Margrabe Formulas for a Simple Bivariate Exponential Variance-Gamma Price Process (II) Statistical Estimation and Application

Werner Hürlimann
Wolters Kluwer Financial Services,
Seefeldstrasse 69,
CH-8008 Zürich, Switzerland.
werner.huerlimann@wolterskluwer.com

Received: 20-06-2013 Revised: 21-07-2013 Accepted: 24-07-2013

Abstract: A multivariate moment method for the simple multivariate variance-gamma distribution is considered. It takes into account the star products of both the coskewness and cokurtosis tensors. The model parameters depend upon the solution of a sextic equation, and the covariance matrix is functionally dependent upon coskewness and cokurtosis. The method enables simultaneous estimation of the parameters given sample estimates of the mean vector, coskewness vector and cokurtosis matrix. Application to the estimation of the bivariate variance-gamma model for the Standard & Poors 500 and NASDAQ 100 stock market indices is undertaken. The statistical fitting results are used to compare the original Margrabe formula with a variance-gamma exchange option pricing formula derived through application of the state-price deflator approach.

Keywords: multivariate variance-gamma, coskewness, cokurtosis, exchange option, stock market.

1. INTRODUCTION

In the first part of the present study, a simple multivariate exponential variance-gamma price process has been considered as alternative to the usual multivariate exponential Gaussian process. Based on it a state-price deflator has been constructed and applied to derive pricing formulas for the exchange option with a bivariate exponential variance-gamma deflated price process. It has been observed that the simple model is easy to work with but has some serious drawbacks. For example, linear correlation cannot be fitted once the margins are fixed. Moreover, great difficulty has been encountered with it in the joint calibration to option prices on the margins. To eliminate some of these disadvantages, it is possible to design an alternative estimation method. This is the main purpose of the present follow-up. In particular, statistical estimation of the multivariate variance-gamma model is discussed, and an application to stock market indices is presented.

An overview of our approach and the obtained main results follows. Based on the first four moments of the common gamma subordinator and the Theorem of Isserlis (on the third and fourth order multivariate normal moments), we obtain in Theorem 2.1 the explicit expressions for the mean, covariance, coskewness and cokurtosis parameters of the simple multivariate variance-gamma model. Then, a multivariate moment method is designed, which takes into account the star products of both the coskewness and cokurtosis tensors. The method is a one-parameter extension of the novel approach presented in [1]. Two features of it might be mentioned:

(i) The covariance matrix of the multivariate variance-gamma distribution is functionally dependent upon coskewness and cokurtosis.
The considered multivariate moment method enables simultaneous estimation of the parameters given sample estimates of the mean vector, coskewness vector and cokurtosis matrix. To demonstrate the practicability of the new approach, a real-life case study is presented in Section 3. It concerns the statistical estimation of the eight parameter bivariate variance-gamma model for the Standard & Poors 500 and NASDAQ 100 stock market indices. The model is successfully fitted to seven bivariate daily data sets over different time periods. The goodness-of-fit of the margins is optimized and compared with the goodness-of-fit of the bivariate normal, which turns out to be rather poor. The results are used in Section 4 to compare the original Margrabe formula with the variance-gamma exchange option pricing formula derived in [2]. Section 5 is devoted to further discussion and conclusions. The Appendix derives in a simple way an expression for the variance-gamma density that turns out to be more symmetrical than the original density formula by [3]. It is used to compute the Anderson-Darling goodness-of-fit statistics in Section 3.

2. A MULTIVARIATE MOMENT METHOD

A random vector \( X = (X_1, X_2, \ldots, X_n) \) has a simple multivariate variance-gamma (VG) distribution with parameter vectors \( \xi = (\xi_i), \theta = (\theta_i), i = 1, \ldots, n \), parameter matrix \( \Sigma = (\sigma_{ij}), 1 \leq i, j \leq n \), and parameter \( \nu \), for short \( X \sim VG(\xi, \theta, \Sigma, \nu) \), if its cumulant generating function (cgf) is given by

\[
C_u(u) = \xi^T u - \nu^{-1} \cdot \ln\left[ 1 - \nu \cdot (\theta^T u - \frac{1}{2} u^T \Sigma u) \right],
\]

(2.1)

for all values of \( u = (u_1, u_2, \ldots, u_n) \) for which the expression (2.1) exists. The special case \( \nu = 1 \) is the three-parameter class of the asymmetric Laplace \( AL(\xi, \theta, \Sigma) \), for which a moment method has been designed by [4]. The random vector \( X \) satisfies the stochastic representation

\[
X = \xi + \theta \cdot G + \sqrt{G} \cdot Y,
\]

(2.2)

where \( G \sim \Gamma(1/\nu, 1/\nu) \) is a gamma random variable with mean rate one and variance \( \nu \), \( Y \sim N(0, \Sigma) \) is a multivariate normal with vanishing mean vector, and \( Y, G \) are independent. In a first step, we determine the mean vector \( E[X] = (E[X_1], E[X_2], \ldots, E[X_n]) \), abbreviated \( \mu = (\mu_1, \ldots, \mu_n) \), and the matrix of \( k \)-th order central moments, denoted by \( \bar{M}_k[X], k = 2, 3, 4 \). For \( k = 2 \) the \( nxn \) matrix \( \bar{M}_2[X] = D[X] = (V_{ij}), 1 \leq i, j \leq n \), is the covariance matrix with elements \( V_{ij} = E[(X_i - \mu_i)(X_j - \mu_j)] \). The \( nxn^2 \) matrix \( \bar{M}_3[X] = (S_{ijk}), 1 \leq i, j, k \leq n \), consists of the coskewness elements \( S_{ijk} = E[(X_i - \mu_i)(X_j - \mu_j)(X_k - \mu_k)] \), and the \( nxn^3 \) matrix \( \bar{M}_4[X] = (K_{ijkl}), 1 \leq i, j, k, \ell \leq n \), consists of the cokurtosis elements \( K_{ijkl} = E[(X_i - \mu_i)(X_j - \mu_j)(X_k - \mu_k)(X_\ell - \mu_\ell)] \). In general, one has the relationships

\[
\begin{align*}
S_{ijk} &= E[X_i X_j X_k] - (\mu_i V_{jk} + \mu_j V_{ik} + \mu_k V_{ij}) - \mu_i \mu_j \mu_k, \\
K_{ijkl} &= E[X_i X_j X_k X_\ell] - (\mu_i S_{jk} + \mu_j S_{ik} + \mu_k S_{ij} + \mu_\ell S_{ij}) - (\mu_i \mu_j V_{kl} + \mu_i \mu_k V_{jl} + \mu_j \mu_k V_{il} + \mu_i \mu_\ell V_{jk} + \mu_j \mu_\ell V_{ik} + \mu_k \mu_\ell V_{ij}) - \mu_i \mu_j \mu_k \mu_\ell.
\end{align*}
\]

