ \sum\limits_{j=1}^{n} a_{nj}^2\left(v_j(t) + \frac{b_j^2}{v_j(t)} + 2b_j\right) \end{pmatrix} = \begin{pmatrix} \sum\limits_{j=1}^{n} a_{1j}^2 v_j(t) \\ \vdots \\ \sum\limits_{j=1}^{n} a_{nj}^2 v_j(t) \end{pmatrix} + \begin{pmatrix} \sum\limits_{j=1}^{n} a_{1j}^2\left(\frac{b_j^2}{v_j(t)}\right) \\ \vdots \\ \sum\limits_{j=1}^{n} a_{nj}^2\left(\frac{b_j^2}{v_j(t)}\right) \end{pmatrix} + \begin{pmatrix} \sum\limits_{j=1}^{n} a_{1j}^2 2b_j \\ \vdots \\ \sum\limits_{j=1}^{n} a_{nj}^2 2b_j \end{pmatrix}$$

This model for the variance can be interpreted as factor model with n 4/2 factors. Due to the popularity of factor models for explaining asset classes, it stands to reason that volatility indexes (these variances) can also be expressed in terms of factors, which could reflect intrinsic and systemic economical movements.

One can obtain the dynamics of $\boldsymbol{\sigma}_t^2$ more explicitly:

$$\mathbf{d\boldsymbol{\sigma}_t^2} = \begin{pmatrix} \sum\limits_{j=1}^{n} a_{1j}^2\left[\left(\alpha_j(\theta_j - v_j(t)) + \frac{b_j^2}{v_j^2(t)}(\xi_j^2 - \alpha_j\theta_j + \alpha_j v_j(t))\right)dt + \xi_j(\sqrt{v_j(t)} - b_j^2 v_j^{\frac{3}{2}}(t))dB_j^{\mathbb{P}}(t)\right] \\ \vdots \\ \sum\limits_{j=1}^{n} a_{nj}^2\left[\left(\alpha_j(\theta_j - v_j(t)) + \frac{b_j^2}{v_j^2(t)}(\xi_j^2 - \alpha_j\theta_j + \alpha_j v_j(t))\right)dt + \xi_j(\sqrt{v_j(t)} - b_j^2 v_j^{\frac{3}{2}}(t))dB_j^{\mathbb{P}}(t)\right] \end{pmatrix}$$

From the above stochastic differential Equation (SDE), we are able to obtain the variance and covariance of the vector $\mathbf{d\boldsymbol{\sigma}_t^2}$ via quadratic variations. Note that these can be interpreted as the volatility of variance and the correlation among variances (co-volatility movement), respectively:

$$\langle d\sigma_i^2(t), d\sigma_i^2(t) \rangle = \sum_{j=1}^{n} a_{ij}^4 \left[\xi_j\left(\sqrt{v_j(t)} - b_j^2 v_j^{\frac{3}{2}}(t)\right)\right]^2 dt \quad (10)$$

$$\langle d\sigma_i^2(t), d\sigma_j^2(t) \rangle = \sum_{k=1}^{n} a_{ik}^2 a_{jk}^2 \left[\xi_j \left(\sqrt{v_j(t)} - b_j^2 v_j^{\frac{3}{2}}(t) \right) \right]^2 dt \tag{11}$$

Equations (10) and (11) suggest that the instantaneous variance and covariance of $\boldsymbol{\sigma}_t^2$ are locally stochastic (i.e., driven by the same Brownian as the underlying).

3. Characteristic Functions and Approximations

In this section, we derive the c.f.s for the previously presented cases involving spillover effects and no spillover effects, using Proposition 2 and Corollary 1 from Cheng et al. (2019), in line with the approximation approach from Escobar-Anel and Gong (2020). In Escobar-Anel and Gong (2020), the authors obtained analytical approximations of the c.f.'s for the special cases $\rho = 0, b \neq 0$; $b = 0, \rho \neq 0$ using results from Grasselli (2016). Taking advantage of the principal component structure of the model, we demonstrate that the c.f. representations boil down to a multiplication of one-dimensional approximations.

Next, we show the c.f. for the general model and its submodels described in Section 2, namely the model with general spillover effects (Section 3.1) and the model with separable spillover effects (Section 3.2). Then, in Section 3.3, we present the principle used to approximate the c.f.'s.

3.1. Characteristic Function for Model with Spillover Effects

Let us first define $\mathbf{Z_t} = e^{\mathbf{B}t} \mathbf{Y_t}$ such that $e^{\mathbf{B}t}$ is a matrix exponential; then $Z_i(t)$ is represented as

$$dZ_i(t) = \sum_{j=1}^{n} \left(e^{\mathbf{B}t} \right)_{ij} \left\{ L_j + \sum_{k=1}^{n} a_{ij} c_j \left(\sqrt{v_k(t)} + \frac{b_k}{\sqrt{v_k(t)}} \right)^2 \right\} dt + \sum_{j=1}^{n} \left(e^{\mathbf{B}t} \right)_{ij} \left\{ \sum_{k=1}^{n} a_{jk} \left(\sqrt{v_k(t)} + \frac{b_k}{\sqrt{v_k(t)}} \right) dW_k(t) \right\} \tag{12}$$

For convenience, we use $\left(e^{\mathbf{B}t} \right)_{ij}$ as the ij-th component of the matrix $e^{\mathbf{B}t}$. Note that $Z_i(t)$ is no longer a mean-reverting process, although it accounts for time-dependent coefficients.

Corollary 1. *Let* $(\mathbf{Z(t)})_{t \geq 0}$ *evolve according to the model in Equation* (12). *The c.f.* $\Phi_{\mathbf{Z}(t), \mathbf{v}(t)}$ *is then given as follows:*

$$\Phi_{\mathbf{Z(t)}, \boldsymbol{\cdot}_t}(T, \omega) = E \left[\exp i\omega'(\mathbf{Z}(T) - \mathbf{Z}(t)) \mid \mathbf{Z}(t) = z_t, v(t) = \mathbf{v_t} \right]$$

$$= \prod_{k=1}^{n} \Phi_{GG}(T, 1; L(\omega), h_k(\omega), g_k(\omega), \alpha_k, \theta_k, \xi_k, \rho_k, b_k, c_k, v_{k,t}, Z(t))$$

where $h_k(\omega, t) = \sum_{j=1}^{n} a_{jk} c_k f_j(\omega, t)$, $L(\omega, t) = \sum_{j=1}^{n} \frac{L_j}{n} f_j(\omega, t)$, $g_k(\omega) = \sum_{j=1}^{n} a_{jk} f_j(\omega, t)$ *and* $f_j(\omega, t) = \sum_{m=1}^{n} i\omega_m \left(e^{\beta t} \right)_{mj}$. *Moreover,* Φ_{GG} *is a one-dimensional generalization of the c.f. from Grasselli (2016) provided in Lemma A1 of Cheng et al. (2019).*

The proof follows as a direct application of the proof of Proposition 2 in Cheng et al. (2019).

3.2. Characteristic Function for Model with Separable Spillover Effects

To derive the c.f. of $M_i(t)$, we perform the transformation $S_j(t) = e^{\tilde{\beta}_j t} M_j(t)$, recognizing that the c.f. of $S_j(t)$ has been derived in Escobar-Anel and Gong (2020). Hence, the c.f. of $\mathbf{Y_t}$ is a product of the corresponding c.f. of $S_j(t)$. The result is summarized in the following corollary.

Corollary 2. *Let* $\Phi_{MR}\left(T, u; L, c, a, \beta, \alpha, \theta, \xi, \rho, b, v(t), Z_{(t)}\right)$ *denote the characteristic function provided in Proposition 2.1 in* Escobar-Anel and Gong (2020)*; then, the characteristic function of* $\mathbf{Y_t}$ *is given by the following equation:*

$$\mathbb{E}(e^{i\mathbf{u}'\mathbf{Y_T}}|\mathcal{F}_t) = \prod_{j=1}^{n} \Phi_{MR}\left(T, \mathbf{u}^*; \tilde{L}_j, \tilde{c}_j, 1, \tilde{\beta}_j, \alpha_j, \theta_j, \xi_j, \rho_j, b_j, v_j(t), S_j(t)\right) \tag{13}$$

where $\mathbf{u}^*$ *is a new vector of real numbers with element* $u_j^* = \sum\limits_{i=1} u_i a_{ij} e^{-\tilde{\beta}_j t}$.

The proof is straightforward; using the relationship $\mathbf{Y_t} = \mathbf{AM_t}$, we know that each individual process $Y_i(t)$ is a linear combination of $M_j(t)$, $j = 1, 2, \dots, n$, and therefore $S_j(t)$, processes. The product $\mathbf{u}'\mathbf{Y_T}$ can be further written in terms of $S_j(t)$:

$$\mathbf{u}'\mathbf{Y_T} = \mathbf{u}'\mathbf{AM_T} = \sum_{j=1}^{n}\sum_{i=1}^{n} u_i a_{ij} M_j(T) = \sum_{j=1}^{n}\sum_{i=1}^{n} u_i a_{ij} e^{-\tilde{\beta}_j t} S_j(T) = \sum_{j=1}^{n} u_j^* S_j(T).$$

The independence property of random variables leads to Equation (13).

3.3. Approximation Principle and Results

We have learned that the c.f. can be written in terms of a product of the c.f.'s of n independent one-dimensional processes thanks to principal component decomposition. These one-dimensional processes are only different in the structure of the matrix exponential term (i.e., $e^{\mathbf{B}t}$), which is deterministic, and they resemble the same $Z(t)$ process seen in Escobar-Anel and Gong (2020). Therefore, the principles to approximate $\Phi_{Z(t), \cdot_t}(T, \omega)$ follow those adopted in Escobar-Anel and Gong (2020). In other words, we only need to calculate an approximation to the individual c.f. Φ_{GG}, and the approximation can be realized under three scenarios, as described in Section 2: $b_j = 0$; $\rho_j = 0$ and the trivial case of $b_j = 0, \rho_j = 0$.

