Part III
Financial Engineering

Basket Option Pricing and Implied Correlation in a One-Factor Lévy Model

Daniël Linders and Wim Schoutens

Abstract In this paper we employ a one-factor Lévy model to determine basket option prices. More precisely, basket option prices are determined by replacing the distribution of the real basket with an appropriate approximation. For the approximate basket we determine the underlying characteristic function and hence we can derive the related basket option prices by using the Carr–Madan formula. We consider a three-moments-matching method. Numerical examples illustrate the accuracy of our approximations; several Lévy models are calibrated to market data and basket option prices are determined. In the last part we show how our newly designed basket option pricing formula can be used to define implied Lévy correlation by matching model and market prices for basket options. Our main finding is that the implied Lévy correlation smile is flatter than its Gaussian counterpart. Furthermore, if (near) at-the-money option prices are used, the corresponding implied Gaussian correlation estimate is a good proxy for the implied Lévy correlation.

Keywords Basket option · Implied correlation · One-factor Lévy model · Variance-Gamma

1 Introduction

Nowadays, an increased volume of multi-asset derivatives is traded. An example of such a derivative is a *basket option*. The basic version of such a multivariate product has the same characteristics as a vanilla option, but now the underlying is a basket of stocks instead of a single stock. The pricing of these derivatives is not a trivial task because it requires a model that jointly describes the stock prices involved.

D. Linders (✉)
Faculty of Business and Economics, KU Leuven, Naamsestraat 69,
3000 Leuven, Belgium
e-mail: daniel.linders@kuleuven.be

W. Schoutens
Faculty of Science, KU Leuven, Celestijnenlaan 200, 3001 Heverlee, Belgium
e-mail: Wim@Schoutens.be

K. Glau et al. (eds.), *Innovations in Derivatives Markets*, Springer Proceedings
in Mathematics & Statistics 165, DOI 10.1007/978-3-319-33446-2_16

Stock price models based on the lognormal model proposed in Black and Scholes [6] are popular choices from a computational point of view; however, they are not capable of capturing the skewness and kurtosis observed for log returns of stocks and indices. The class of Lévy processes provides a much better fit for the observed log returns and, consequently, the pricing of options and other derivatives in a Lévy setting is much more reliable. In this paper we consider the problem of pricing multi-asset derivatives in a multivariate Lévy model.

The most straightforward extension of the univariate Black and Scholes model is based on the *Gaussian copula model*, also called the multivariate Black and Scholes model. In this framework, the stocks composing the basket at a given point in time are assumed to be lognormally distributed and a Gaussian copula is connecting these marginals. Even in this simple setting, the price of a basket option is not given in a closed form and has to be approximated; see e.g. Hull and White [23], Brooks et al. [8], Milevsky and Posner [39], Rubinstein [42], Deelstra et al. [18], Carmona and Durrleman [12] and Linders [29], among others. However, the normality assumption for the marginals used in this pricing framework is too restrictive. Indeed, in Linders and Schoutens [30] it is shown that calibrating the Gaussian copula model to market data can lead to non-meaningful parameter values. This dysfunctioning of the Gaussian copula model is typically observed in distressed periods. In this paper we extend the classical Gaussian pricing framework in order to overcome this problem.

Several extensions of the Gaussian copula model are proposed in the literature. For example, Luciano and Schoutens [32] introduce a multivariate Variance Gamma model where dependence is modeled through a common jump component. This model was generalized in Semeraro [44], Luciano and Semeraro [33], and Guillaume [21]. A stochastic correlation model was considered in Fonseca et al. [19]. A framework for modeling dependence in finance using copulas was described in Cherubini et al. [14]. The pricing of basket options in these advanced multivariate stock price models is not a straightforward task. There are several attempts to derive closed form approximations for the price of a basket option in a non-Gaussian world. In Linders and Stassen [31], approximate basket option prices in a multivariate Variance Gamma model are derived, whereas Xu and Zheng [48, 49] consider a local volatility jump diffusion model. McWilliams [38] derives approximations for the basket option price in a stochastic delay model. Upper and lower bounds for basket option prices in a general class of stock price models with known joint characteristic function of the logreturns are derived in Caldana et al. [10].

In this paper we start from the one-factor Lévy model introduced in Albrecher et al. [1] to build a multivariate stock price model with correlated Lévy marginals. Stock prices are assumed to be driven by an idiosyncratic and a systematic factor. The idea of using a common market factor is not new in the literature and goes back to Vasicek [47]. Conditional on the common (or market) factor, the stock prices are independent. We show that our model generalizes the Gaussian model (with single correlation). Indeed, the idiosyncratic and systematic components are constructed from a Lévy process. Employing a Brownian motion in that construction delivers the Gaussian copula model, but other Lévy models arise by employing different Lévy

processes like VG, NIG, Meixner, etc. As a result, this new *one-factor Lévy model* is more flexible and can capture other types of dependence.