(2.3)

The following result generalizes Proposition 2.1 in [1] (special case \( \nu = 1 \)).

**Theorem 2.1 (Moments of the simple multivariate VG)** The mean, covariance, coskewness and cokurtosis parameters of the multivariate VG random vector \( X \sim VG(\xi, \theta, \Sigma, \nu) \) are given by
\[ \mu_i = \xi_i + \theta_i, \quad V_{ij} = \nu \cdot \theta_i \theta_j + \sigma_{ij}, \quad 1 \leq i, j \leq n, \]
\[ S_{ik} = 2\nu^2 \cdot \theta_i \theta_j \theta_k + \nu \cdot (\theta_i \sigma_{jk} + \theta_j \sigma_{ik} + \theta_k \sigma_{ij}), \quad 1 \leq i, j, k \leq n, \]
\[ K_{ijkl} = 3\nu^2 \cdot (1 + 2\nu) \cdot \theta_i \theta_j \theta_k \theta_l + \nu(1 + 2\nu) \cdot (\theta_i \theta_j \sigma_{kl} + \theta_j \theta_k \sigma_{il} + \theta_k \theta_i \sigma_{jl} + \theta_i \theta_k \sigma_{jl} + \theta_j \theta_i \sigma_{kl} + \theta_k \theta_j \sigma_{il}) + (1 + \nu) \cdot (\sigma_{ij} \sigma_{kl} + \sigma_{ik} \sigma_{jl} + \sigma_{il} \sigma_{jk}), \quad 1 \leq i, j, k, \ell \leq n. \]

**Proof:** One uses the representation (2.2). The expression for the mean vector is immediate. For the central moments it suffices to consider the case \( \xi = 0 \). With (2.2) the vector components of \( X \) satisfy the representation \( X_i = \theta G + \sqrt{G} \cdot Y_i, i = 1, \ldots, n \), where \( G \sim \Gamma(1/\nu, 1/\nu) \) is independent of \( Y_i \sim N(0, \sigma_{ii}) \). The moment relationships

\[ E[G] = 1, \quad E[G^2] = 1 + \nu^2, \quad E[G^3] = 1 + 3\nu + 2\nu^2, \quad E[G^4] = 1 + 6\nu + 1\nu^2 + 6\nu^3, \]

will be used repeatedly without further mention. One has

\[ E[X, X] = E[\theta_i \theta_j G^2 + (\theta_i Y_k + \theta_j Y_k)G^{3/2} + Y_k G] = \theta_i \theta_j (1 + \nu) + \sigma_{ij}, \]

which implies the expression for the covariance. Similarly, one has

\[ E[X, X, X] = E[\theta_i \theta_j \theta_k G^3 + (\theta_i Y_k + \theta_j Y_k + \theta_i \theta_k Y_j + \theta_j \theta_k Y_i)G^{5/2} + (\theta_i Y_k + \theta_j Y_k + \theta_i \theta_j Y_k)G^2 + Y_k Y_k G^{3/2}] + (\theta_i Y_k + \theta_j Y_k + \theta_i \theta_k Y_j + \theta_j \theta_k Y_i)G^{7/2}. \]

With the fact that \( E[Y_k Y_k] = 0 \) (theorem of Isserlis) one sees that

\[ E[X, X, X] = (1 + 3\nu + 2\nu^2) \cdot \theta_i \theta_j \theta_k + (1 + \nu) \cdot (\theta_i \sigma_{jk} + \theta_j \sigma_{ik} + \theta_k \sigma_{ij}). \]

Insert this and the fact that \( \mu_i = \theta_i, V_{ij} = \nu \theta_i \theta_j + \sigma_{ij} \) into the first relation of (2.3) to obtain the coskewness formula in (2.4). Proceeding in the same way, one shows that

\[ E[X, X, X] = E[\theta_i \theta_j \theta_k \theta_l G^4 + (\theta_i \theta_j \theta_k Y_l + \theta_i \theta_j \theta_l Y_k + \theta_i \theta_k \theta_l Y_j + \theta_j \theta_k \theta_l Y_i)G^{7/2} + (\theta_i \theta_j Y_k + \theta_j \theta_k Y_j + \theta_i \theta_j Y_k + \theta_j \theta_k Y_k)G^3 + (\theta_i Y_k + \theta_j Y_k + \theta_i \theta_j Y_k + \theta_j \theta_i Y_k)G^{3/2} + Y_k Y_k G^2]. \]

Since \( E[Y_k Y_k] = \sigma_{ij} \sigma_{kl} + \sigma_{ik} \sigma_{jl} + \sigma_{il} \sigma_{jk} \) (theorem of Isserlis) one gets

\[ E[X, X, X] = (1 + 6\nu + 1\nu^2 + 6\nu^3) \cdot \theta_i \theta_j \theta_k \theta_l + (1 + 3\nu + 2\nu^2) \cdot (\theta_i \theta_j \sigma_{kl} + \theta_j \theta_i \sigma_{jl} + \theta_i \theta_k \sigma_{jl} + \theta_j \theta_k \sigma_{il} + \theta_i \theta_k \sigma_{il} + \theta_j \theta_k \sigma_{il}) + (1 + \nu) \cdot (\sigma_{ij} \sigma_{kl} + \sigma_{ik} \sigma_{jl} + \sigma_{il} \sigma_{jk}). \]