For completeness, the formula for Φ_{GG} in Cheng et al. (2019) is as follows:

$$\Phi_{GG}(T, u; L, h, g, \kappa, \theta, \xi, \rho, b, c, v_t, Z_t) = \exp\left\{iu\int_t^T A(s)ds\right\} v(t)^{iu\frac{b\rho}{\xi}g(t)} \exp\left\{-iu\rho\frac{g(t)v(t)}{\xi}\right\}$$

$$\times \mathbb{E}\left[v(T)^{iu\frac{b\rho}{\xi}g(T)} \exp\left\{iu\left(\int_t^T B(s)v(s)ds + \int_t^T C(s)\frac{1}{v(s)}ds + \int_t^T D(s)\ln(v(s))ds\right.\right.\right.$$

$$\left.\left.\left.+\rho\frac{g(T)v(T)}{\xi}\right)\right\}\,\Big|\,\mathcal{F}_t\right].$$

Φ_{GG} cannot be solved in closed-form due to the lack of a representation of the moment generating function of an integrated Cox-Ingersoll-Ross (CIR) process with time-dependent integrands. Therefore, we propose an analytic function that approximates the unsolvable conditional expectation:

$$\mathbb{E}\left[v(T)^{\frac{b\rho}{\xi}g(T)} \exp\left\{\int_t^T B(s)v(s)ds + \int_t^T C(s)\frac{1}{v(s)}ds + \int_t^T D(s)\ln(v(s))ds + \rho\frac{g(T)v(T)}{\xi}\right\}\,\Big|\,\mathcal{F}_t\right]$$

$$\approx \mathbb{E}\left[v(T)^{\frac{b\rho}{\xi}g(T)} \exp\left\{\rho\frac{g(T)v(T)}{\xi} - n\int_t^T v(s)ds - m\int_t^T \frac{1}{v(s)}ds\right\}\,\Big|\,\mathcal{F}_t\right],$$

for some complex constants m and n: $m \approx -\int_t^T C(s)ds$, $n \approx -\int_t^T B(s)ds$, and $D(s) = 0$, for $s \in [t, T]$. We propose the following two approximations:

- **Midpoint:** $m = \dfrac{\min\limits_{[t,T]}(C(s)) + \max\limits_{[t,T]}(C(s))}{2}$, $n = \dfrac{\min\limits_{[t,T]}(B(s)) + \max\limits_{[t,T]}(B(s))}{2}$.
- **Average:** $m = \dfrac{1}{T-t}\int_t^T C(s)ds$, $n = \dfrac{1}{T-t}\int_t^T B(s)ds$.

The approximated conditional expectation is solvable, as it fits the framework of Grasselli (2016). We summarize the results in the following corollary for the general model with spillover effects, which includes separable spillover effects as a special case.

Corollary 3. *Given deterministic functions $B_j(s)$ and $C_j(s)$, defined in Lemma A1 in Cheng et al. (2019), and $g_j(s)$, defined in Corollary 1 for $s \in [t, T]$,*

$$\mathbb{E}\left[v_j(T)^{\frac{b_j \rho_j}{\xi_j} g_j(T)} \exp\left\{ \int_t^T B_j(s) v_j(s) ds + \int_t^T C_j(s) \frac{1}{v_j(s)} ds + \int_t^T D_j(s) \ln(v_j(s)) ds \right. \right.$$
$$\left. \left. + \rho_j \frac{g_j(T) v_j(T)}{\xi_j} \right\} \mid \mathcal{F}_t \right]$$

can be approximated by analytic functions for constants m_j and n_j satisfying

$$n_j = \begin{cases} \dfrac{\min\limits_{[t,T]}(B_j(s)) + \max\limits_{[t,T]}(B_j(s))}{2}, & \text{if midpoint approach is considered} \\[2mm] \dfrac{1}{T-t} \int_t^T B_j(s) ds, & \text{if average approach is considered} \end{cases}$$

$$m_j = \begin{cases} \dfrac{\min\limits_{[t,T]}(C_j(s)) + \max\limits_{[t,T]}(C_j(s))}{2}, & \text{if midpoint approach is considered} \\[2mm] \dfrac{1}{T-t} \int_t^T C_j(s) ds, & \text{if average approach is considered} \end{cases}$$

under three scenarios:

- $\mathbf{b_j = 0}$: *Given $b_j = 0$, $C(s) = 0$ and $D(s) = 0$, $s \in [t, T]$. If $n_j \geq -\frac{\alpha_j^2}{2\xi_j^2}$, then*

$$\mathbb{E}\left[\exp\left\{ \int_t^T B_j(s) v_j(s) ds + \rho_j \frac{g_j(T) v_j(T)}{\xi_j} \right\} \mid \mathcal{F}_t \right] \approx \mathbb{E}\left[\exp\left\{ \rho_j \frac{g_j(T) v_j(T)}{\xi_j} - n_j \int_t^T v_j(s) ds \right\} \mid \mathcal{F}_t \right]$$

$$= \left(\frac{(B_j \xi_j^2 + \alpha)(e^{\sqrt{A_j}(T-t)} - 1) + \sqrt{A_j}(e^{\sqrt{A_j}(T-t)} + 1)}{2\sqrt{A_j} e^{\frac{\sqrt{A_j} + \alpha_j}{2}(T-t)}} \right)^{-\frac{2\alpha_j \theta_j}{\xi_j^2}} e^{v_j(t) \left(\frac{(B_j \alpha_j - 2n_j)(e^{\sqrt{A_j}(T-t)} - 1) - B_j \sqrt{A_j}(e^{\sqrt{A_j}(T-t)} + 1)}{(B_j \xi_j^2 + \alpha_j)(e^{\sqrt{A_j}(T-t)} - 1) + \sqrt{A_j}(e^{\sqrt{A_j}(T-t)} + 1)} \right)}, \quad (14)$$

$$B_j = -\frac{\rho_j g_j(T)}{\xi_j}, A_j = \alpha_j^2 + 2n_j \xi_j^2,$$

- $\mathbf{b_j \neq 0, æ_j = 0}$: *Given $b_j = 0$, $\rho_j \neq 0$ and $D(s) = 0$, $s \in [t, T]$. If $m_j > -\frac{(2\alpha_j \theta_j - \xi_j^2)^2}{8\xi_j^2}$, $n_j \geq -\frac{\alpha_j^2}{2\xi_j^2}$, then*

$$\mathbb{E}\left[\exp\left\{ \int_t^T B_j(s) v_j(s) ds + \int_t^T C_j(s) \frac{1}{v_j(s)} ds \right\} \mid \mathcal{F}_t \right] \approx \mathbb{E}\left[\exp\left\{ -n_j \int_t^T v_j(s) ds - m_j \int_t^T \frac{1}{v_j(s)} ds \right\} \mid \mathcal{F}_t \right]$$

$$= \left(\frac{\gamma_j(T, v_j(t))}{2} \right)^{k_j+1} v_j(t)^{-\frac{\alpha_j \theta_j}{\xi_j^2}} K_j(T)^{-\left(\frac{1}{2} + \frac{k_j}{2} + \frac{\alpha_j \theta_j}{\xi_j^2}\right)} e^{\frac{1}{\xi_j^2}\left(\theta_j(T-t) - \sqrt{H_j} v_j(t) \coth\left(\frac{\sqrt{H_j}(T-t)}{2}\right) + \alpha_j v_j(t)\right)} \frac{\Gamma\left(\frac{1}{2} + \frac{k_j}{2} + \frac{\alpha_j \theta_j}{\xi_j^2}\right)}{\Gamma(k_j + 1)}$$

$$\times {}_1F_1\left(\frac{1}{2} + \frac{k_j}{2} + \frac{\alpha_j \theta_j}{\xi_j^2}, k_j + 1, \frac{\gamma_j(T, v_j(t))^2}{4K_j(T)} \right), \quad (15)$$

$$k_j = \frac{1}{\xi_j^2} \sqrt{(2\alpha_j \theta_j - \xi_j^2)^2 + 8m_j \xi_j^2}, H_j = \alpha_j^2 + 2n_j \xi_j^2, \gamma_j(T, v_j(t)) = \frac{2\sqrt{H_j v_j(t)}}{\xi_j^2 \sinh\left(\frac{\sqrt{H_j}(T-t)}{2}\right)},$$

$$K_j(T) = \frac{1}{\xi_j^2} \left(\sqrt{H_j} v_j(t) \coth\left(\frac{\sqrt{H_j}(T-t)}{2}\right) + \alpha_j \right).$$

Corollary 3 follows directly from Propositions 2.2 and 2.3 in Escobar-Anel and Gong (2020). The approximation for the c.f. when there are no spillover effects follows the same procedure as presented in Corollary 3. For the case when the spillover effects are separable, we obtain a sum of independent mean-reverting 4/2 stochastic volatility processes, as indicated in Equation (6).

As a result, Propositions 2.2 and 2.3 in Escobar-Anel and Gong (2020) can be directly applied to approximate the c.f.s for these processes.

4. Estimation

In this section, we consider the model with separable spillover effects as the underlying model for estimation. In this way, on the one hand, we fulfill the purpose of studying spillover effects among assets, and on the other hand, we avoid the complexity of a matrix exponential[1]. Recall that the model with separable spillover effects can be expressed as follows:

$$\mathbf{dY_t} = (\mathbf{L} + \mathbf{A}\tilde{\mathbf{C}}\mathbf{V_t} - \mathbf{BY_t})dt + \mathbf{A}diag(\mathbf{V_t})^{\frac{1}{2}}\mathbf{dW_t^{\mathbb{P}}}.$$

where $\mathbf{B}$ is constructed such that it can be decomposed into a product of three matrices: $\mathbf{B} = \mathbf{A}\tilde{\mathbf{B}}\mathbf{A}^{-1}$. In this case, $dY_i(t)$ is a linear combination of independent processes $dM_j(t)$, that is, $d\mathbf{Y_t} = \mathbf{A}d\mathbf{M_t}$ with

$$\begin{cases} dM_j(t) & = \left[\tilde{L}_j + \tilde{c}_j \left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}} \right)^2 - \tilde{\beta}_j M_j(t) \right] dt + \left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}} \right) dW_j^{\mathbb{P}}(t) \\ dv_j(t) & = \alpha_j(\theta_j - v_j(t))dt + \xi_j \sqrt{v_j(t)} dB_j^{\mathbb{P}}(t) \\ \langle dW_j^{\mathbb{P}}(t), dB_j^{\mathbb{P}}(t) \rangle & = \rho_j dt \end{cases}$$

For simplicity, we focus on two dimensions, hence studying pairs of assets with their respective volatility indexes. For example, VIX (VVIX) and VSTOXX (VVSTOXX), USO (OVX) and GLD (GVZ), or USO (OVX) and SLV (VXSLV). Then, we follow the same estimation procedure outlined in Escobar-Anel and Gong (2020), splitting the parameters into two groups: volatility group and drift group.

After a data description in Section 4.1, Section 4.2 estimates tEstVolGpnder the model with separable spillover effects, we first need to estimate covariance matrix ($\hat{\mathbf{\Sigma}}$) from asset data as a long-term average of the SC matrix ($\mathbf{\Sigma_t}$). This permits us to produce and estimate for the constant eigenvectors, denoted as $\hat{\mathbf{A}}$. With the estimated eigenvectors, we decompose our original asset processes into the sum of independent mean-reverting 4/2 models. The volatility group then consists of parameters for the underlying CIR processes driving the principal components: $b_j, \alpha_j, \theta_j, \xi_j$. Section 4.3 tackles the estimation of drift group parameters, using least squares.