The correlation is by construction always positive and, moreover, we assume a single correlation. Stocks can, in reality, be negatively correlated and correlations between different stocks will differ. From a tractability point of view, however, reporting a single correlation number is often preferred over $n(n-1)/2$ pairwise correlations. The single correlation can be interpreted as a mean level of correlation and provides information about the general dependence among the stocks composing the basket. Such a single correlation appears, for example, in the construction of a correlation swap. Therefore, our framework may have applications in the pricing of such correlation products. Furthermore, calibrating a full correlation matrix may require an unrealistically large amount of data if the index consists of many stocks.

In the first part of this paper, we consider the problem of finding accurate approximations for the price of a basket option in the one-factor Lévy model. In order to value a basket option, the distribution of this basket has to be determined. However, the basket is a weighted sum of dependent stock prices and its distribution function is in general unknown or too complex to work with. Our valuation formula for the basket option is based on a moment-matching approximation. To be more precise, the (unknown) basket distribution is replaced by a shifted random variable having the same first three moments than the original basket. This idea was first proposed in Brigo et al. [7], where the Gaussian copula model was considered. Numerical examples illustrating the accuracy and the sensitivity of the approximation are provided.

In the second part of the paper we show how the well-established notions of implied volatility and implied correlation can be generalized in our multivariate Lévy model. We assume that a finite number of options, written on the basket and the components, are traded. The prices of these derivatives are observable and will be used to calibrate the parameters of our stock price model. An advantage of our modeling framework is that each stock is described by a volatility parameter and that the marginal parameters can be calibrated separately from the correlation parameter. We give numerical examples to show how to use the vanilla option curves to determine an implied Lévy volatility for each stock based on a Normal, VG, NIG, and Meixner process and determine basket option prices for different choices of the correlation parameter.

An *implied Lévy correlation* estimate arises when we tune the single correlation parameter such that the model price exactly hits the market price of a basket option for a given strike. We determine implied correlation levels for the stocks composing the Dow Jones Industrial Average in a Gaussian and a Variance Gamma setting. We observe that implied correlation depends on the strike and in the VG model, this implied Lévy correlation *smile* is flatter than in the Gaussian copula model. The standard technique to price non-traded basket options (or other multi-asset derivatives) is by interpolating on the implied correlation curve. It is shown in Linders and Schoutens [30] that in the Gaussian copula model, this technique can sometimes lead to non-meaningful correlation values. We show that the Lévy version of the implied correlation solves this problem (at least to some extent). Several papers consider the problem of measuring implied correlation between stock prices;

see e.g. Fonseca et al. [19], Tavin [46], Ballotta et al. [4], and Austing [2]. Our approach is different in that we determine implied correlation estimates in the one-factor Lévy model using multi-asset derivatives consisting of many assets (30 assets for the Dow Jones). When considering multi-asset derivatives with a low dimension, determining the model prices of these multi-asset derivatives becomes much more tractable. A related paper is Linders and Stassen [31], where the authors also use high-dimensional multi-asset derivative prices for calibrating a multivariate stock price model. However, whereas the current paper models the stock returns using correlated Lévy distributions, the cited paper uses time-changed Brownian motions with a common time change.

2 The One-Factor Lévy Model

We consider a market where n stocks are traded. The price level of stock j at some future time t, $0 \leq t \leq T$ is denoted by $S_j(t)$. Dividends are assumed to be paid continuously and the dividend yield of stock j is constant and deterministic over time. We denote this dividend yield by q_j. The current time is $t = 0$. We fix a future time T and we always consider the random variables $S_j(T)$ denoting the time-T prices of the different stocks involved. The price level of a basket of stocks at time T is denoted by $S(T)$ and given by

$$S(T) = \sum_{j=1}^{n} w_j S_j(T),$$

where $w_j > 0$ are weights which are fixed upfront. In case the basket represents the price of the Dow Jones, the weights are all equal. If this single weight is denoted by w, then $1/w$ is referred to as the Dow Jones Divisor.[1] The pay-off of a basket option with strike K and maturity T is given by $(S(T) - K)_+$, where $(x)_+ = \max(x, 0)$. The price of this basket option is denoted by $C[K, T]$. We assume that the market is arbitrage-free and that there exists a risk-neutral pricing measure $\mathbb{Q}$ such that the basket option price $C[K, T]$ can be expressed as the discounted risk-neutral expected value. In this pricing formula, discounting is performed using the risk-free interest rate r, which is, for simplicity, assumed to be deterministic and constant over time. Throughout the paper, we always assume that all expectations we encounter are well-defined and finite.

[1] More information and the current value of the Dow Jones Divisor can be found here: http://www.djindexes.com.

2.1 The Model

The most straightforward way to model dependent stock prices is to use a Black and Scholes model for the marginals and connect them with a Gaussian copula. A crucial (and simplifying) assumption in this approach is the normality assumption. It is well-known that log returns do not pass the test for normality. Indeed, log returns exhibit a skewed and leptokurtic distribution which cannot be captured by a normal distribution; see e.g. Schoutens [43].