Insert this, \( \mu_i = \theta_i, V_{ij} = \nu \theta_i \theta_j + \sigma_{ij} \), and the coskewness relation in (2.4), into the second part of (2.3) to obtain after rearrangement and counting of equal terms the cokurtosis formula. \( \Box \)

We are ready for the generalization of the moment method in [1], Section 3. For any fixed \( \nu > 0 \) and given the mean parameters \( \mu_i \), the coskewness and cokurtosis parameters \( S_{ik} \) and \( K_{ijkl} \), we determine the remaining parameters \( V_{ij}, (\sigma_{ij}), (\xi_i), (\theta_i) \) in terms of them. In particular, it is shown that the covariance matrix \( V_{ij} \) of the multivariate VG distribution functionally depends upon coskewness and cokurtosis. First of all, given the mean \( \mu \) and assuming \( \theta \) has been determined, it is clear that \( \xi \) is obtained from the mean vector equation as \( \xi = \mu - \theta \). Similarly, once \( (V_{ij}), (\theta_i) \) have been determined, the parameter matrix \( \Sigma = (\sigma_{ij}) \) is obtained from the covariance equation as \( \sigma_{ij} = V_{ij} - \nu \theta_i \theta_j \). Next let us examine the
Margrabe Formulas for a Simple Bivariate Exponential Variance-Gamma Price Process (II) Statistical Estimation and Application

coskewness equations. For this, consider the coskewness vector \( \overline{S}(X) = (S_1, \ldots, S_n) \) derived from the star product \( \overline{M}_3[X] \) such that

\[
S_i = \sum_{j,k=1}^{n} S_{ijk}, \quad i = 1, 2, \ldots, n. \tag{2.5}
\]

The following short hand notation for sums of covariances and parameters is used:

\[
V_i = \sum_{j=1}^{n} V_{ij} = \sum_{j=1}^{n} V_{ji}, \quad V = \sum_{i,j=1}^{n} V_{ij}, \quad M = \sum_{i=1}^{n} \theta_i. \tag{2.6}
\]

The evaluation of (2.5) based on the coskewness formula in (2.4) yields the relationships

\[
\nu \cdot (V - \nu \cdot M^2) \cdot \theta_i + 2\nu \cdot M V_i = S_i, \quad i = 1, 2, \ldots, n. \tag{2.7}
\]

Set further \( S = \sum_{i=1}^{n} S_i \) and add the equations in (2.7) to get the equation in \( (M, V) \):

\[
3\nu \cdot VM - \nu^2 \cdot M^3 - S = 0. \tag{2.8}
\]

Consider now the cokurtosis equations and define the cokurtosis matrix \( K(X) = (K_{ij}) \) using the star product \( K(X) = \sum_{i=1}^{n} \overline{M}_4 [X] \) such that

\[
K_{ij} = \sum_{k=1}^{n} K_{ijk}, \quad i, j = 1, \ldots, n. \tag{2.9}
\]

A calculation of (2.9) based on the last equations in (2.4) yields the relationships

\[
\nu^2 \cdot (V - \nu \cdot M^2) \cdot \theta_i \theta_j + 2\nu^2 \cdot (V \theta_j + \theta_j V_j) \cdot M
\]

\[
+ 2(1 + \nu) \cdot V_i V_j + \{\nu^2 \cdot M^2 + (1 + \nu) \cdot V\} \cdot V_j = K_{ij}, \quad i, j = 1, 2, \ldots, n. \tag{2.10}
\]

Further, summing (2.10) with the short hand notation \( K_i = \sum_{j=1}^{n} K_{ij} = \sum_{j=1}^{n} K_{ji}, i = 1, 2, \ldots, n \) one gets

\[
3\nu^2 \cdot (V - \nu \cdot M^2) \cdot M \theta_i + 3\nu^2 \cdot M^2 + (1 + \nu) \cdot V \cdot V_j = K_{ij}, \quad i = 1, \ldots, n. \tag{2.11}
\]

With \( K = \sum_{i=1}^{n} K_i \) one obtains through addition of (2.11) a further equation in \( (M, V) \), namely

\[
3\nu^3 \cdot M^4 - 6\nu^2 \cdot V M^2 - 3(1 + \nu) \cdot V^2 + K = 0. \tag{2.12}
\]

The resulting system of non-linear equations (2.7), (2.8), (2.10), (2.11), (2.12) in the unknowns \( (\theta_i, M, V_j, V_i, V) \) can be solved by applying a three-stage procedure as follows.

**Step 1:** solve the equations (2.8) and (2.12) for the parameters \( (M, V) \)

From (2.8) one gets

\[
V = \frac{\nu^2 \cdot M^3 + S}{3\nu^2 \cdot M}. \tag{2.13}
\]

Insert this expression into (2.12) and multiply with \( 9\nu^2 \cdot M^2 \) to see that \( M \) satisfies the following sextic equation in the parameters \( (S, K) \):

\[
(2\nu - 1)\nu^4 \cdot M^6 - 2(1 + 4\nu) \cdot \nu^2 \cdot S M^3 + 3\nu^2 \cdot K M^2 - (1 + \nu) \cdot S^2 = 0. \tag{2.14}
\]
Step 2: solve the equations (2.7) and (2.11) for the parameters \((\theta_i, V_i), i=1,2,...,n\)

From (2.7) one gets

\[
\theta_i = \frac{S_i \cdot (2\nu \cdot M \hat{V}_i)}{\nu \cdot (V - \nu \cdot M^2)}, \quad i=1,2,...,n. \tag{2.15}
\]

Insert this into (2.11) to see that \(V_i\) is function of the parameters \((M, V, S, K)\):

\[
V_i = \frac{1}{3} \frac{K_i - 3\nu \cdot MS_i}{(1 + \nu \cdot V - \nu ^2 \cdot M^2)}, \quad i=1,2,...,n. \tag{2.16}
\]

It follows that \(\theta_i\) is a function of the same parameters, namely

\[
\theta_i = \frac{S_i \cdot (2\nu \cdot M \hat{V}_i)}{\nu \cdot (V - \nu \cdot M^2)} - \frac{2}{3} \frac{(K_i - 3\nu \cdot MS_i)M}{(1 + \nu \cdot V - \nu ^2 \cdot M^2)}, \quad i=1,2,...,n. \tag{2.17}
\]

where one must assume that

\[
(V - \nu \cdot M^2)((1 + \nu) \cdot V - \nu ^2 \cdot M^2) \neq 0. \tag{2.18}\]

Step 3: the unknowns \((V_i)\) are obtained from the equation (2.10), where one must verify that the covariance matrix \((V_i)\) and its associated correlation matrix are positive semi-definite.