Volatility indexes are functions of implied volatility, model free and directly calculated from option prices from the market. Here, we use volatility indexes data as a convenient proxy for instantaneous volatility. In fact, instantaneous volatility is rather impossible to capture from empirical data, even with high-frequency data, as it requires instantaneous periods rather than the available discrete periods. On the other hand, once a model is specified, volatility indexes can be used to represent instantaneous volatility with some multiplicative (scaling) adjustment or factor; see, for example Luo and Zhang (2012), Zhang and Zhu (2006) and references therein. The relationship between the instantaneous volatility and the volatility index, for example VIX, can be expressed in terms of a closed-form equation, where the difference between the two lies in a multiplicative factor. In a recent paper, see Lin et al. (2017), the author determines the connection between instantaneous and implied volatility, assuming Grasselli's 4/2 model Grasselli (2016) with jumps.

Due to the short horizon of volatility indexes (21-day options), the multiplicative factor could be close to one in a region of the parametric space, which implies that volatility indexes are almost equal to instantaneous volatilities regardless of the structural choice of the underlying model. As precautions and inspired by these pioneering works, we introduce scale parameters to adjust empirical volatility indexes data to estimate instantaneous volatilities. This is done such that the empirical means of the observed variance series ($\mathbf{V_t^{obs}}$) match the corresponding long-term asset variances. These new scaling parameters are estimated at an early stage and are methodologically independent of other parameters.

4.1. Data Description

We consider the following pairs of assets and volatilities: The first study is on VIX (S&P 500 Volatility Index) and VSTOXX (Euro STOXX 50 Volatility Index); here, we also use the volatility indexes VVIX and VVSTOXX, respectively. The second group is comprised of USO (Oil ETF) and GLD (Gold ETF), with OVX and GVZ as the respective volatility indexes. The third and final group is made up of SLV (Silver ETF) and GLD, with VXSLV and GVZ as the respective volatility indexes. All these data sets are daily in the period from late 2010 to early 2020.

Our estimation method consists of two stages: we first estimate the parameters in the assets' data (called drift group), and then the parameters in the assets' volatilities (called the volatility group). The sample size of the raw data is different across all the assets and volatility indexes. Hence, we must further process the data to better suit our estimation purpose, in particular ensuring that we take only the trading days when both assets and their volatilities can be observed. Figures 1–3 depict the pairs of asset data and their volatility indexes. Note that the volatility index data is quoted as annualized volatility multiplied by 100. When we use the volatility index for estimation, we transform the volatility index to daily volatility by dividing by $100 \times \sqrt{250}$.
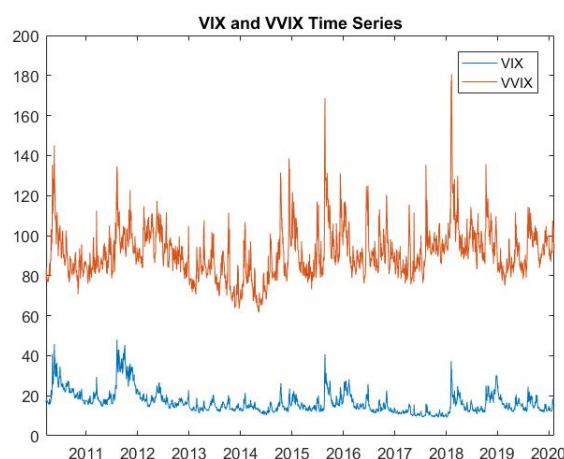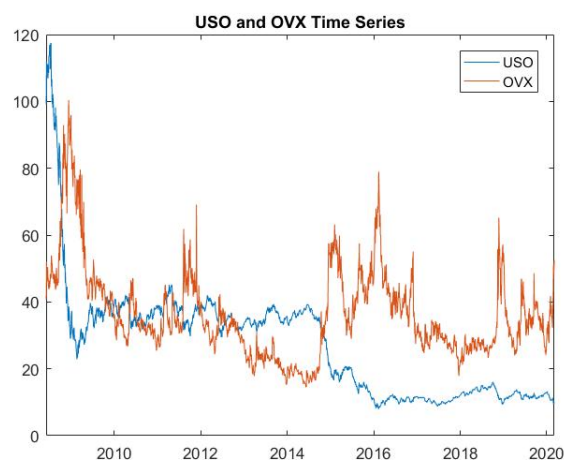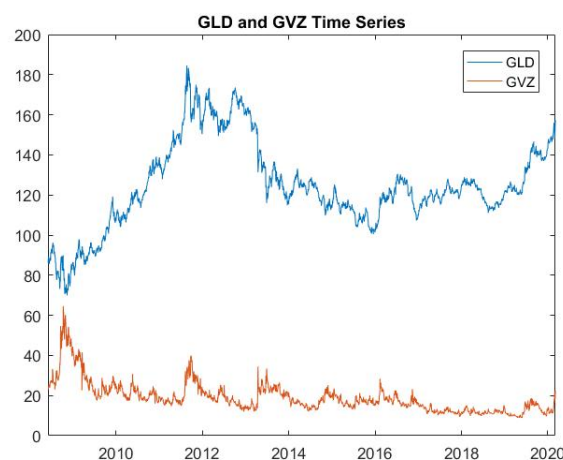

(**a**) Data: VIX and VVIX.
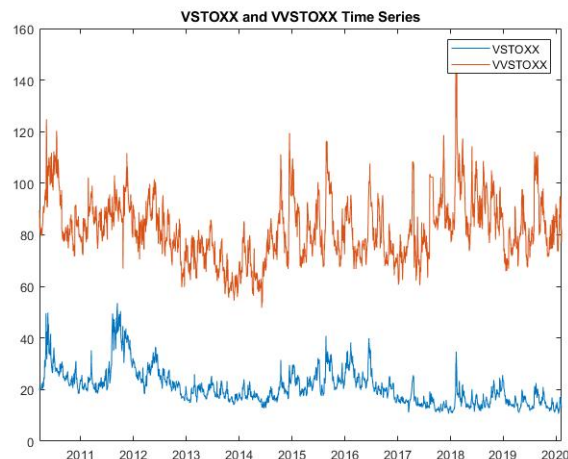
(**b**) Data: VSTOXX and VVSTOXX.

Figure 1. Historical VIX (VVIX) and VSTOXX (VVSTOXX) data.

(**a**) Data: USO and OVX.

(**b**) Data: GLD and GVZ.

Figure 2. Historical USO (OVX) and GLD (GVZ) data.

(**a**) Data: SLV and VXSLV.
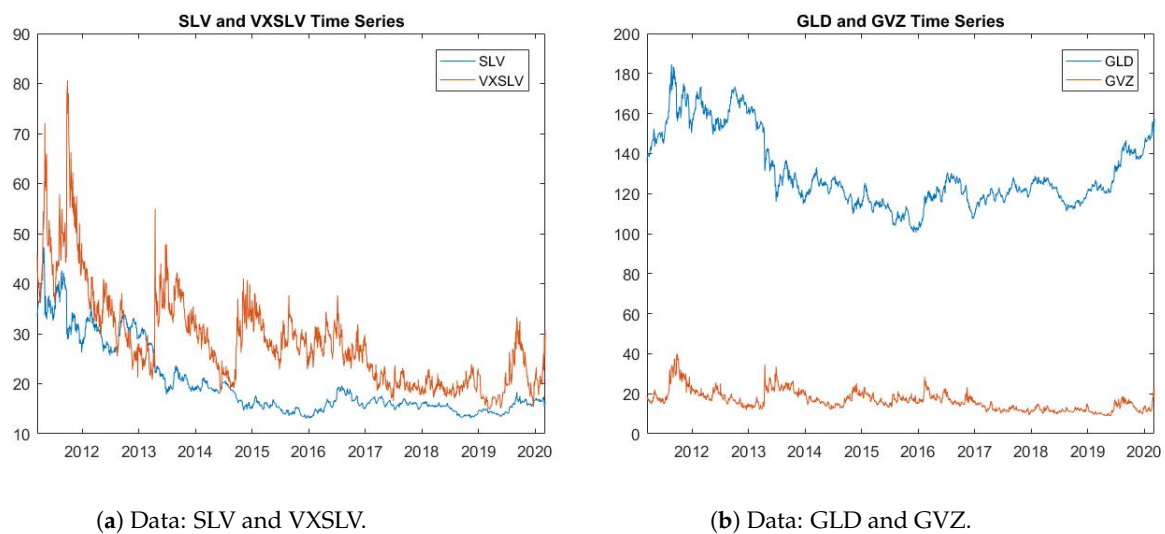
(**b**) Data: GLD and GVZ.

Figure 3. Historical SLV (VXSLV) and GLD (GVZ) data.

4.2. Estimation of Volatility Group Parameters

The model used for estimation of the "volatility group" parameters in this section is the model with separable spillover effects, described next for completeness.

Recall that Equation (6) gives us the representation for each principal component that reflects on our mean-reverting 4/2 stochastic volatility model; that is, $dlog(\mathbf{X_t}) = d\mathbf{Y_t} = \mathbf{A}d\mathbf{M_t}$, with $dM_i(t)$ defined as

$$\begin{cases} dM_i(t) & = \left[\tilde{L}_i + \tilde{c}_i\left(\sqrt{v_i(t)} + \frac{b_i}{\sqrt{v_i(t)}}\right)^2 - \tilde{\beta}_i M_i(t) \right]dt + \left(\sqrt{v_i(t)} + \frac{b_i}{\sqrt{v_i(t)}}\right)dW_i^{\mathbb{P}}(t) \\ dv_i(t) & = \alpha_i(\theta_i - v_i(t))dt + \xi_i\sqrt{v_i(t)}dB_i^{\mathbb{P}}(t) \\ \langle dW_i^{\mathbb{P}}(t), dB_i^{\mathbb{P}}(t) \rangle & = \rho_i dt \end{cases}$$

The estimation procedure for this model setup can be summarized as follows. We first transform the data using matrix $\mathbf{A}$ to produce the $M_i(t)$ process following the relationships among $\mathbf{X_t}$, $\mathbf{Y_t}$ and $\mathbf{M_t}$. Then, we can use the estimation method developed in Escobar-Anel and Gong (2020) for each $M_i(t)$ process. Finally, we recover the parameters for each $X_i(t)$ process.

4.2.1. Estimation of Matrix $\mathbf{A}$ and the Scaling Parameters S

In the next sections, we estimate the parameters in the volatility group. The empirical results are summarized in Table 1. The first step is to estimate matrix $\mathbf{A}$, as it connects log asset prices $\mathbf{Y_t}$ and principal components $\mathbf{M_t}$. Recall that $\mathbf{A}$ is an orthogonal matrix comprising the eigenvectors of covariance matrix $\mathbf{\Sigma_t}$. Given daily data, we estimate $\mathbf{A}$ by first calculating the empirical covariance matrix $\widehat{\mathbf{\Sigma}}$ and applying eigenvalue decomposition:

$$\widehat{\mathbf{\Sigma}} = \hat{\mathbf{A}}diag(\sigma^{(1)}, \ldots, \sigma^{(n)})\hat{\mathbf{A}}',$$

where $(\sigma^{(1)}, \ldots, \sigma^{(n)})$ is a vector of the eigenvalues of $\widehat{\mathbf{\Sigma}}$, and $\hat{\mathbf{A}}$ is the estimate of matrix $\mathbf{A}$. In Table 1, we include the results for $\widehat{\mathbf{\Sigma}}$, $\hat{\mathbf{A}}$ and eigenvalues $(\sigma^{(1)}, \sigma^{(2)})$ from empirical data. Note that $\hat{\mathbf{A}}$ is not unique in that the signs of each element in the matrix can be manipulated such that the column vectors are still the eigenvectors for the corresponding eigenvalues, while $\hat{\mathbf{A}}$ preserves its orthogonality.