We generalize the Gaussian copula approach by allowing the risk factors to be distributed according to any infinitely divisible distribution with known characteristic function. This larger class of distributions increases the flexibility to find a more realistic distribution for the log returns. In Albrecher et al. [1] a similar framework was considered for pricing CDO tranches; see also Baxter [5]. The Variance Gamma case was considered in Moosbrucker [40, 41], whereas Guillaume et al. [22] consider the pricing of CDO-squared tranches in this one-factor Lévy model. A unified approach for these CIID models (conditionally independent and identically distributed) is given in Mai et al. [36].

Consider an infinitely divisible distribution for which the characteristic function is denoted by ϕ. A stochastic process X can be built using this distribution. Such a process is called a Lévy process with mother distribution having the characteristic function ϕ. The Lévy process $X = \{X(t)|t \geq 0\}$ based on this infinitely divisible distribution starts at zero and has independent and stationary increments. Furthermore, for $s, t \geq 0$ the characteristic function of the increment $X(t+s) - X(t)$ is ϕ^s.

Assume that the random variable L has an infinitely divisible distribution and denote its characteristic function by ϕ_L. Consider the Lévy process $X = \{X(t)|t \in [0, 1]\}$ based on the distribution L. We assume that the process is standardized, i.e. $\mathbb{E}[X(1)] = 0$ and $\mathrm{Var}[X(1)] = 1$. One can then show that $\mathrm{Var}[X(t)] = t$, for $t \geq 0$. Define also a series of independent and standardized processes $X_j = \{X_j(t)|t \in [0, 1]\}$, for $j = 1, 2, \ldots, n$. The process X_j is based on an infinitely divisible distribution L_j with characteristic function ϕ_{L_j}. Furthermore, the processes $X_1, X_2, \ldots, X_n$ are independent from X. Take $\rho \in [0, 1]$. The r.v. A_j is defined by

$$A_j = X(\rho) + X_j(1 - \rho), \quad j = 1, 2, \ldots n. \tag{1}$$

In this construction, $X(\rho)$ and $X_j(1-\rho)$ are random variables having the characteristic function ϕ_L^ρ and $\phi_{L_j}^{1-\rho}$, respectively. Denote the characteristic function of A_j by ϕ_{A_j}. Because the processes X and X_j are independent and standardized, we immediately find that

$$\mathbb{E}[A_j] = 0, \quad \mathrm{Var}[A_j] = 1 \quad \text{and} \quad \phi_{A_j}(t) = \phi_L^\rho(t)\phi_{L_j}^{1-\rho}(t), \quad \text{for } j = 1, 2, \ldots, n. \tag{2}$$

Note that if X and X_j are both Lévy processes based on the same mother distribution L, we obtain the equality $A_j \overset{\mathrm{d}}{=} L$.

The parameter ρ describes the correlation between A_i and A_j, if $i \neq j$. Indeed, it was proven in Albrecher et al. [1] that in case A_j, $j = 1, 2, \ldots, n$ is defined by (1), we have that

$$\text{Corr}\left[A_i, A_j\right] = \rho. \tag{3}$$

We model the stock price levels $S_j(T)$ at time T for $j = 1, 2, \ldots, n$ as follows

$$S_j(T) = S_j(0)\mathrm{e}^{\mu_j T + \sigma_j \sqrt{T} A_j}, \quad j = 1, 2, \ldots, n, \tag{4}$$

where $\mu_j \in \mathbb{R}$ and $\sigma_j > 0$. Note that in this setting, each time-T stock price is modeled as the exponential of a Lévy process. Furthermore, a drift μ_j and a volatility parameter σ_j are added to match the characteristics of stock j. Our model, which we will call the *one-factor Lévy model*, can be considered as a generalization of the Gaussian model. Indeed, instead of a normal distribution, we allow for a Lévy distribution, while the Gaussian copula is generalized to a Lévy-based copula.[2] This model can also, at least to some extent, be considered as a generalization to the multidimensional case of the model proposed in Corcuera et al. [17] and the parameter σ_j in (4) can then be interpreted as the Lévy space (implied) volatility of stock j. The idea of building a multivariate asset model by taking a linear combination of a systematic and an idiosyncratic process can also be found in Kawai [26] and Ballotta and Bonfiglioli [3].

2.2 The Risk-Neutral Stock Price Processes

If we take

$$\mu_j = (r - q_j) - \frac{1}{T} \log \phi_L \left(-\mathrm{i}\sigma_j \sqrt{T}\right), \tag{5}$$

we find that

$$\mathbb{E}[S_j(T)] = \mathrm{e}^{(r-q_j)T} S_j(0), \quad j = 1, 2, \ldots, n.$$

From expression (5) we conclude that the risk-neutral dynamics of the stocks in the one-factor Lévy model are given by

$$S_j(T) = S_j(0)\mathrm{e}^{(r-q_j-\omega_j)T + \sigma_j \sqrt{T} A_j}, \quad j = 1, 2, \ldots, n, \tag{6}$$

where $\omega_j = \log \phi_L \left(-\mathrm{i}\sigma_j \sqrt{T}\right) / T$. We always assume ω_j to be finite. The first three moments of $S_j(T)$ can be expressed in terms of the characteristic function ϕ_{A_j}. By

[2] The Lévy-based copula refers to the copula between the r.v.'s $A_1, A_2, \ldots, A_n$ and is different from the Lévy copula introduced in Kallsen and Tankov [25].