The described moment method is useful for parameter estimation. Given a sample \((x_1, x_2, ..., x_N)\) of size \(N\), where each \(x_i\) is an observation of the random vector \(X = (X_1, X_2, ..., X_n)\), one considers the following sample estimates of the coskewness vector and cokurtosis matrix:

\[
\hat{S}(X) = (\hat{S}_1, ..., \hat{S}_n) = N^{-1} \cdot 1_{n \times n} \cdot \sum_{i=1}^{n} (x_i x_i^T \otimes x_i),
\]

\[
\hat{K}(X) = (\hat{K}_i) = N^{-1} \cdot 1_{n \times n} \cdot \sum_{i=1}^{n} (x_i x_i^T \otimes x_i x_i^T). \tag{2.19}
\]

Samples estimates of the quantities \(S, K, K, \) are obtained through summation as

\[
\hat{S} = \sum_{i=1}^{n} \hat{S}_i, \quad \hat{K}_i = \sum_{j=1}^{n} \hat{K}_{ij} = \sum_{j=1}^{n} \hat{K}_{ji}, i=1,2,...,n, \quad \hat{K} = \sum_{i=1}^{n} \hat{K}_i. \tag{2.20}
\]

Inserting these estimates into the derived formulas, one obtains for any fixed \(\nu > 0\) estimates of the multivariate VG parameters in terms of the sample mean vector, coskewness vector and cokurtosis matrix. The next Section illustrates a real-world application of this procedure.

### 3. Estimation of Bivariate Logarithmic Returns

We consider now two stock market indices for which all the mean, coskewness and cokurtosis quantities can be estimated. Return observations stem from the following seven different pairs of bivariate data from the Standard & Poors 500 (SP500) and the NASDAQ 100 (NDX) data sets:

- **SP500/NDX/3Y:** 754 daily closing prices over 3 years from 04.01.2010 to 31.12.2012
- **SP500/NDX/5Y:** 1259 daily closing prices over 5 years from 02.01.2008 to 31.12.2012
- **SP500/NDX/10Y:** 2516 daily closing prices over 10 years from 02.01.2003 to 31.12.2012
- **SP500/NDX/15Y:** 3773 daily closing prices over 15 years from 02.01.1998 to 31.12.2012
- **SP500/NDX/20Y:** 5093 daily closing prices over 20 years from 04.01.1993 to 31.12.2012
- **SP500/NDX/25Y:** 6302 daily closing prices over 25 years from 04.01.1988 to 31.12.2012

*International Journal of Scientific and Innovative Mathematical Research (IJSIMR)*
These data sets are typical as they contain short to medium high volatile periods (recent 3 and 5 years), moderate long term periods (10 and 15 years), and long term periods (20,25 and 27 years). The last data set has been included because it contains the highest and lowest daily changes observed so far (drop in 22.9% and 16.3% for SP500 respectively NDX on 19.10.1987, increase of 17.2% for NDX on 03.01.2001).

The following Table 3.1 lists the required sample moment estimates for the bivariate logarithmic returns obtained from each of these combinations.

Table 3.1 Sample moment estimates of bivariate log-returns from two stock market indices

<table>
<thead>
<tr>
<th>unit</th>
<th>10^-4</th>
<th>10^-6</th>
<th>10^-6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\mu_1)</td>
<td>(\mu_2)</td>
<td>(S_1)</td>
</tr>
<tr>
<td>3Y</td>
<td>3.05639</td>
<td>4.56635</td>
<td>-2.53599</td>
</tr>
<tr>
<td>5Y</td>
<td>-0.11603</td>
<td>2.07454</td>
<td>-3.14410</td>
</tr>
<tr>
<td>10Y</td>
<td>1.79008</td>
<td>3.78059</td>
<td>-1.88119</td>
</tr>
<tr>
<td>15Y</td>
<td>1.00817</td>
<td>2.57285</td>
<td>-0.55355</td>
</tr>
<tr>
<td>20Y</td>
<td>2.35612</td>
<td>4.00594</td>
<td>-0.87526</td>
</tr>
<tr>
<td>25Y</td>
<td>2.72626</td>
<td>4.45475</td>
<td>-1.14828</td>
</tr>
</tbody>
</table>

Up to the 15Y and 20Y periods the coskewness vector has always negatively skewed components. The exception is the NDX. In the 15Y case one has also \(S = S_1 + S_2 > 0\). Over the longest period of 27Y the coskewness components take the highest negative values. Up to the shortest 3Y period the overall cokurtosis coefficient \(K = K_{11} + 2K_{12} + K_{22}\) exceeds 5 and is highest for the 5Y and 27Y periods. For specific fixed values of \(\nu > 0\) the bivariate VG is fitted to the data following the moment method described in Section 2.