As mentioned at the beginning of this section, volatility indexes are a useful proxy for instantaneous volatility; however, they may require scaling adjustment. Let $\mathbf{V_t^{obs}} = (\mathbf{V_t^{(1)}}, \ldots, \mathbf{V_t^{(n)}})'$ denote the squared observed volatility indexes data for n as-

sets. We introduce a scale parameter s_i to bridge observed volatility indexes' series $\mathbf{V_t^{(i)}}$s to theoretical variances via the following relationship:

$$\mathbf{SV_t^{obs}} = (\mathbf{A} \circ \mathbf{A})\mathbf{V_t} \tag{16}$$

$$\mathbf{V_t} = (\mathbf{A} \circ \mathbf{A})^{-1}\mathbf{SV_t^{obs}}, \tag{17}$$

where $\mathbf{S}$ is a diagonal matrix with diagonal vector $(s_1, \ldots, s_n)$. In theory, s_i is a function of $t, T, b_i, \alpha_i, \theta_i, \xi_i$, see Luo and Zhang (2012), and Zhang and Zhu (2006), some of which fall into the volatility group and are to be estimated. Hence, devising an estimate that does not depend on these parameters is crucial. We propose an estimate that matches the long-run first empirical moment of both sides of Equation (16). The long-term average of the left-hand side of Equation (16) can be directly calculated from squared volatility indexes' data. The long-term average of the right-hand side may not seem as straightforward, as it deals with a stochastic process. In our definition, $(\mathbf{A} \circ \mathbf{A})\mathbf{V_t}$ is the diagonal of the covariance matrix and refers to the instantaneous variance process. Therefore, in the long run, the expectation of the variance process should converge to the variance of the underlying asset. Let $\hat{\sigma}_i^2$ denote the empirical long term variance of asset i, and let $\hat{\mu}_i$ denote the long-term average of the corresponding squared volatility index data. We estimate s_i as

$$\hat{s}_i = \frac{\hat{\sigma}_i^2}{\hat{\mu}_i}. \tag{18}$$

Substituting Equation (18) and $\hat{\mathbf{A}}$ back to the right-hand side of Equation (17), the long-term average is

$$\begin{pmatrix} \sigma^{(1)} \\ \vdots \\ \sigma^{(n)} \end{pmatrix} = \mathbb{E}\left[(\mathbf{A} \circ \mathbf{A})^{-1}\hat{\mathbf{S}}\mathbf{V_t^{obs}}\right],$$

which matches the left-hand side of Equation (17) in terms of the long-term average, as $\mathbf{V_t}$ is the eigenvalue of $\mathbf{\Sigma}_t$ and converges to $(\sigma^{(1)}, \ldots, \sigma^{(n)})$ in the long run. Table 1 also displays the results for $\hat{\mathbf{\Sigma}}$, $(\sigma^{(1)}, \sigma^{(2)})$ and $(\hat{s}_1, \hat{s}_2)$.

Table 1. Empirical results (**a**) covariance matrix and long term average of squared volatility indexes; (**b**) eigenvectors, eigenvalues and scaling factors

(a)			
	VIX & VSTOXX	**USO & GLD**	**SLV & GLD**
$\hat{\mathbf{\Sigma}}$	$\begin{pmatrix} 0.0062 & 0.0030 \\ 0.0030 & 0.0048 \end{pmatrix}$	$\begin{pmatrix} 4.78 \times 10^{-4} & 4.62 \times 10^{-5} \\ 4.62 \times 10^{-5} & 1.255 \times 10^{-5} \end{pmatrix}$	$\begin{pmatrix} 0.0724 & 0.0331 \\ 0.0331 & 0.0237 \end{pmatrix}$
$(\hat{\mu}_1, \hat{\mu}_2)$	$(0.0034, 0.0028)$	$(6.01 \times 10^{-4}, 1.67 \times 10^{-4})$	$(3.62 \times 10^{-4}, 1.18 \times 10^{-4})$

(b)			
	VIX & VSTOXX	**USO & GLD**	**SLV & GLD**
$\hat{\mathbf{A}}$	$\begin{pmatrix} 0.7825 & -0.6226 \\ 0.6226 & 0.7825 \end{pmatrix}$	$\begin{pmatrix} 0.9918 & -0.1278 \\ 0.1278 & 0.9918 \end{pmatrix}$	$\begin{pmatrix} 0.8925 & -0.451 \\ 0.451 & 0.8925 \end{pmatrix}$
$\dot{\mathbf{A}}$	$\begin{pmatrix} 2.725 & -1.725 \\ -1.725 & 2.725 \end{pmatrix}$	$\begin{pmatrix} 1.0169 & -0.0169 \\ -0.0169 & 1.0169 \end{pmatrix}$	$\begin{pmatrix} 1.3429 & -0.3429 \\ -0.3429 & 1.3429 \end{pmatrix}$
$(\sigma^{(1)}, \sigma^{(2)})$	$(2.1372, 0.592)$	$(0.121, 0.0299)$	$(0.0891, 0.007)$
$(\hat{s}_1, \hat{s}_2)$	$(1.8074, 1.7229)$	$(0.795, 0.7518)$	$(0.799, 0.8)$

4.2.2. Estimation of Volatility Group

Let $(\hat{\mathbf{A}} \circ \hat{\mathbf{A}})^{-1} = \dot{\mathbf{A}} = \{\dot{a}_{ij}\}_{i,j=1}^n$, and let the j-th eigenvalue of $\mathbf{\Sigma_t}$ be defined as $V_j(t) = v_j(t) + \frac{b_j^2}{v_j(t)} + 2b_j$. Then, $V_j(t)$ is represented by $\sum_{i=1}^n \dot{a}_{ji}\hat{s}_i\mathbf{V_t^{(i)}}$ according to Equation (17). Suppose $\mathbf{V_t^{(i)}} = (V^{(i)}(t_1), \ldots, V^{(i)}(t_n))$ is a series of squared volatility

indexes for asset i observed on $\Omega_T = \{t_i\}_{i=0}^m, t_0 = 0, t_m = T$; then, at time $0 \leq t_k \leq T$, we have $V_j(t_k) = \sum\limits_{i=1}^n \dot{a}_{ji} \hat{s}_i V^{(i)}(t_k)$.

In theory, we expect the $\mathbf{V_t}$ series to be non-negative, since it is related to the series of instantaneous variances $(\mathbf{A} \circ \mathbf{A}) \mathbf{V_t}$ for the underlying assets. In practice, however, we observe inconsistencies in some cases. For example, as Figure 4a illustrates, $V_2(t)$ has a number of negative values (labeled by "V2"), which are non-negligible. Next, we perform a preliminary analysis to locate the root of the issue.

(**a**) Principal components. VIX (VVIX) and VSTOXX (VVSTOXX).

(**b**) Principal components after modification. VIX (VVIX) and VSTOXX (VVSTOXX).
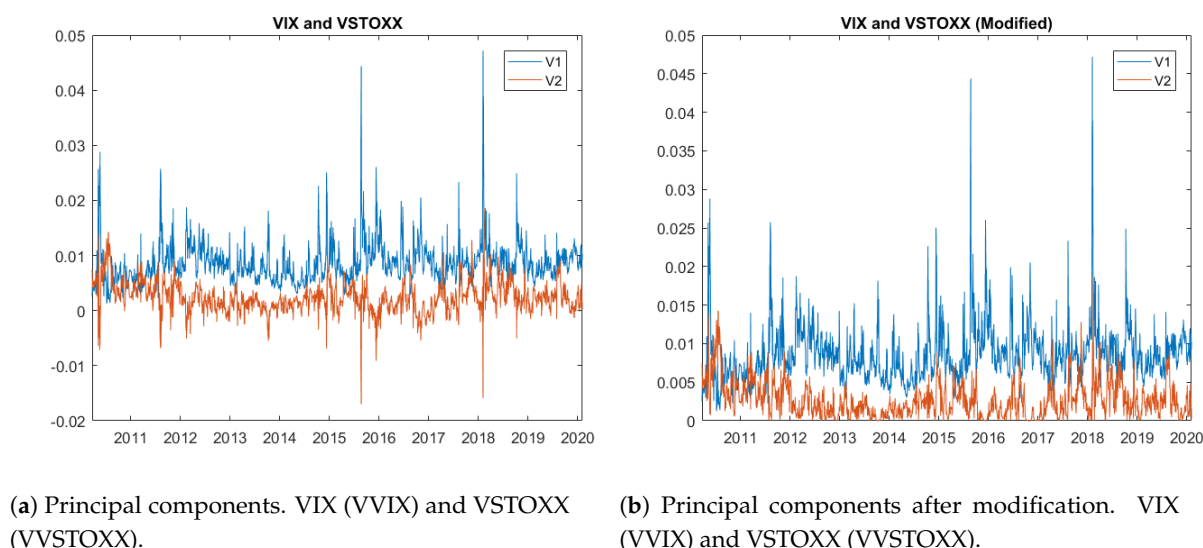
Figure 4. Principal components with volatility indexes' data.