the martingale property, we have that $\mathbb{E}\left[S_j(T)\right] = S_j(0)\mathrm{e}^{(r-q_j)T}$. The risk-neutral variance $\mathrm{Var}\left[S_j(T)\right]$ can be written as follows

$$\mathrm{Var}\left[S_j(T)\right] = S_j(0)^2\mathrm{e}^{2(r-q_j)T}\left(\mathrm{e}^{-2\omega_j T}\phi_{A_j}\left(-\mathrm{i}2\sigma_j\sqrt{T}\right) - 1\right).$$

The second and third moment of $S_j(T)$ are given by:

$$\mathbb{E}\left[S_j(T)^2\right] = \mathbb{E}[S_j(T)]^2\frac{\phi_{A_j}\left(-\mathrm{i}2\sigma_j\sqrt{T}\right)}{\phi_{A_j}\left(-\mathrm{i}\sigma_j\sqrt{T}\right)^2},$$

$$\mathbb{E}\left[S_j(T)^3\right] = \mathbb{E}[S_j(T)]^3\frac{\phi_{A_j}\left(-\mathrm{i}3\sigma_j\sqrt{T}\right)}{\phi_{A_j}\left(-\mathrm{i}\sigma_j\sqrt{T}\right)^3}.$$

We always assume that these quantities are finite. If the process X_j has mother distribution L, we can replace ϕ_{A_j} by ϕ_L in expression (5) and in the formulas for $\mathbb{E}\left[S_j(T)^2\right]$ and $\mathbb{E}\left[S_j(T)^3\right]$. From here on, we always assume that all Lévy processes are built on the same mother distribution. However, all results remain to hold in the more general case.

3 A Three-Moments-Matching Approximation

In order to price a basket option, one has to know the distribution of the random sum $S(T)$, which is a weighted sum of dependent random variables. This distribution is in most situations unknown or too cumbersome to work with. Therefore, we search for a new random variable which is sufficiently 'close' to the original random variable, but which is more attractive to work with. More concretely, we introduce in this section a new approach for approximating $C[K, T]$ by replacing the sum $S(T)$ with an appropriate random variable $\widetilde{S}(T)$ which has a simpler structure, but for which the first three moments coincide with the first three moments of the original basket $S(T)$. This moment-matching approach was also considered in Brigo et al. [7] for the multivariate Black and Scholes model.

Consider the Lévy process $Y = \{Y(t) \mid 0 \leq t \leq 1\}$ with infinitely divisible distribution L. Furthermore, we define the random variable A as

$$A = Y(1).$$

In this case, the characteristic function of A is given by ϕ_L. The sum $S(T)$ is a weighted sum of dependent random variables and its cdf is unknown. We approximate the sum $S(T)$ by $\widetilde{S}(T)$, defined by

$$\widetilde{S}(T) = \bar{S}(T) + \lambda, \tag{7}$$

where $\lambda \in \mathbb{R}$ and

$$\bar{S}(T) = S(0) \exp\left\{(\bar{\mu} - \bar{\omega})T + \bar{\sigma}\sqrt{T}A\right\}. \tag{8}$$

The parameter $\bar{\mu} \in \mathbb{R}$ determines the drift and $\bar{\sigma} > 0$ is the volatility parameter. These parameters, as well as the shifting parameter λ, are determined such that the first three moments of $\widetilde{S}(T)$ coincide with the corresponding moments of the real basket $S(T)$. The parameter $\bar{\omega}$, defined as follows

$$\bar{\omega} = \frac{1}{T} \log \phi_L\left(-\mathrm{i}\bar{\sigma}\sqrt{T}\right),$$

is assumed to be finite.

3.1 Matching the First Three Moments

The first three moments of the basket $S(T)$ are denoted by m_1, m_2, and m_3 respectively. In the following lemma, we express the moments m_1, m_2, and m_3 in terms of the characteristic function ϕ_L and the marginal parameters. A proof of this lemma is provided in the appendix.

Lemma 1 *Consider the one-factor Lévy model (6) with infinitely divisible mother distribution L. The first two moments m_1 and m_2 of the basket $S(T)$ can be expressed as follows*

$$m_1 = \sum_{j=1}^{n} w_j \mathbb{E}\left[S_j(T)\right], \tag{9}$$

$$m_2 = \sum_{j=1}^{n}\sum_{k=1}^{n} w_j w_k \mathbb{E}\left[S_j(T)\right]\mathbb{E}\left[S_k(T)\right]\left(\frac{\phi_L\left(-i(\sigma_j+\sigma_k)\sqrt{T}\right)}{\phi_L\left(-i\sigma_j\sqrt{T}\right)\phi_L\left(-i\sigma_k\sqrt{T}\right)}\right)^{\rho_{j,k}} \tag{10}$$

where

$$\rho_{j,k} = \begin{cases} \rho, & \text{if } j \neq k; \\ 1, & \text{if } j = k. \end{cases}$$