Though the simple multivariate VG is easy to work with and has been theoretically justified (see [2], Section 2), it has been remarked that linear correlation cannot be fitted once the margins are fixed. However, the proposed moment method does not fit the margins separately, but provides an overall parsimonious fit of all its parameters regardless of the margins and the dependence structure. For this reason, it is important to discuss its goodness-of-fit capabilities as compared to the bivariate normal, which underlies Margrabe’s original pricing formula. In particular, it is important to analyze the goodness-of-fit of the estimated margins. To do so our goodness-of-fit (GoF) measure is based on statistics, which measure the difference between the empirical distribution functions \(F_n(x)\) and the estimated marginal distribution functions \(F(x)\). We use the Cramér-von Mises family of statistics defined by (e.g. [5]-[7])

\[
T = n \cdot \int_{-\infty}^{\infty} \left[ F_n(x) - F(x) \right]^2 w(x) \, dF(x)
\]

(3.1)

where \(w(x)\) is a suitable weighting function. If \(w(x) = 1 / [F(x)\hat{F}(x)]\) one gets the \(A^2\) Anderson-Darling [8] statistic. Consider the order statistics of the return data such that \(r_1 \leq r_2 \leq \ldots \leq r_n\) and let \(\hat{F}(r_i), i = 1,2,...,n\), be the estimated values of a marginal distribution function. Then one has

\[
A^2 = -n - \sum_{i=1}^{n} \frac{2i-1}{n} \cdot \ln \left( \frac{\hat{F}(r_i)}{\hat{F}(r_{n-i+1})} \right)
\]

(3.2)

The values \(\hat{F}(r_i)\) are obtained numerically by integration of the VG density. Under the so-called bilateral gamma representation this is expression (A.8) in the Appendix. The Anderson-Darling
statistic yields one of the most powerful test if the fitted distribution departs from the true distribution in the tails (e.g. [5]), and is recommended in this situation. Now, the observed sample return marginal data is skewed and has a much higher kurtosis than is allowed by a normal distribution, which indicates that the fit in the tails matters and justifies the use of the GoF statistics (3.2). Needless to say, the proposed moment method is only a starting point for improved GoF estimation methods. However, a more complex data analysis is beyond the scope of the present study. To weight the influence of the margins, we use the Euclidean distance to define an overall GoF measure as $\|A\|^2 = (A_i^2)^2 + (A_i^2)^2$, with $A_i^2, i=1,2$, the Anderson-Darling statistics of the margins. To calculate the GoF statistics of the variance-gamma margins, it is simpler to use the bilateral gamma representation of the margins defined by (e.g. [9], [10])

\[ X^{(k)} = \xi_k + \alpha_k^{-1} \cdot G_1 - \beta_k^{-1} \cdot G_2, \quad k = 1,2, \quad G_i \sim \Gamma(\gamma,1), \quad i = 1,2, \quad \text{independent}, \quad (3.3) \]

with the one-to-one parameter transformation

\[ \gamma = \nu^{-1}, \quad \alpha_k^{-1} = \frac{1}{2} \left( (v \theta_k)^2 + 2v \tau_k^2 + v \theta_k \right), \quad \beta_k^{-1} = \frac{1}{4} \left( (v \theta_k)^2 + 2v \tau_k^2 - v \theta_k \right). \quad (3.4) \]

where the changed parameter notations $\tau_k^2 = \sigma_{\theta_k}, k=1,2, \rho = \sigma_{\theta_k} / \tau_{\theta_k}$, are used. The estimated values of the marginal distribution functions in (3.2) are obtained numerically by integration of the expression (A.8) for the VG pdf. The estimated parameters and GoF statistics are summarized and compared in the Table 3.2.

As a first observation, one remarks that the bivariate normal yields a rather poor GoF in terms of (3.2) whatever the considered time period of analysis is. To limit the possible output, one determines by trial and error values of $\nu > 0$, for which the overall GoF measure $\|A\|$ is smallest (bold numbers in the Table). It is interesting to compare the “optimal” fits with the ones from a simple bivariate variance-gamma distribution, which determines the correlation coefficient between the margins (see the later Table 4.1). The estimated values of the marginal distribution functions in (3.2) are obtained numerically by integration of the expression (A.8) for the VG pdf. The estimated parameters and GoF statistics are summarized and compared in the Table 3.2.

4. **BLACK-SCHOLES VERSUS VARIANCE-GAMMA MARGRABE OPTION PRICES**

It is interesting to compare the classical closed-form Margrabe exchange option pricing formula with the corresponding formula for the exponential bivariate VG model. Recall that in the bivariate Black-Scholes model the future prices after one unit time are described by the equations

\[ S^{(k)} = S^{(k)}_0 \exp(\mu_k - \frac{1}{2} \sigma_k^2 + \sigma_k Z_k), \quad k = 1,2, \quad (4.1) \]

where the $Z_k$’s are correlated standard normal with correlation coefficient $\rho^{BS}$. In the special case $S^{(k)}_0 = 1, \mu_k = 0, k=1,2$, Margrabe’s formula reads

\[ M^{BS} = E[D(S^{(1)} - S^{(2)}),] = 2 \cdot \Phi(\frac{1}{2} \sigma) - 1, \quad \sigma = \sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho^{BS} \sigma_1 \sigma_2}. \quad (4.2) \]
If the future prices are exponential bivariate variance-gamma 

\( S^{(i)} = \exp(\xi + X^{(i)}) \), \( k = 1,2 \), 

with \( \xi + X \sim VG(\xi, \theta, \Sigma, \nu) \), one uses the VG deflator 

\[ D = \exp(-\alpha - \beta X) \), \( X \sim \text{bivariate normal} \), sets 

\[ \rho = \rho_{12}, \gamma = v^{-1}, \omega^2 = \tau_2^2 - \rho \tau_1 \tau_2, k = 1,2, \omega = \sqrt{\omega_1^2 + \omega_2^2} \] 
to obtain the pricing formula (see Hürlimann (2013a), Theorem 6.1)

\[ M^{VG} = E[D(S^{(1)} - S^{(2)})], = e^{\hat{\xi} - \alpha} \cdot \Psi(a_1, b_1, c_1, \gamma) - e^{\hat{\xi} - \alpha} \cdot \Psi(a_2, b_2, c, \gamma) \]

\[ \Psi(a, b, c, \gamma) = \int_0^\infty z^{v^{-1}} e^{-z} z^{\gamma-1} \Phi(z) e^{\frac{z^2}{2} c} / \sqrt{\pi} \frac{d}{dz}. \]

\[ a_1 = \gamma^{-1}((1 - \beta_2) \rho \tau_1 \tau_2 - \theta^T \beta + \frac{1}{2} \beta^T \Sigma \beta), \quad a_2 = \gamma^{-1}((\beta - 1) \rho \tau_1 \tau_2 - \theta^T \beta + \frac{1}{2} \beta^T \Sigma \beta), \]

\[ b_k = \sqrt{\gamma^{-1} \cdot (1 - \beta_k^- 1) \omega + (\beta - 1) \rho \tau_1 \tau_2 \omega^{-1}}, \quad k = 1,2, \quad c = \sqrt{\gamma^{-1} (\xi - \xi^2) \omega^{-1}}. \]