To solve the problem of inconsistency without modifying our model, we deal with the negative values as if they are missing values. We thus replace the negative values by the weekly averages centered on those negative values. Figure 4b illustrates the series of $V_1(t)$ and $V_2(t)$ after this modification. Furthermore, Figure 5 presents two series: $V_1(t)$ and $V_2(t)$ after transforming the original OVX and GVZ data. In this case, we do not observe the inconsistency shown in Figure 4b, which means that the data supports our model. The figure also illustrates the trend as expected, with the first principal component generating the largest variation ($V_1(t)$) in asset price compared to the second principal component ($V_2(t)$).
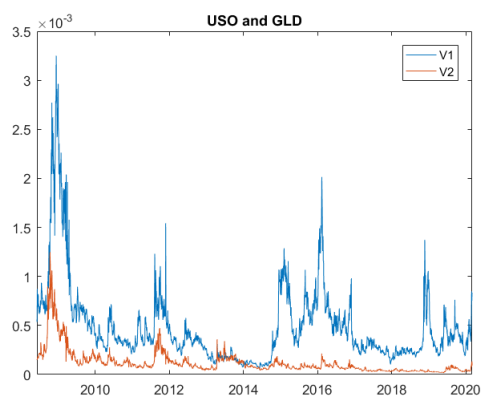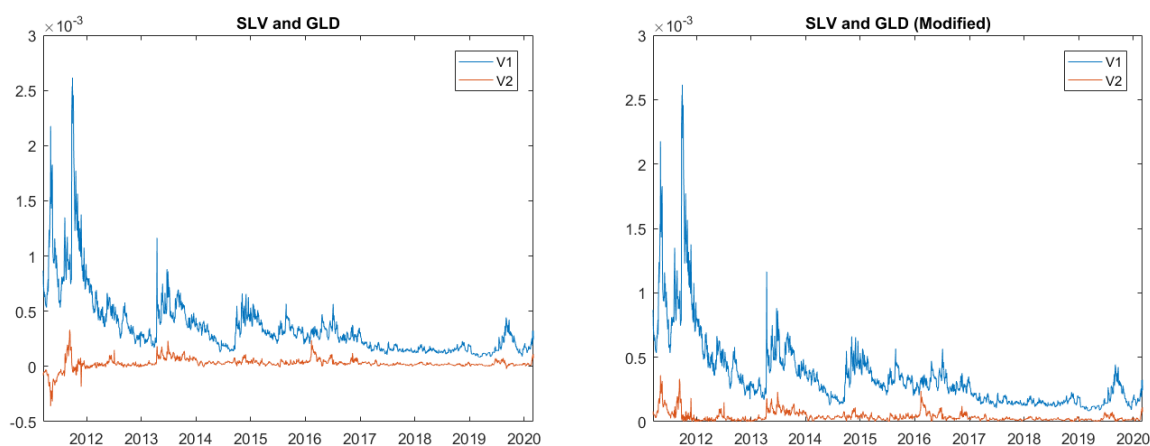

Figure 5. Principal components. Data: USO (OVX) and GLD (GVZ).

In Figure 6a, we also observe some inconsistency in silver ETF (SLV) and gold ETF (GLD) data between 2011 and 2012. Given that the correlation between SLV and GLD is

large, $V_2(t)$ series stays close to 0, which implies that the two assets are likely driven by the same random factor. Since the negative values do not appear as often as in Figure 4a and are close to 0, we simply take the absolute value of the negative values and show the modified series in Figure 6b.

Now that we have prepared all the data for estimation, we apply the estimation method developed in Escobar-Anel and Gong (2020) to $V_1(t)$ and $V_2(t)$ to estimate $b_1, \alpha_1, \theta_1, \xi_1$ and $b_2, \alpha_2, \theta_2, \xi_2$. Note that, in all three scenarios, the minimum of $V_2(t)$ is approximately zero, which implies that b_2 is 0 as seen from Figures 4b, 5 and 6. Therefore, it is sufficient to model $V_2(t)$ as a CIR (1/2) process instead of a 4/2 process. On the other hand, the "spikes" occurred frequently in $V_1(t)$—labeled as "V1" and shown by the figures—are signals that $V_1(t)$ should be a 4/2 process given all three pairs of assets-and-volatility index data. Since we assume that $V_2(t)$ follows a CIR process, we estimate $\alpha_2, \theta_2, \xi_2$ using maximum likelihood. Tables 2–4 list the estimated parameters and their standard errors (s.es) with the chosen data sets for parameters in the volatility group.

The inference on the parameters (asymptotic mean and variance) is performed via parametric bootstrap. In other words, we simulate the corresponding processes with the estimated parameters 1000 times and repeat the estimation procedure for each simulation. In the end, we obtain a pool of 1000 sets of estimates. The law of large numbers suggests that the means calculated from the pool of estimates are the asymptotic means for each estimator. It is interesting to see how the first principal component not only accounts for most of the variation, but it also absorbs the complexity of the problem. In other words. the tables show that the first component requires the advanced 4/2 modeling (i.e., $b \neq 0$), while the second component can be better explained with the simpler 1/2 model ($b = 0$).

(**a**) Principal components. Data: SLV (VXSLV) and GLD (GVZ).

(**b**) Principal components after modification. Data: SLV (VXSLV) and GLD (GVZ).

Figure 6. Principal components with volatility indexes' data.

Table 2. Estimated volatility group parameters with empirical data (VIX and VSTOXX).

	Data: $V_1(t)$	Data: $V_2(t)$
$\hat{\mathbf{b}}$	3.11×10^{-4}	~ 0
Mean of $\hat{\mathbf{b}}$ (s.e)	$4.511 \times 10^{-4} (1.1581 \times 10^{-6})$	$0(0)$
$\hat{\alpha}$	42.1811	19.1624
Mean of $\hat{\alpha}$ (s.e)	$41.6498(0.0695)$	$19.5827(0.0552)$
$\hat{\theta}$	0.0079	0.0027
Mean of $\hat{\theta}$ (s.e)	$0.0076(4.8578 \times 10^{-6})$	$0.0027(8.0796 \times 10^{-6})$
$\hat{\xi}$	0.3436	0.3885
Mean of $\hat{\xi}$ (s.e)	$0.3537(1.3937 \times 10^{-4})$	$0.3877(5.1103 \times 10^{-4})$

Table 3. Estimated volatility group parameters with empirical data (USO and GLD).

	Data: $V_1(t)$	Data: $V_2(t)$
$\hat{\mathbf{b}}$	1.6776×10^{-5}	~ 0
Mean of $\hat{\mathbf{b}}$ (s.e)	$7.5578 \times 10^{-5}(1.3892 \times 10^{-6})$	$0(0)$
$\hat{\alpha}$	3.62	5.3597
Mean of $\hat{\alpha}$ (s.e)	$5.9253(0.0465)$	$5.6994(0.0186)$
$\hat{\theta}$	8.9803×10^{-4}	1.1859×10^{-4}
Mean of $\hat{\theta}$ (s.e)	$4.6621 \times 10^{-4}(7.4231 \times 10^{-6})$	$1.1779 \times 10^{-4}(2.4962 \times 10^{-7})$
$\hat{\hat{\xi}}$	0.0271	0.0231
Mean of $\hat{\hat{\xi}}$ (s.e)	$0.02(2.4446 \times 10^{-4})$	$0.0231(5.3704 \times 10^{-6})$

Table 4. Estimated volatility group parameters with empirical data (SLV and GLD).

	Data: $V_1(t)$	Data: $V_2(t)$
$\hat{\mathbf{b}}$	2.0968×10^{-5}	~ 0
Mean of $\hat{\mathbf{b}}$ (s.e)	$7.4123 \times 10^{-5}(2.85 \times 10^{-7})$	$0(0)$
$\hat{\alpha}$	5.178	24.3083
Mean of $\hat{\alpha}$ (s.e)	$7.7687(0.0664)$	$24.8349(0.057)$
$\hat{\theta}$	8.3026×10^{-4}	3.8644×10^{-5}
Mean of $\hat{\theta}$ (s.e)	$4.6417 \times 10^{-4}(7.5518 \times 10^{-6})$	$3.8753 \times 10^{-5}(6.5274 \times 10^{-8})$
$\hat{\hat{\xi}}$	0.0307	0.0343
Mean of $\hat{\hat{\xi}}$ (s.e)	$0.024(3.0304 \times 10^{-4})$	$0.0344(1.1187 \times 10^{-5})$

4.3. Estimation of Drift Group

Similarly, we use the least square approach to estimate parameters in the drift group. Tables 5–7 display the results. Some parameters are assessed to be non-significant based on the p-values. We decide to keep all the parameters because our sample sizes are not large enough to draw concrete conclusions on the significance of the parameters.

Note that the estimated parameters reported in the tables are for the parameters of the $M_1(t)$ and $M_2(t)$ processes. We can recover the estimates for original parameters using the relationship we defined earlier; that is, $\hat{\mathbf{L}} = \hat{\mathbf{A}}\tilde{\mathbf{L}}$, $\hat{\mathbf{B}} = \hat{\mathbf{A}}\tilde{\mathbf{B}}\hat{\mathbf{A}}^{-1}$, and $\hat{\mathbf{C}} = \hat{\mathbf{A}}\tilde{\mathbf{C}} + \frac{1}{2}(\hat{\mathbf{A}} \circ \hat{\mathbf{A}})$. The estimates for the original parameters are reported in Table 8. The diagonal entries in the $\hat{\mathbf{B}}$ matrices provide information on the mean-reverting speed for all the assets. Note that for the pair SLV and GLD, one of the eigenvalues (-0.5401) of $\hat{\mathbf{B}}$ does not satisfy the assumption imposed on the eigenvalues of $\mathbf{B}$, which is a sign that the data does not support this particular model. The correlation coefficients are not included because they are not affected by the transformation.

Table 5. Estimated drift group parameters (VIX and VSTOXX).

Data	$\hat{L}_i$	p-Value	$\hat{c}_i$	p-Value	$\hat{\beta}_i$	p-Value	$\hat{\rho}$	p-Value
$M_1(t) \& V_1(t)$	25.6041	0	-114.6521	0.436	6.1077	0	0.5621	0
$M_2(t) \& V_2(t)$	9.4459	0	73.6448	0.734	15.6266	0	0.00419	0

Table 6. Estimated drift group parameters (USO and GLD).

Data	$\hat{L}_i$	p-Value	$\hat{c}_i$	p-Value	$\hat{\beta}_i$	p-Value	$\hat{\rho}$	p-Value
$M_1(t) \& V_1(t)$	0.8096	0.154	-416.2006	0.214	0.214	0.134	-0.3723	0
$M_2(t) \& V_2(t)$	2.6418	0.07	-646.7339	0.384	0.5701	0.079	-0.00294	0

Table 7. Estimated drift group parameters (SLV and GLD).

Data	$\hat{L}_i$	*p*-Value	$\hat{c}_i$	*p*-Value	$\hat{\beta}_i$	*p*-Value	$\hat{\rho}$	*p*-Value
$M_1(t) \& V_1(t)$	4.5459	0.009	842.936	0.153	1.0183	0.008	−0.2323	0
$M_2(t) \& V_2(t)$	−1.5092	0.23	−1787.0546	0.16	−0.5401	0.191	0.1228	0

Table 8. Estimated original drift group parameters.