The third moment m_3 of the basket $S(T)$ is given by

$$m_3 = \sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} w_j w_k w_l \mathbb{E}\left[S_j(T)\right]\mathbb{E}\left[S_k(T)\right]\mathbb{E}\left[S_l(T)\right] \times \frac{\phi_L\left(-i\left(\sigma_j+\sigma_k+\sigma_l\right)\sqrt{T}\right)^{\rho}}{\phi_L\left(-i\sigma_j\sqrt{T}\right)\phi_L\left(-i\sigma_k\sqrt{T}\right)\phi_L\left(-i\sigma_l\sqrt{T}\right)} A_{j,k,l}, \tag{11}$$

where

$$A_{j,k,l} = \begin{cases} \left(\phi_L\left(-i\sigma_j\sqrt{T}\right)\phi_L\left(-i\sigma_k\sqrt{T}\right)\phi_L\left(-i\sigma_l\sqrt{T}\right)\right)^{1-\rho}, & \text{if } j\neq k, k\neq l \text{ and } j\neq l; \\ \left(\phi_L\left(-i(\sigma_j+\sigma_k)\sqrt{T}\right)\phi_L\left(-i\sigma_l\sqrt{T}\right)\right)^{1-\rho}, & \text{if } j=k, k\neq l; \\ \left(\phi_L\left(-i(\sigma_k+\sigma_l)\sqrt{T}\right)\phi_L\left(-i\sigma_j\sqrt{T}\right)\right)^{1-\rho}, & \text{if } j\neq k, k=l; \\ \left(\phi_L\left(-i(\sigma_j+\sigma_l)\sqrt{T}\right)\phi_L\left(-i\sigma_k\sqrt{T}\right)\right)^{1-\rho}, & \text{if } j=l, k\neq l; \\ \phi_L\left(-i\left(\sigma_j+\sigma_k+\sigma_l\right)\sqrt{T}\right)^{1-\rho}, & \text{if } j=k=l. \end{cases}$$

In Sect. 2.2 we derived the first three moments for each stock $j, j = 1, 2, \ldots, n$. Taking into account the similarity between the price $S_j(T)$ defined in (6) and the approximate r.v. $\bar{S}(T)$, defined in (8), we can determine the first three moments of $\bar{S}(T)$:

$$\mathbb{E}\left[\bar{S}(T)\right] = S(0)\mathrm{e}^{\bar{\mu}T} =: \xi,$$

$$\mathbb{E}\left[\bar{S}(T)^2\right] = \mathbb{E}\left[\bar{S}(T)\right]^2 \frac{\phi_L\left(-\mathrm{i}2\bar{\sigma}\sqrt{T}\right)}{\phi_L\left(-\mathrm{i}\bar{\sigma}\sqrt{T}\right)^2} =: \xi^2\alpha,$$

$$\mathbb{E}\left[\bar{S}(T)^3\right] = \mathbb{E}\left[\bar{S}(T)\right]^3 \frac{\phi_L\left(-\mathrm{i}3\bar{\sigma}\sqrt{T}\right)}{\phi_L\left(-\mathrm{i}\bar{\sigma}\sqrt{T}\right)^3} =: \xi^3\beta.$$

These expressions can now be used to determine the first three moments of the approximate r.v. $\widetilde{S}(T)$:

$$\begin{aligned} \mathbb{E}\left[\widetilde{S}(T)\right] &= \mathbb{E}\left[\bar{S}(T)\right] + \lambda, \\ \mathbb{E}\left[\widetilde{S}(T)^2\right] &= \mathbb{E}\left[\bar{S}(T)^2\right] + \lambda^2 + 2\lambda\mathbb{E}\left[\bar{S}(T)\right], \\ \mathbb{E}\left[\widetilde{S}(T)^3\right] &= \mathbb{E}\left[\bar{S}(T)^3\right] + \lambda^3 + 3\lambda^2\mathbb{E}\left[\bar{S}(T)\right] + 3\lambda\mathbb{E}\left[\bar{S}(T)^2\right]. \end{aligned}$$

Determining the parameters $\bar{\mu}$, $\bar{\sigma}$ and the shifting parameter λ by matching the first three moments, results in the following set of equations

$$\begin{aligned} m_1 &= \xi + \lambda, \\ m_2 &= \xi^2\alpha + \lambda^2 + 2\lambda\xi, \\ m_3 &= \xi^3\beta + \lambda^3 + 3\lambda^2\xi + 3\lambda\xi^2\alpha. \end{aligned}$$

These equations can be recast in the following set of equations

$$\begin{aligned}\lambda &= m_1 - \xi,\\ \xi^2 &= \frac{m_2 - m_1^2}{\alpha - 1},\\ 0 &= \left(\frac{m_2 - m_1^2}{\alpha - 1}\right)^{3/2}(\beta + 2 - 3\alpha) + 3m_1m_2 - 2m_1^3 - m_3.\end{aligned}$$

Remember that α and β are defined by

$$\alpha = \frac{\phi_L\left(-\mathrm{i}2\bar{\sigma}\sqrt{T}\right)}{\phi_L\left(-\mathrm{i}\bar{\sigma}\sqrt{T}\right)^2} \quad \text{and} \quad \beta = \frac{\phi_L\left(-\mathrm{i}3\bar{\sigma}\sqrt{T}\right)}{\phi_L\left(-\mathrm{i}\bar{\sigma}\sqrt{T}\right)^3}.$$