International Journal of Scientific and Innovative Mathematical Research (IJSIMR)
We have evaluated (4.3) for the seven pairs in Section 3 and compared the results with (4.2). We assume that (4.2) is calculated with either the sample covariance matrix, i.e. \( \rho_{BS} = \rho \) is the sample correlation coefficient, or the estimated covariance matrix, i.e. \( \rho_{BS} = \rho F \) is the fitted correlation coefficient). Since \( \alpha = r + C(X_{10}, X_{23}, \beta) \) (see [2], formula (4.23)), the formula (4.3), in contrast to (4.2), depends upon the risk-free rate of return. Without loss of generality it suffices to compare the case \( r = 0 \). Table 4.1 below lists the percentage signed relative deviation of the VG Margrabe formula from the original Margrabe formula in the Black-Scholes model.

One notes the following properties. The asymmetric Laplace Margrabe formula is always on the safe side, whether the sample or estimated covariance matrices are used in Black-Scholes model. Except for the 5Y and 10Y periods, the original Margrabe formula with sample covariance matrix underestimates the exchange option price evaluated with the estimated VG model. Using in both formulas the estimated covariance matrix, one observe that they diverge considerably over the smallest 3Y time period (>40% and >80% relative deviation) and partly for the asymmetric Laplace special case. Otherwise, the VG and the original Margrabe prices are reasonably close with changing signs depending on the period of analysis. Furthermore, in parallel to the diverging feature observed in Section 3 about the correlation coefficient, the VG option prices for periods up to 10Y remain stable in the interval \( (1.97 \times 10^{-3}, 2.38 \times 10^{-3}) \) while over the longer time periods from 15Y on these prices are much higher and vary in the interval \( (5.03 \times 10^{-3}, 5.67 \times 10^{-3}) \).

5. RESULTS, FURTHER DISCUSSION AND CONCLUSIONS

Our starting point has been the observation that the considered multivariate variance-gamma model is easy to work with but has some drawbacks. For example, linear correlation cannot be fitted once the margins are fixed. To circumvent this disadvantage, we have designed a multivariate moment method that does not fit the margins separately, but provides an overall parsimonious fit of all its parameters regardless of the margins and the dependence structure. The algorithmic structure of this novel statistical method is simple. Given the coskewness vector and the cokurtosis matrix, it follows a three-stage procedure. First, one solves the sextic equation (2.14) in the overall VG parameter \( M \) (sum of the marginal VG parameters \( \theta_i \)) and the parameter \( \nu \) of the gamma subordinator. In step 2, one determines the marginal VG and covariance parameters \( (\theta_i, V_i), i = 1,2,...,n \), using the equations (2.16)-(2.17). In the third step, one obtains the covariance matrix \( (V_{ij}) \), which necessarily must be positive semi-definite. The remaining model parameters are then straightforward functions of the obtained quantities. The derivation of this multivariate moment method uses the explicit expressions for the mean, covariance, coskewness and cokurtosis parameters of the multivariate VG model (Theorem 2.1) as well as the star products of the coskewness and cokurtosis tensors. It is important to note that the star product has also been applied in related statistical methods by [13]-[15].

We have illustrated the usefulness of this moment method with a case study, namely the statistical estimation of the eight parameter bivariate variance-gamma model for the Standard & Poors 500 and NASDAQ 100 stock market indices. The model is successfully fitted to seven bivariate daily data sets over different time periods, and the goodness-of-fit of the margins has been optimized. The fitting results have been used in Section 4 to compare the original Margrabe formula with the variance-gamma exchange option pricing formula derived in [2].

As a mode of conclusions let us give an outlook on some further possible studies in this area. First of all, there is a need for an improved goodness-of-fit estimation method in Section 3 that is not only restricted to the margins. Second, we expect applications of the state-price deflator approach to the pricing of more complex financial derivatives and insurance products with embedded options. Finally, it is possible to extend the present work to related multivariate non-Gaussian variance-mean mixture models that include multivariate normal inverse Gaussian and multivariate normal tempered stable distributions. The latter extension will be presented elsewhere.
### Margrabe Formulas for a Simple Bivariate Exponential Variance-Gamma Price Process (II) Statistical Estimation and Application