Data	$\hat{\mathbf{L}}$	$\hat{\mathbf{C}}$	$\hat{\mathbf{B}}$
VIX&VSTOXX	(25.9172, 8.5495)	$\begin{pmatrix} -89.4091 & 46.0451 \\ -71.1886 & -57.3209 \end{pmatrix}$	$\begin{pmatrix} 9.7978 & -4.6378 \\ -4.6378 & 11.9365 \end{pmatrix}$
USO&GLD	(0.4653, 2.7236)	$\begin{pmatrix} -412.2959 & -82.6444 \\ -53.1823 & 641.9225 \end{pmatrix}$	$\begin{pmatrix} 0.2198 & -0.0451 \\ -0.0451 & 0.5643 \end{pmatrix}$
SLV&GLD	(4.7379, 0.7032)	$\begin{pmatrix} 752.7178 & -805.8599 \\ 380.2658 & 1595.3 \end{pmatrix}$	$\begin{pmatrix} 0.7013 & 0.6273 \\ 0.6273 & -0.2231 \end{pmatrix}$

It is worth noting that $\hat{\mathbf{L}}$ in Table 8 does not reflect the actual mean-reverting level. Therefore, to determine the mean-reverting level for each asset, we must go back to Equation (4) and rewrite it in following format:

$$\mathbf{dY_t} = \mathbf{B}(\mathbf{B}^{-1}\mathbf{L} + \mathbf{B}^{-1}\mathbf{A}\tilde{\mathbf{C}}\mathbf{V_t} - \mathbf{Y_t})dt + \mathbf{A}diag(\mathbf{V_t})^{\frac{1}{2}}\mathbf{dW_t^{\mathbb{P}}}.$$

We can now see that the mean-reverting level is $\mathbf{B}^{-1}\mathbf{L}$ plus a random component, which we define as $H(\mathbf{V_t}) = \mathbf{B}^{-1}\mathbf{A}\tilde{\mathbf{C}}\mathbf{V_t}$. The long term mean indicated by the model is basically $\mathbf{B}^{-1}\mathbf{L} + \mathbb{E}\big[H(\mathbf{V_t})\big]$. We report these estimates in Table 9 and compare them with the averages calculated from empirical data.

As Table 9 shows, the estimated MRLs are close to the empirical log price averages, except for the USO case, where the estimated mean is smaller than the empirical mean. This latest point might be due to the impact of the initial value on the stationary value of a 4/2 process. Moreover, the VIX and VSTOXX pair has the largest mean-reverting speed compared to the other two commodity ETF pairs. This is not a surprise, as evidenced by empirical data. Volatility indexes tend to return to the mean faster due to an economic cycle, whilst commodities normally have a longer time horizon to revert to the mean level due to scarcity, demand and supply.

Table 9. Estimated original drift group parameters.

Data	$\mathbf{B}^{-1}\mathbf{L}$	$\mathbb{E}\big[H(\mathbf{V_t})\big]$	Estimated Mean-Reverting Level	Empirical Averages
VIX&VSTOXX	(2.9042, 3.0832)	(−0.0952, −0.0631)	(2.809, 3.0219)	(2.6497, 2.9139)
USO&GLD	(3.1598, 5.0794)	(−0.961, −0.2547)	(2.244, 4.8247)	(3.1342, 4.8096)
SLV&GLD	(2.7242, 4.5075)	(0.2215, 0.2158)	(2.9457, 4.7232)	(2.9468, 4.8645)

5. Application to Risk Measures

Risk measures in financial risk management are used to determine the minimum amount of capital to be kept in reserve in worst-case scenarios as a way of protecting financial institutions. There are many risk measures in the literature, see, for example, Artzner et al. (1999) and McNeil et al. (2005), one of which is considered fundamental: Value-at-Risk (VaR), which is a distribution-based risk measure. In other words, a VaR calculation takes into account the distribution of the underlying (VaR is in fact a quantile). It is more robust to outliers than mean and variance.

In this section, we compute the VaR of a portfolio consisting of two assets and a cash account, in line with the previous estimation section. We must first find the distribution of this portfolio, which might not be available due to, in particular, the correlations among the underlyings. In the language of mathematical statistics, we must find the joint distribution

of USO and GLD to compute VaR. In general, finding closed-form expressions for the joint distribution of two non-Gaussian stochastic processes is theoretically difficult. In fact, USO and GLD have complex distribution functions under our multidimensional 4/2 model setting. Fortunately, this is feasible in our model, as we can express the joint distribution at any given date of USO and GLD in terms of two independent random variables, which simplifies our problem significantly and allows for the use of c.f.s to compute the properties of the portfolio distribution.

5.1. Portfolio Setup

Suppose that we have a portfolio Π consisting of two assets $X_1(t)$ and $X_2(t)$:

$$\Pi(t) = \phi_1(t)X_1(t) + \phi_2(t)X_2(t) + \phi_3(t)B(t), \tag{19}$$

where ϕ_1 and ϕ_2 represent the weights of $X_1(t)$ and $X_2(t)$ in the portfolio, and $B(t)$ is a cash account with interest rate r. In a short period of time, we can also write the problem using the self-financing condition and relative portfolio weights π_1, π_2 and $(1 - \pi_1 - \pi_2)$; that is, the proportion allocated to the assets and cash account, allocations see Campell et al. (1997):

$$\frac{d\Pi(t)}{\Pi(t)} = \pi_1 \frac{dX_1(t)}{X_1(t)} + \pi_2 \frac{dX_2(t)}{X_2(t)} + (1 - \pi_1 - \pi_2)rdt. \tag{20}$$

Constant allocations will be considered (i.e., constant π), as they constitute the most popular investment strategy in the market, supported by Merton (1975). From the process $\frac{dX_i(t)}{X_i(t)}$, we can easily obtain $dY_i(t)$ ($Y_i(t) = log(X_i(t))$) by using Ito's lemma. When comparing $\frac{dX_i(t)}{X_i(t)}$ and $dY_i(t)$, we observe that only the drift term is adjusted, while diffusion terms stay the same. Assuming that $(X_1(t), X_2(t))$ are modeled by Equation (2), the log prices $(Y_1(t), Y_2(t))$ then have the SDE specified in Equation (4) under the PCSV framework. Moreover, we can also write $\frac{dX_i(t)}{X_i(t)}$ in terms of $dY_i(t)$:

$$\frac{dX_i(t)}{X_i(t)} = dY_i(t) + \frac{1}{2} \sum_j a_{ij}^2 \left(\sqrt{v_j(t)} + \frac{b_j}{v_j(t)} \right)^2 dt.$$

Hence, we rewrite Equation (20) as follows:

$$\frac{d\Pi(t)}{\Pi(t)} = \pi_1 \left[dY_1(t) + \frac{1}{2} \sum_{j=1}^{2} a_{1j}^2 \left(\sqrt{v_j(t)} + \frac{b_j}{v_j(t)} \right)^2 dt \right] + \pi_2 \left[dY_2(t) + \frac{1}{2} \sum_{j=1}^{2} a_{2j}^2 \left(\sqrt{v_j(t)} + \frac{b_j}{v_j(t)} \right)^2 dt \right]$$
$$+ (1 - \pi_1 - \pi_2)rdt. \tag{21}$$

It is known that $Y_1(t)$ and $Y_2(t)$ are linear combinations of two independent stochastic processes or random variables $M_1(t)$ and $M_2(t)$:

$$dY_1(t) = a_{11}dM_1(t) + a_{12}dM_2(t), \tag{22}$$
$$dY_2(t) = a_{21}dM_1(t) + a_{22}dM_2(t). \tag{23}$$

We now substitute $dY_1(t)$ and $dY_2(t)$ in Equation (21) with Equations (22) and (23):

$$\frac{d\Pi(t)}{\Pi(t)} = (\pi_1 a_{11} + \pi_2 a_{21})dM_1(t) + \frac{1}{2}(\pi_1 a_{11}^2 + \pi_2 a_{21}^2) \left(\sqrt{v_1(t)} + \frac{b_1}{v_1(t)} \right)^2 dt$$
$$+ (\pi_1 a_{12} + \pi_2 a_{22})dM_2(t) + \frac{1}{2}(\pi_1 a_{12}^2 + \pi_2 a_{22}^2) \left(\sqrt{v_2(t)} + \frac{b_2}{v_2(t)} \right)^2 dt + (1 - \pi_1 - \pi_2)rdt. \tag{24}$$

From Equation (24), we can conclude that $\frac{d\Pi(t)}{\Pi(t)}$ is also a linear combination of $dM_1(t)$ and $dM_2(t)$, with adjustment to the drift terms, which does not affect the independent relationship between $dM_1(t)$ and $dM_2(t)$. We organize Equation (24) into the following expression:

$$\frac{d\Pi(t)}{\Pi(t)} = dM_1^*(t) + dM_2^*(t) \tag{25}$$

where $dM_1^*(t)$ and $dM_2^*(t)$ are independent, with

$$dM_i^*(t) = \left[L_i^* + c_i^* \left(\sqrt{v_i(t)} + \frac{b_i}{\sqrt{v_i(t)}} \right)^2 - \beta_i^* M_i^*(t) \right] dt + a_i^* \left(\sqrt{v_i(t)} + \frac{b_i}{\sqrt{v_i(t)}} \right) dW_i^{\mathbb{P}}(t),$$

where $L_1^* = \tilde{L}_1 a_1^* + (1 - \pi_1 - \pi_2)r$, $L_2^* = \tilde{L}_2 a_2^*$, $c_i^* = \tilde{c}_i a_i^* + \frac{1}{2} a_i^{**}$, $\beta_i^* = \tilde{\beta}_i a_i^*$, $a_i^* = \sum_{j=1}^2 \pi_j a_{ji}$, and $a_i^{**} = \sum_{j=1}^2 \pi_j a_{ji}^2$.

Note that for convenience, we included the growth rate on the cash account into the long-term average of $M_1^*(t)$. From a mathematical perspective, $M_1^*(t)$ is constructed based on $M_1(t)$, which is the first principal component that determines the most variation among the assets. As a result, $M_1^*(t)$ affects the performance of $\Pi(t)$ more than $M_2^*(t)$ financially. For this reason, the effect of the growth rate on the portfolio can also be interpreted as if it impacts the long-term average of $M_1^*(t)$. Now, we apply Ito's lemma to Equation (25) to obtain the dynamics for $ln(\Pi(t))$:

$$dln(\Pi(t)) = d\tilde{M}_1^*(t) + d\tilde{M}_2^*(t) \tag{26}$$

where

$$d\tilde{M}_i^*(t) = \left[L_i^* + (c_i^* - \frac{1}{2}(a_i^*)^2) \left(\sqrt{v_i(t)} + \frac{b_i}{\sqrt{v_i(t)}} \right)^2 - \beta_i^* \tilde{M}_i^*(t) \right] dt + a_i^* \left(\sqrt{v_i(t)} + \frac{b_i}{\sqrt{v_i(t)}} \right) dW_i^{\mathbb{P}}(t).$$

It is straightforward to find the characteristic function of $dln(\Pi(t))$ using the above mentioned results.

5.2. The Density Function of the Portfolio $\Pi(t)$

From Equation (26), our portfolio now essentially contains two new "assets" that are independent of each other. Thanks to this independence, we can derive the characteristic function as well as the density function of our portfolio. Since our goal is to calculate the VaR, it is convenient to use a density function and integrate numerically. In this section, we list two approaches to obtain such a density function.