Solving the third equation results in the parameter $\bar{\sigma}$. Note that this equation does not always have a solution. This issue was also discussed in Brigo et al. [7] for the Gaussian copula case. However, in our numerical studies we did not encounter any numerical problems. If we know $\bar{\sigma}$, we can also determine ξ and λ from the first two equations. Next, the drift $\bar{\mu}$ can be determined from

$$\bar{\mu} = \frac{1}{T}\log\frac{\xi}{S(0)}.$$

3.2 Approximate Basket Option Pricing

The price of a basket option with strike K and maturity T is denoted by $C[K, T]$. This unknown price is approximated in this section by $C^{MM}[K, T]$, which is defined as

$$C^{MM}[K, T] = \mathrm{e}^{-rT}\mathbb{E}\left[\left(\widetilde{S}(T) - K\right)_+\right].$$

Using expression (7) for $\widetilde{S}(T)$, the price $C^{MM}[K, T]$ can be expressed as

$$C^{MM}[K, T] = \mathrm{e}^{-rT}\mathbb{E}\left[\left(\bar{S}(T) - (K - \lambda)\right)_+\right].$$

Note that the distribution of $\bar{S}(T)$ is also depending on the choice of λ. In order to determine the price $C^{MM}[K, T]$, we should be able to price an option written on $\bar{S}(T)$, with a shifted strike $K - \lambda$. Determining the approximation $C^{MM}[K, T]$ using the Carr–Madan formula requires knowledge about the characteristic function $\phi_{\log\bar{S}(T)}$ of $\log\bar{S}(T)$:

$$\phi_{\log\bar{S}(T)}(u) = \mathbb{E}\left[\mathrm{e}^{\mathrm{i}u\log\bar{S}(T)}\right].$$

Using expression (8) we find that

$$\phi_{\log \bar{S}(T)}(u) = \mathbb{E}\left[\exp\left\{\mathrm{i}u\left(\log S(0) + (\bar{\mu} - \bar{\omega})T + \bar{\sigma}\sqrt{T}A\right)\right\}\right].$$

The characteristic function of A is ϕ_L, from which we find that

$$\phi_{\log \bar{S}(T)}(u) = \exp\left\{\mathrm{i}u\left(\log S(0) + (\bar{\mu} - \bar{\omega})T\right)\right\}\phi_L\left(u\bar{\sigma}\sqrt{T}\right).$$

Note that nowhere in this section we used the assumption that the basket weights w_j are strictly positive. Therefore, the three-moments-matching approach proposed in this section can also be used to price, e.g. spread options. However, for pricing spread options, alternative methods exist; see e.g. Carmona and Durrleman [11], Hurd and Zhou [24] and Caldana and Fusai [9].

3.3 The FFT Method and Basket Option Pricing

Consider the random variable X. In this section we show that if the characteristic function $\phi_{\log X}$ of this r.v. X is known, one can approximate the discounted stop-loss premium

$$\mathrm{e}^{-rT}\mathbb{E}\left[(X - K)_+\right],$$

for any $K > 0$.

Let $\alpha > 0$ and assume that $\mathbb{E}\left[X^{\alpha+1}\right]$ exists and is finite. It was proven in Carr and Madan [13] that the price $\mathrm{e}^{-rT}\mathbb{E}\left[(X - K)_+\right]$ can be expressed as follows

$$\mathrm{e}^{-rT}\mathbb{E}\left[(X - K)_+\right] = \frac{\mathrm{e}^{-\alpha\log(K)}}{\pi}\int_0^{+\infty}\exp\left\{-\mathrm{i}v\log(K)\right\}g(v)\mathrm{d}v, \tag{12}$$

where

$$g(v) = \frac{\mathrm{e}^{-rT}\phi_{\log X}\left(v - (\alpha+1)\mathrm{i}\right)}{\alpha^2 + \alpha - v^2 + \mathrm{i}(2\alpha+1)v}. \tag{13}$$

The approximation $C^{MM}[K, T]$ was introduced in Sect. 3 and the random variable X now denotes the moment-matching approximation $\widetilde{S}(T) = \bar{S}(T) + \lambda$. The approximation $C^{MM}[K, T]$ can then be determined as the option price written on $\bar{S}(T)$ and with shifted strike price $K - \lambda$.