**Table 4.1** Comparison of bivariate BS and VG Margrabe option prices (case $r = 0$)

<table>
<thead>
<tr>
<th>unit</th>
<th>market uncertainty</th>
<th>correlation coefficient</th>
<th>Margrabe price</th>
<th>percentage signed relative deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SP500/NDX</td>
<td></td>
<td>MBS/S</td>
<td>MBS/F</td>
</tr>
<tr>
<td>3Y</td>
<td>1</td>
<td>0.96516</td>
<td>1.27244</td>
<td>2.41537</td>
</tr>
<tr>
<td></td>
<td>1.02</td>
<td>0.96516</td>
<td>1.26927</td>
<td>2.39017</td>
</tr>
<tr>
<td></td>
<td><strong>1.03</strong></td>
<td><strong>0.94184</strong></td>
<td><strong>1.67889</strong></td>
<td><strong>2.16769</strong></td>
</tr>
<tr>
<td></td>
<td>1.04</td>
<td>0.96516</td>
<td>1.26613</td>
<td>2.36570</td>
</tr>
<tr>
<td>5Y</td>
<td>1</td>
<td>0.95787</td>
<td>2.19706</td>
<td>2.55543</td>
</tr>
<tr>
<td></td>
<td>1.55</td>
<td>0.95790</td>
<td>2.06681</td>
<td>2.11297</td>
</tr>
<tr>
<td></td>
<td><strong>1.60</strong></td>
<td><strong>0.94078</strong></td>
<td><strong>2.32237</strong></td>
<td><strong>2.05674</strong></td>
</tr>
<tr>
<td></td>
<td>1.65</td>
<td>0.95790</td>
<td>2.04690</td>
<td>2.05433</td>
</tr>
<tr>
<td>10Y</td>
<td>1</td>
<td>0.94556</td>
<td>2.13992</td>
<td>2.42629</td>
</tr>
<tr>
<td></td>
<td>1.60</td>
<td>0.94560</td>
<td>2.00335</td>
<td>1.99342</td>
</tr>
<tr>
<td></td>
<td><strong>1.65</strong></td>
<td><strong>0.91693</strong></td>
<td><strong>2.34718</strong></td>
<td><strong>1.99378</strong></td>
</tr>
<tr>
<td></td>
<td>1.70</td>
<td>0.94560</td>
<td>1.98443</td>
<td>1.94106</td>
</tr>
<tr>
<td>15Y</td>
<td>1</td>
<td>0.69439</td>
<td>6.22402</td>
<td>6.25483</td>
</tr>
<tr>
<td></td>
<td>1.30</td>
<td>0.69450</td>
<td>6.00922</td>
<td>5.74297</td>
</tr>
<tr>
<td></td>
<td><strong>1.35</strong></td>
<td><strong>0.83383</strong></td>
<td><strong>4.73601</strong></td>
<td><strong>5.97683</strong></td>
</tr>
<tr>
<td></td>
<td>1.40</td>
<td>0.69454</td>
<td>5.94528</td>
<td>5.59845</td>
</tr>
<tr>
<td>20Y</td>
<td>1</td>
<td>0.69609</td>
<td>5.83313</td>
<td>5.90215</td>
</tr>
<tr>
<td></td>
<td>1.40</td>
<td>0.69622</td>
<td>5.57213</td>
<td>5.27523</td>
</tr>
<tr>
<td></td>
<td><strong>1.45</strong></td>
<td><strong>0.82110</strong></td>
<td><strong>4.52216</strong></td>
<td><strong>5.54335</strong></td>
</tr>
<tr>
<td></td>
<td>1.50</td>
<td>0.69625</td>
<td>5.51530</td>
<td>5.14709</td>
</tr>
<tr>
<td>25Y</td>
<td>1</td>
<td>0.69694</td>
<td>5.54498</td>
<td>5.71098</td>
</tr>
<tr>
<td></td>
<td>1.40</td>
<td>0.69707</td>
<td>5.29688</td>
<td>5.09090</td>
</tr>
<tr>
<td></td>
<td><strong>1.45</strong></td>
<td><strong>0.81890</strong></td>
<td><strong>4.22923</strong></td>
<td><strong>5.26952</strong></td>
</tr>
<tr>
<td></td>
<td>1.50</td>
<td>0.69711</td>
<td>5.24285</td>
<td>4.96479</td>
</tr>
<tr>
<td>27Y</td>
<td>1</td>
<td>0.69032</td>
<td>6.00787</td>
<td>7.34234</td>
</tr>
<tr>
<td></td>
<td>1.95</td>
<td>0.69083</td>
<td>5.44521</td>
<td>5.58410</td>
</tr>
<tr>
<td></td>
<td><strong>2.00</strong></td>
<td><strong>0.81446</strong></td>
<td><strong>4.19757</strong></td>
<td><strong>5.42208</strong></td>
</tr>
<tr>
<td></td>
<td>2.05</td>
<td>0.69085</td>
<td>5.39942</td>
<td>5.46314</td>
</tr>
</tbody>
</table>

**APPENDIX:** Special function representation of the variance-gamma density

A five parameter bilateral gamma (BG) random variable is defined by

$$X = \xi + \alpha^{-1} \cdot G_1 - \beta^{-1} \cdot G_2 \sim BG(\xi, \gamma, \alpha, \delta, \beta), \gamma, \alpha, \delta, \beta > 0, -\infty < \xi < \infty,$$

with independent $G_1 \sim \Gamma(\gamma, 1), G_2 \sim \Gamma(\delta, 1)$ (standardized gamma’s with scale parameter 1). It suffices to restrict the attention to the BG with vanishing location $\xi = 0$. The BG pdf, denoted by $f(x) = f(x; \gamma, \alpha, \delta, \beta)$, is the convolution $f(x) = (f_1 * f_2)(x)$ of the two gamma pdf’s:

$$f_1(x) = \Gamma(\gamma)^{-1} \alpha^{-1} x^{\gamma-1} e^{-\alpha x} \cdot 1[x \geq 0], \quad f_2(x) = \Gamma(\delta)^{-1} \beta^{-1} x^{\delta-1} e^{-\beta |x|} \cdot 1[x \leq 0].$$  \hspace{1cm} (A.1)

The following “generalized gamma function” representation seems new. It is equivalent to the representation (A.6) below in terms of the confluent hyper-geometric function of the 2nd kind.