5.2.1. Density Function via Convolution

One way in which to obtain the conditional density function for $ln(\Pi(T))$ is via convolution of two conditional density functions for $\tilde{M}_1^*(T)$ and $\tilde{M}_2^*(T)$. In probability, if two random variables X and Y are independent, with density functions $f_X(x)$ and $f_Y(y)$, respectively, then the density for $Z = X + Y$, $f_Z(z)$, can be found via convolution; that is,

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x) dx. \tag{27}$$

If X and Y have analytical density functions, then the convolution method is straightforward. In our case, we obtain the conditional c.f. first. The Fourier inversion of the conditional c.f. theoretically gives the density function. However, due to the structure of our original c.f. and the approximations, we need to invert both the original c.f. and the approximated c.f.s numerically for the corresponding density functions. For $\tilde{M}_1^*(T)$ and $\tilde{M}_2^*(T)$, we can obtain their conditional corresponding density functions $f_1(m_1|\mathcal{F}_t)$ and $f_2(m_2|\mathcal{F}_t)$ by inverting the c.f. of $\tilde{Z}_1^*(t) = e^{\beta_1^* t} \tilde{M}_1^*(t)$ and $\tilde{Z}_2^*(t) = e^{\beta_2^* t} \tilde{M}_2^*(t)$:

$$f_1(m_1|\mathcal{F}_t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ium_1} \Phi_{MR}\Big(T, u; L_1^*, c_1^*, a_1^*, \beta_1^*, \alpha_1, \theta_1, \xi_1, \rho_1, a_1^* b_1, \nu_1(t), e^{-\beta_1^* t} \tilde{Z}_1^*(t)\Big) du. \tag{28}$$

$$f_2(m_2|\mathcal{F}_t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ium_2} \Phi_{MR}\Big(T, u; L_2^*, c_2^*, a_2^*, \beta_2^*, \alpha_2, \theta_2, \xi_2, \rho_2, a_2^* b_2, \nu_2(t), e^{-\beta_2^* t} \tilde{Z}_2^*(t)\Big) du. \tag{29}$$

We can now write Equation (32) as a convolution of Equations (28) and (29):

$$f(p|\mathcal{F}_t) = \int_{-\infty}^{\infty} f_1(m_1|\mathcal{F}_t) f_2(p - m_1|\mathcal{F}_t) dm_1. \tag{30}$$

A challenging part of this method is that we must first invert the semi-closed c.f.s to obtain the density functions for $\tilde{M}_1^*(t)$ and $\tilde{M}_2^*(t)$—artificial assets—which involves approximations. As we have well-developed approximation approaches for Φ_{MR}, we can apply the results to obtain the analytic function as an approximation of the c.f. for individual artificial assets and then find the density via Fourier inversion. Then, we can use Equation (30) to obtain the density of the portfolio. The approximation approaches work well in a parametric region, as demonstrated in Escobar-Anel and Gong (2020), for three scenarios ($b_i = 0, \rho_i \neq 0; b_i \neq 0, \rho_i = 0; b_i = 0, \rho_i = 0$); the goodness of approximations depends on $L_i^*, c_i^*, a_i^*, \beta_i^*, \alpha_i, \theta_i, \xi_i, \rho_i, b_i$.

5.2.2. Density Function via Fourier Inversion

Another, more direct means of obtaining the density function is to apply an inverse Fourier transform to the characteristic function. Before we provide the formula for the characteristic function, we consider the transformation $\tilde{Z}_i^*(t) = e^{\beta_i^* t} \tilde{M}_i^*(t)$. By Ito's lemma, we have

$$d\tilde{Z}_i^*(t) = e^{\beta_i^* t}\left[L_i^* + (c_i^* - \frac{1}{2}(a_i^*)^2)\Big(\sqrt{\nu_i(t)} + \frac{b_i}{\sqrt{\nu_i(t)}}\Big)^2\right]dt + e^{\beta_i^* t} a_i^*\Big(\sqrt{\nu_i(t)} + \frac{b_i}{\sqrt{\nu_i(t)}}\Big)dW_i^{\mathbb{P}}(t).$$

The next corollary explains how to derive the characteristic function.

Corollary 4. *Let* $\Phi_{MR}\Big(T, u; L, c, a, \beta, \alpha, \theta, \xi, \rho, b, \nu(t), Z_(t)\Big)$ *denote the characteristic function provided in Proposition 2.1 in* Escobar-Anel and Gong (2020)*; then, the characteristic function of* $ln(\Pi(T))$ *is given by the following equation:*

$$\mathbb{E}(e^{iuln(\Pi(T))}|\mathcal{F}_t) = \prod_{i=1}^{2} \Phi_{MR}\Big(T, u; L_i^*, c_i^*, a_i^*, \beta_i^*, \alpha_i, \theta_i, \xi_i, \rho_i, a_i^* b_i, \nu_i(t), \tilde{Z}_i^*(t)\Big) \tag{31}$$

where $L_1^* = \tilde{L}_1 a_1^* + (1 - \pi_1 - \pi_2)r$, $L_2^* = \tilde{L}_2 a_2^*$, $c_i^* = \tilde{c}_i a_i^* + \frac{1}{2}a_i^{**}$, $\beta_i^* = \tilde{\beta}_i a_i^*$, $a_i^* = \sum_{j=1}^{2} \pi_j a_{ji}$, $a_i^{**} = \sum_{j=1}^{2} \pi_j a_{ji}^2$.

Let p denote all possible values in the domain of $ln(\Pi(T))$; the density function hence follows from

$$f(p|\mathcal{F}_t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iup} \mathbb{E}(e^{iuln(\Pi(T))}|\mathcal{F}_t) du. \tag{32}$$

It can be seen that the c.f. of the portfolio does not have a closed-form representation, since it is a product of semi-closed c.f.s (Φ_{MR}). As a result, we could use the approximations developed in Section 3.3.

5.2.3. Numerical Implementation of Selected Method

The Fourier inversion method and the convolution method theoretically yield the same density function for the portfolio. Moreover, in a portfolio that only consists of two (artificial) risky assets, both methods are not complicated to implement. We implement the convolution method with partial simulation for applications with only two assets.

However, it would be more efficient to use the Fourier inversion method when the portfolio has a large pool of assets (e.g., over 100). To see this, note that, for $n >> 2$ assets, the convolution method involves the simulation of n processes with $n + 1$ integrations; in contrast, the Fourier inversion method reduces the number of integrations to just one.

We summarize the numerical implementation to compute the conditional density function of the portfolio in the following steps:

- Step 1: Simulate two CIR processes $v_1(t)$ and $v_2(t)$ and compute Φ_{MR} for $\tilde{M}_1^*(t)$ and $\tilde{M}_2^*(t)$.
- Step 2: Invert the c.f.s obtained in Step 1 to obtain $f_1(m_1|\mathcal{F}_t)$ and $f_2(m_2|\mathcal{F}_t)$.
- Step 3: Numerically integrate the product of the conditional density of $\tilde{M}_1^*(T)$ and $\tilde{M}_2^*(T)$ for the conditional density function of $ln(\Pi(T))$.

Even though we use partial simulation to obtain the density function of the portfolio, partial simulation is not time-consuming, as efficient methods exist to simulate CIR processes; see, for example, Andersen (2007). Moreover, both the convolution method and the direct Fourier inversion method require either fewer simulations or no simulations (via approximations), compared to a full simulation approach, which would require the simulation of four processes. In the case where semi-closed c.f.s are involved, we would only need to simulate at most $n \geq 2$ processes ($v_i(t)$ for $i = 1, \ldots, n$ process), instead of a simulation of the $\tilde{M}_1^*(t)$ processes altogether ($2n$ processes). Most importantly, thanks to PCA decomposition, we would likely need m such volatility drivers to explain the SC of n assets with $m << n$. This means a substantial reduction in computational complexity (partial, simulations, integrations or approximations).

In summary, under a PCSV framework, partial simulation is a good choice in terms of efficiency. The PCA reduces computational complexity, as fewer diffusions may be required to explain the variation of all assets. Our approximations further improve the efficiency for computing the c.f.s with analytic functions.

In the next application section, we apply the convolution method from Section 5.2.1 to compute the VaR at popular quantile $\alpha_q = 0.95$ ($VaR_{0.95}$).

5.3. The VaR for a Portfolio of USO and GLD

In this section, we consider a pair of risky assets: USO and GLD, and we study $VaR_{0.95}$ under an investment strategy: equally weighted risky assets only ($\pi_1 = \pi_2 = 0.5$)[2]. These have been proven to be robust and reliable strategies in the seminal work of DeMiguel et al. (2009). In Table 10, we report the $VaR_{0.95}$ values and their s.es. We consider a well-known asymptotic result for quantiles to calculate the s.es for VaR_{α_q} as derived in Stuart et al. (1994):

$$s.e(VaR_{\alpha_q}) = \sqrt{\frac{\alpha_q(1 - \alpha_q)}{nf(VaR_{\alpha_q})^2}},$$

where α_q is the quantile of the portfolio distribution (in this case, it is 95%); n is the sample size, and $f(VaR_{\alpha_q})$ is the probability distribution function (density function) evaluated at VaR_{α_q}.

Case: $\pi_1 = \pi_2 = 0.5$

In this case, we use the information from Tables 1b, 3 and 6 to obtain parameters that generate $\tilde{M}_2^*(T)$ or $\tilde{M}_1^*(T)$. Figures 7 and 8 confirm that density functions from theory are in line with the simulations. We compute and compare $VaR_{0.95}$ values from four sources: simulation of the portfolio (Simulation), density function without approximation (Density w/o Approximation), approximated density function using the midpoint approach (Approx. Density (M)) and approximated density function using the average approach (Approx. Density (A)).
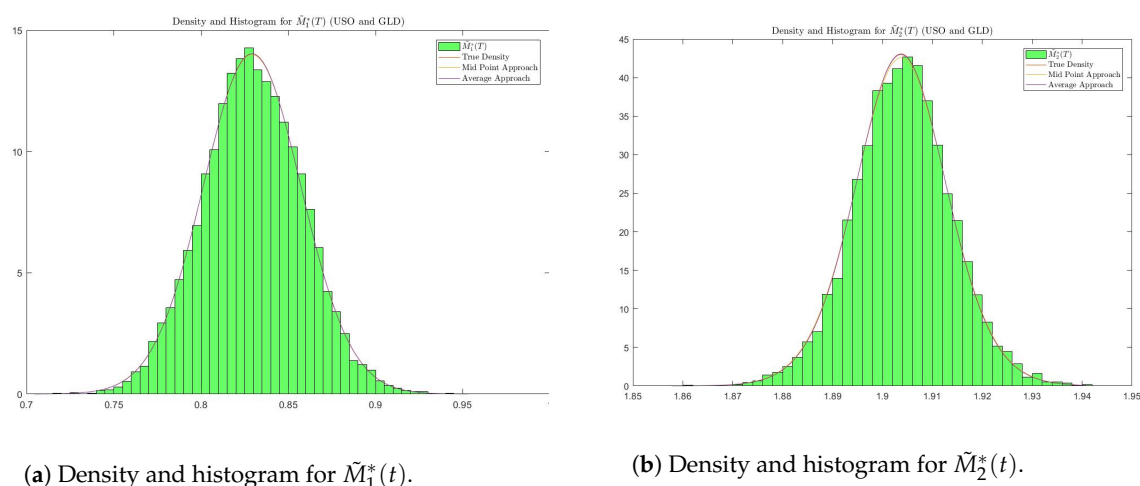
(**a**) Density and histogram for $\tilde{M}_1^*(t)$.