Table 1 Overview of infinitely divisible distributions

	Gaussian	Variance Gamma
Parameters	$\mu \in \mathbb{R}, \sigma > 0$	$\mu, \theta \in \mathbb{R}, \sigma, \nu > 0$
Notation	$\mathcal{N}(\mu, \sigma^2)_\sigma$	$VG(\sigma, \nu, \theta, \mu)$
$\phi(u)$	$e^{iu\mu + \frac{1}{2}\sigma^2 u_\sigma}$	$e^{iu\mu}\left(1 - iu\theta\nu + u^2\sigma^2\nu/2\right)^{-1/\nu}$
Mean	μ	$\mu + \theta$
Variance	σ^2	${\sigma^2}_\sigma + \nu\theta^2$
Standardized version	$\mathcal{N}(0, 1)$	$VG(\kappa\sigma, \nu, \kappa\theta, -\kappa\theta)$ where $\kappa = \frac{1}{\sqrt{\sigma^2 + \theta^2{}_\sigma \nu}}$
	Normal Inverse Gaussian	Meixner
Parameters	$\alpha, \delta > 0, \beta \in (-\alpha, \alpha), \mu \in \mathbb{R}$	$\alpha, \delta > 0, \beta \in (-\pi, \pi), \mu \in \mathbb{R}$
Notation	$NIG(\alpha, \beta, \delta, \mu)$	$MX(\alpha, \beta, \delta, \mu)$
$\phi(u)$	$e^{iu\mu - \delta\left(\sqrt{\alpha^2 - (\beta + iu)^2} - \sqrt{\alpha^2 - \beta^2{}_\sigma}\right)}$	$e^{iu\mu}\left(\frac{\cos(\beta/2)}{\cosh((\alpha u - i\beta)/2)}\right)^{2\delta}$
Mean	$\mu + \frac{\delta\beta}{\sqrt{\alpha^2 - \beta^2{}_\sigma}}$	$\mu + \alpha\delta\tan(\beta/2)$
Variance	$\alpha^2\delta\left(\alpha^2 - \beta^2\right)^{-3/2}$	$\cos^{-2}(\beta/2)\alpha^2{}_\sigma\delta/2$
Standardized version	$NIG\left(\alpha, \beta, (\alpha^2 - \beta^2)^{3/2}, \frac{-(\alpha^2 - \beta^2)\beta}{\alpha^2}\right)$	$MX\left(\alpha, \beta, \frac{2\cos^2(\frac{\beta}{2})}{\alpha^2{}_\sigma}, \frac{-\sin(\beta)}{\alpha}\right)$

4 Examples and Numerical Illustrations

The Gaussian copula model with equicorrelation is a member of our class of one-factor Lévy models. In this section we discuss how to build the Gaussian, Variance Gamma, Normal Inverse Gaussian, and Meixner models. However, the reader is invited to construct one-factor Lévy models based on other Lévy-based distributions; e.g. CGMY, Generalized hyperbolic, etc. distributions.

Table 1 summarizes the Gaussian, Variance Gamma, Normal Inverse Gaussian, and the Meixner distributions, which are all infinitely divisible. In the last row, it is shown how to construct a standardized version for each of these distributions. We assume that L is distributed according to one of these standardized distributions. Hence, L has zero mean and unit variance. Furthermore, the characteristic function ϕ_L of L is given in closed form. We can then define the Lévy processes X and X_j, $j = 1, 2, \ldots, n$ based on the mother distribution L. The random variables A_j, $j = 1, 2, \ldots, n$, are modeled using expression (1).

Table 2 Basket option prices in the one-factor VG model with $S_1(0) = 40$, $S_2(0) = 50$, $S_3(0) = 60$, $S_4(0) = 70$, and $\rho = 0$

K	$C^{mc}[K,T]$	$C^{MM}[K,T]$	Length CI
$\sigma_1 = 0.2; \sigma_2 = 0.2; \sigma_3 = 0.2; \sigma_4 = 0.2$			
50	6.5748	6.5676	4.27E-03
55	2.4363	2.4781	3.05E-03
60	0.2651	0.2280	9.29E-04
$\sigma_1 = 0.5; \sigma_2 = 0.5; \sigma_3 = 0.5; \sigma_4 = 0.5$			
55	4.1046	4.2089	6.31E-03
60	1.7774	1.7976	4.13E-03
65	0.5474	0.4637	2.16E-03
$\sigma_1 = 0.8; \sigma_2 = 0.8; \sigma_3 = 0.8; \sigma_4 = 0.8$			
60	3.2417	3.3371	7.16E-03
65	1.6806	1.6429	5.08E-03
70	0.7581	0.6375	3.30E-03
$\sigma_1 = 0.6; \sigma_2 = 1.2; \sigma_3 = 0.3; \sigma_4 = 0.9$			
55	5.5067	5.6719	9.44E-03
60	3.2266	3.3305	7.31E-03
65	1.6972	1.6750	5.26E-03
70	0.7889	0.6830	3.52E-03

4.1 Variance Gamma

Although pricing basket option under a normality assumption is tractable from a computational point of view, it introduces a high degree of model risk; see e.g. Leoni and Schoutens [28]. The Variance Gamma distribution has already been proposed as a more flexible alternative to the Brownian setting; see e.g. Madan and Seneta [34] and Madan et al. [35].