**Theorem A.1** (Generalized gamma function representation) The probability density function of the bilateral gamma $BG(\xi = 0, \gamma, \alpha, \delta, \beta)$ is given by

International Journal of Scientific and Innovative Mathematical Research (IJSIMR)
with the generalized gamma function
\[ \Gamma(a, b, x) = \int_0^\infty (1 + x^{-1}t)^{b-1} e^{-t} dt. \] (A.3)

**Proof:** Using the symmetry relation \( f(x; \gamma, \alpha, \delta, \beta) = f(-x; \delta, \beta, \gamma, \alpha) \) it suffices to consider the case \( x \in (0, \infty) \). Through elementary integration (change of variables \( y = -tx \)) one obtains

\[
\begin{align*}
  f(x) &= (f_1 \ast f_2)(x) = \int_{-\infty}^0 f_1(x-y)f_2(y)dy = \Gamma(\gamma)^{-1} \Gamma(\delta)^{-1} \alpha^\gamma \beta^\delta e^{-\alpha x} \cdot I(x), \\
  I(x) &= \int_{-\infty}^0 (x-y)^{y^{-1}} (-y)^{\delta^{-1}} e^{(\alpha+\beta)y}dy = \frac{\gamma}{\beta} x^{\gamma+\delta-1}(1+t)^{\gamma-1} e^{-(\alpha+\beta)t} dt,
\end{align*}
\]

The transformation \( t = c(x)^{-1}u \) with \( c(x) = (\alpha + \beta) x \) yields further

\[
I(x) = x^{\gamma+\delta-1} c(x)^{-\delta} \int_0^\infty (1+c(x)^{-1}u)^{\gamma-1} u^{\delta-1} e^{-u} du = x^{\gamma+\delta-1}(\alpha + \beta)^{-\delta} \cdot \Gamma(\delta, \gamma, c(x)).
\]

Insert into the first integral expression for \( f(x) \) to get (A.2). \( \square \)

In virtue of the limiting property \( \lim_{x \to \infty} \Gamma(a, b, x) = \int_0^\infty e^{-t} dt = \Gamma(a) \) the naming of the integral (A.3) is justified. Furthermore, one has also trivially \( \Gamma(a, 1, x) = \Gamma(a) \). Another justification arises from the fact that when \( \alpha \to \infty \) or \( \beta \to \infty \) the pdf converges to a left- and right-tail gamma pdf respectively, as should be. Moreover, a close look at the confluent hyper-geometric function of the 2nd kind, introduced by [11] and also called Tricomi function, shows the relationship

\[ \Gamma(a, b, x) = \Gamma(a)x^a U(a, a + b, x), \] (A.4)

where the Tricomi function is defined by (e.g. [12], 48:3:6 and 48:3.7)

\[ U(a, b, x) = \left[ \frac{1}{\beta} \right] t^{a-1} (1+t)^{1-ba^{-1}} e^{-u} dt = \Gamma(a)^{-1} \left[ \frac{1}{\beta} \right] t^{a-1} (1-t)^{-b} e^{-a(1-t)^{-1}} dt. \] (A.5)

The generalized gamma function is a transformed Tricomi function and (A.2) rewrites as

\[
\begin{align*}
  f(x) &= \Gamma(\gamma)^{-1} \alpha^\gamma x^{\gamma-1} e^{-ax} \cdot U(\delta, \gamma + \alpha, (\alpha + \beta)x), \quad x > 0, \\
  f(x) &= \Gamma(\delta)^{-1} \beta^\delta |x|^{\delta-1} e^{-\beta|x|} \cdot U(\gamma, \gamma + \delta, (\alpha + \beta)|x|), \quad x < 0.
\end{align*}
\] (A.6)

In the variance-gamma special case \( \text{VG}(\rho, \alpha, \beta) = \text{BG}(\xi = 0, \gamma = \rho, \alpha, \delta = \rho, \beta) \) the relevant Tricomi function reduces to a Macdonald function (modified Bessel function of the 2nd kind, hyperbolic Bessel function of the 3rd kind, Basset function, modified Hankel function) of the type ([12], 48:4:3 and 48:13:6)

\[ U(\rho, 2\alpha, x) = \frac{x^{\frac{3}{2}-\alpha}}{\sqrt{\pi}} e^{\frac{x}{2}} K_{\frac{1}{2}-\frac{3}{2}}(\frac{1}{2}x). \] (A.7)

Inserting these expressions into the Tricomi representation (A.6) one obtains the VG pdf

\[
f(x) = \frac{1}{\sqrt{\pi} \Gamma(\rho)} \left( \frac{|x|}{\alpha + \beta} \right)^{\alpha+\frac{1}{2}} \cdot \exp(-\frac{1}{2}(\alpha - \beta)x) \cdot K_{\frac{1}{2}-\frac{3}{2}}(\frac{1}{2}(\alpha + \beta)|x|), \quad x \neq 0.
\] (A.8)

This closed-form expression has been first derived in [3] for the parameterization

**International Journal of Scientific and Innovative Mathematical Research (IJSIMR)**
Margrabe Formulas for a Simple Bivariate Exponential Variance-Gamma Price Process (II) Statistical Estimation and Application

\[
(\theta, \sigma^2, \nu) = \left( (\alpha^{-1} - \beta^{-1}) \rho, 2(\alpha \beta)^{-1} \rho, \rho^{-1} \right).
\]

However, in its original form the VG pdf takes the less symmetrical form

\[
f(x) = \frac{2 \exp(-\sigma^2/2\theta)}{\sqrt{2\pi} \sigma} \cdot \Gamma(\nu^{-1}) \left( \frac{x^2}{\theta^2 + 2\nu^{-1} \sigma^2} \right)^{1/2} K_{\nu^{-1}/2} \left( \sigma^{-2} \sqrt{(\theta^2 + 2\nu^{-1} \sigma^2)x^2} \right), \quad x \neq 0.
\]

REFERENCES


AUTHOR’S BIOGRAPHY

Werner Hürlimann has studied mathematics and physics at ETHZ, where he obtained his PhD in 1980 with a thesis in higher algebra. After postdoctoral fellowships at Yale University and at the Max Planck Institute in Bonn, he became 1984 an actuary at Winterthur Life and Pensions, a senior actuary at Aon Re and IRMG Switzerland 2003-06, a senior consultant at IRIS in Zürich 2006-2008, and is currently employed at FRSGlobal Switzerland (a Wolters Kluwer Company). He has been visiting associate professor in actuarial science at the University of Toronto during the academic year 1988-89. More information is found at his homepage https://sites.google.com/site/whurlimann/.