(**b**) Density and histogram for $\tilde{M}_2^*(t)$.

Figure 7. Case 1: Density and histogram for $\tilde{M}_1^*(t)$ and $\tilde{M}_2^*(t)$.
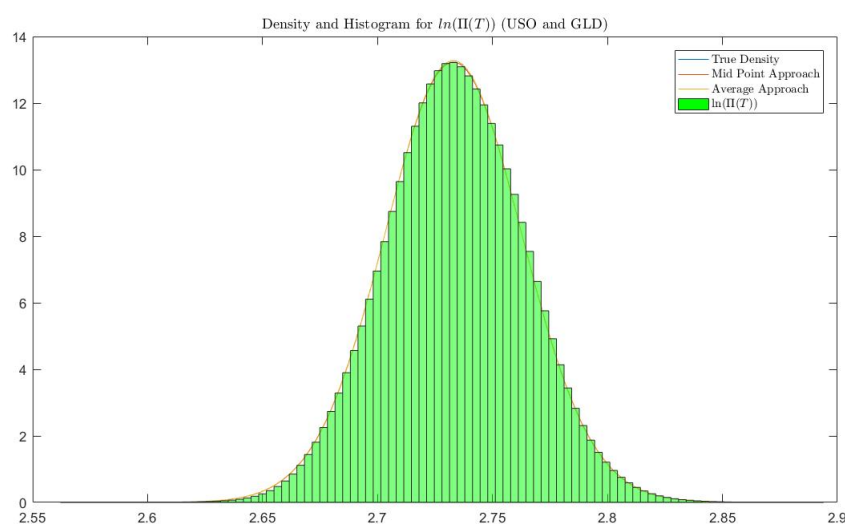

Figure 8. Case 1: Density and histogram for $ln(\Pi(T))$.

Table 10. $VaR_{0.05}$ for $ln(\Pi(t))$ from four sources.

	Simulation	Density w/o Approximation	Approx. Density (M)	Approx. Density (A)
$VaR_{0.05}$	2.7833 (0.004006)	2.783 (0.00401)	2.7832 (0.003988)	2.783 (0.004008)

We use linear interpolation here to calculate the quantile if α_q falls between two critical levels calculated from histogram and density functions. Standard errors are reported in parantheses.

6. Conclusions

This paper studied the properties of a multivariate mean-reverting 4/2 stochastic volatility model based on principal component decomposition. In particular, we studied the variance and covariance processes as well as several submodels of interest to the industry (e.g., separable or no spillover effects, multivariate mean-reverting Heston models). We also obtain an expectation representation for the c.f. of the asset prices with respect to the paths of the stochastic volatility process. Two closed-form approximations to the c.f. are presented in Section 3, these are the first efficient calculations of c.f. for multivariate

mean-reverting stochastic covariance models. In Section 4, we implemented a two-step estimation methodology to three sets of data involving two asset classes, commodities and volatility indexes. The study confirms stylized facts commonly attributed to commodities, like spillover effects, are also observed in the joint dynamic of volatility indexes, which has not been previously reported; it also displays the role and need of scaling parameters between instantaneous variance and volatility indexes in a multidimensional setting.

In Section 5, we further tested our approximation methods in a risk management setting by computing one of the most popular risk measures, namely, VaR. Since VaR is a distribution-based risk measure, our analysis confirms the effectiveness of the c.f. approximations in a multidimensional setting for a portfolio of advanced stochastic processes. Although our analysis was in two dimensions, the methodology is transferable to any dimension, e.g., a portfolio with a large number of underlying assets. In such case, the average-based approximation can greatly save time in calculating distribution-based risk measures, the alternative is the MontCarlo simulation of a high number of continuous-time processes with the subsequent loss in precision.

Author Contributions: Conceptualization, M.E.-A. and Z.G.; methodology, M.E.-A. and Z.G.; software, Z.G.; validation, M.E.-A. and Z.G.; formal analysis, M.E.-A. and Z.G.; investigation, M.E.-A. and Z.G.; resources, Z.G.; data curation, Z.G.; writing—original draft preparation, M.E.-A. and Z.G.; writing—review and editing, ME.-A.; visualization, Z.G.; supervision, M.E.-A.; project administration, M.E.-A.; funding acquisition, M.E.-A. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Notes

1. The estimation methodology for volatility group parameters (Section 4.2) can also be applied to the general model, as for drift group parameters (Section 4.3) some modifications are needed to account for the vector autoregressive structure coming from spillover effects.
2. Similar results were obtained for equally weighted assets case ($\pi_1 = \pi_2 = \frac{1}{3}$).

References

Alexander, Carol. 2001. Orthogonal garch. *In Mastering Risk. Financial Times–Prentice Hall: London* 2: 21–38

Andersen, Leif B. G. 2007. Efficient Simulation of the Heston Stochastic Volatility Model; SSRN 946405. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=946405 (accessed on 23 July 2021).

Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David Heath. 1999. Coherent measures of risk. *Mathematical Finance* 9: 203–28. [CrossRef]

Benth, Fred Espen. 2011. The stochastic volatility model of barndorff-nielsen and shephard in commodity markets. *Mathematical Finance* 21: 595–625. [CrossRef]

Black, Fischer. 1976. Studies of stock market volatility changes. Paper presented at 1976 American Statistical Association Bisiness and Economic Statistics Section, Washington, DC, USA, pp. 177–81. Available online: https://www.scirp.org/(S(i43dyn45teexjx455qlt3d2q))/reference/ReferencesPapers.aspx?ReferenceID=2030459 (accessed on 23 July 2021)

Campell, John Y., Andrew W. Lo, and A. Craig MacKinlay. 1997. *The Econometrics of Financial Markets*. Princeton: Princeton University Press

Carr, Peter, and Dilip Madan. 1999. Option valuation using the fast fourier transform. *Journal of Computational Finance* 2: 61–73. [CrossRef]

Cheng, Peng, and Olivier Scaillet. 2007. Linear-quadratic jump-diffusion modeling. *Mathematical Finance* 17: 575–98. [CrossRef]

Cheng, Yuyang, Marcos Escobar-Anel, and Zhenxian Gong. 2019. Generalized mean-reverting 4/2 factor model. *Journal of Risk and Financial Management* 12: 159. [CrossRef]

Christoffersen, Peter, Steven Heston, and Kris Jacobs. 2009. The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well. *Management Science* 55: 1914–32. [CrossRef]

Cui, Zhenyu, Justin Lars Kirkby, and Duy Nguyen. 2021 Efficient simulation of generalized sabr and stochastic local volatility models based on markov chain approximations. *European Journal of Operational Research* 290: 1046–62. [CrossRef]

Da Fonseca, José, Martino Grasselli, and Claudio Tebaldi. 2007. Option pricing when correlations are stochastic: An analytical framework. *Review of Derivatives Research* 10: 151–80. [CrossRef]

De Col, Alvise, Alessandro Gnoatto, and Martino Grasselli. 2013. Smiles all around: Fx joint calibration in a multi-heston model. *Journal of Banking & Finance* 37: 3799–818.

DeMiguel, Victor, Lorenzo Garlappi, and Raman Uppal. 2009. Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *The Review of Financial Studies* 22: 1915–53. [CrossRef]

Escobar, Marcos. 2018. A stochastic volatility factor model of heston type. statistical properties and estimation. *Stochastics* 90: 172–99. [CrossRef]

Escobar, Marcos, Barbara Götz, Luis Seco, and Rudi Zagst. 2010. Pricing a cdo on stochastically correlated underlyings. *Quantitative Finance* 10: 265–77. [CrossRef]

Escobar, Marcos, and Christoph Gschnaidtner. 2018. A multivariate stochastic volatility model with applications in the foreign exchange market. *Review of Derivatives Research* 21: 1–43. [CrossRef]

Escobar, Marcos, and Pablo Olivares. 2013. Pricing of mountain range derivatives under a principal component stochastic volatility model. *Applied Stochastic Models in Business and Industry* 29: 31–44. [CrossRef]

Escobar-Anel, Marcos, and Zhenxian Gong. 2020. The mean-reverting 4/2 stochastic volatility model: Properties and financial applications. *Applied Stochastic Models in Business and Industry* 36: 836–856 [CrossRef]

Escobar-Anel, Marcos, and Harold A. Moreno-Franco. 2019. Dynamic portfolio strategies under a fully correlated jump-diffusion process. *Annals of Finance* 15: 421–53. [CrossRef]

Gouriéroux, Christian. 2006. Continuous time wishart process for stochastic risk. *Econometric Reviews* 25: 177–217. [CrossRef]

Grasselli, Martino. 2016. The 4/2 stochastic volatility model: A unified approach for the heston and the 3/2 model. In *Mathematical Finance*. Hoboken: Wiley Online Library

Heston, Steven L. 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies* 6: 327–43. [CrossRef]

Kirkby, J. Lars, and Duy Nguyen. 2020. Efficient asian option pricing under regime switching jump diffusions and stochastic volatility models. *Annals of Finance* 16: 307–51. [CrossRef]

Langetieg, Terence C. 1980. A multivariate model of the term structure. *The Journal of Finance* 35: 71–97.

Larsen, Linda Sandris. 2010. Optimal investment strategies in an international economy with stochastic interest rates. *International Review of Economics & Finance* 19: 145–65.

Lin, Wei, Shenghong Li, Xingguo Luo, and Shane Chern. 2017. Consistent pricing of vix and equity derivatives with the 4/2 stochastic volatility plus jumps model. *Journal of Mathematical Analysis and Applications* 447: 778–97. [CrossRef]

Luo, Xingguo, and Jin E. Zhang. 2012. The term structure of vix. *Journal of Futures Markets* 32: 1092–123. [CrossRef]

McNeil, Alexander, Rudiger Frey, and Paul Embrechts. 2005. *Quantitative Risk Management: Concepts, Techniques, and Tools*. Princeton Series in Finance. Princeton: Princeton University Press.

Merton, Robert C. 1975. Optimum consumption and portfolio rules in a continuous-time model. In *Stochastic Optimization Models in Finance*. Amsterdam: Elsevier, pp. 621–61.

Muhle-Karbe, Johannes, Oliver Pfaffel, and Robert Stelzer. 2012. Option pricing in multivariate stochastic volatility models of ou type. *SIAM Journal on Financial Mathematics* 3: 66–94. [CrossRef]

Platen, Eckhard. 1997. A non-linear stochastic volatility model. In *Financial Mathematics Research Report No. FMRR005-97*. Canberra: Center for Financial Mathematics, Australian National University.

Stuart, Alan, Steven Arnold, J. Keith Ord, Anthony O'Hagan, and Jonathan Forster. 1994. *Kendall's Advanced Theory of Statistics*. Hoboken: Wiley.

Zhang, Jin E., and Yingzi Zhu. 2006. Vix futures. *Journal of Futures Markets* 26: 521–31. [CrossRef]