We consider two numerical examples where L has a Variance Gamma distribution with parameters $\sigma = 0.5695$, $\nu = 0.75$, $\theta = -0.9492$, $\mu = 0.9492$. Table 2 contains the numerical values for the first illustration, where a four-basket option paying $\left(\frac{1}{4}\sum_{j=1}^{4} S_j(T) - K\right)_+$ at time T is considered. We use the following parameter values: $r = 6\,\%, T = 0.5, \rho = 0$ and $S_1(0) = 40$, $S_2(0) = 50$, $S_3(0) = 60$, $S_4(0) = 70$. These parameter values are also used in Sect. 5 of Korn and Zeytun [27]. We denote by $C^{mc}[K,T]$ the corresponding Monte Carlo estimate for the price $C[K,T]$. Here, 10^7 number of simulations are used. The approximation of the basket option price $C[K,T]$ using the moment-matching approach outlined in Sect. 3 is denoted by $C^{MM}[K,T]$. A comparison between the empirical density and the approximate density is provided in Fig. 1.

In the second example, we consider the basket $S(T) = w_1 X_1(T) + w_2 X_2(T)$, written on two non-dividend paying stocks. We use as parameter values the ones

also used in Sect. 7 of Deelstra et al. [18], hence $r = 5\,\%$, $X_1(0) = X_2(0) = 100$, and $w_1 = w_2 = 0.5$. Table 3 gives numerical values for these basket options. Note that strike prices are expressed in terms of forward moneyness. A basket strike price K has forward moneyness equal to $K/\mathbb{E}[S]$. We can conclude that the three-moments-matching approximation gives acceptable results. For far out-of-the-money call options, the approximation is not always able to closely approximate the real basket option price.

We also investigate the sensitivity with respect to the Variance Gamma parameters σ, ν, and θ and to the correlation parameter ρ. We consider a basket option consisting of 3 stocks, i.e. $n = 3$. From Tables 2 and 3, we observe that the error is the biggest in case we consider different marginal volatilities and the option under consideration is an out-of-the-money basket call. Therefore, we put $\sigma_1 = 0.2$, $\sigma_2 = 0.4$, $\sigma_3 = 0.6$ and we determine the prices $C^{mc}[K, T]$ and $C^{MM}[K, T]$ for $K = 105.13$. The other parameter values are: $r = 0.05$, $\rho = 0.5$, $w_1 = w_2 = w_3 = 1/3$ and $T = 1$. The first panel of Fig. 2 shows the relative error for varying σ. The second panel of Fig. 2 shows the relative error in function of ν. The sensitivity with respect to θ is shown in the third panel of Fig. 2. Finally, the fourth panel of Fig. 2 shows the relative error in function of ρ.

The numerical results show that the approximations do not always manage to closely approximate the true basket option price. Especially when some of the volatilities deviate substantially from the other ones, the accuracy of the approximation deteriorates. The dysfunctioning of the moment-matching approximation in the Gaussian copula model was already reported in Brigo et al. [7]. However, in order to calibrate the Lévy copula model to available option data, the availability of a basket option pricing formula which can be evaluated in a fast way, is of crucial importance. Table 4 shows the CPU times[3] for the one-factor VG model for different basket dimensions. The calculation time of approximate basket option prices when 100 stocks are involved is less than one second. Therefore, the moment-matching approximation is a good candidate for calibrating the one-factor Lévy model.

4.2 Pricing Basket Options

In this subsection we explain how to determine the price of a basket option in a realistic situation where option prices of the components of the basket are available and used to calibrate the marginal parameters. In our example, the basket under consideration consists of 2 major stock market indices ($n = 2$), the S&P500 and the Nasdaq:

$$\text{Basket} = w_1 \text{S\&P 500} + w_2 \text{Nasdaq}.$$

The pricing date is February 19, 2009 and we determine prices for the Normal, VG, NIG, and Meixner case. The details of the basket are listed in Table 5. The weights

[3] The numerical illustrations are performed on an Intel Core i7, 2.70 GHz.

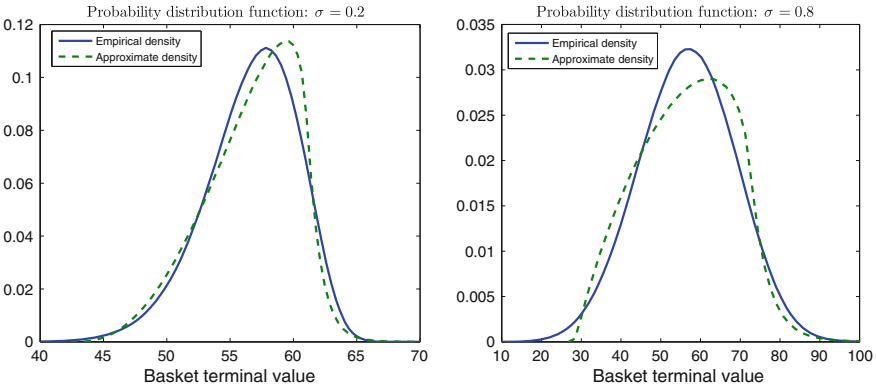

Fig. 1 Probability density function of the real basket (*solid line*) and the approximate basket (*dashed line*). The basket option consists of 4 stocks and $r = 0.06$, $\rho = 0$, $T = 1/2$, $w_1 = w_2 = w_3 = w_4 = \frac{1}{4}$. All volatility parameters are equal to σ

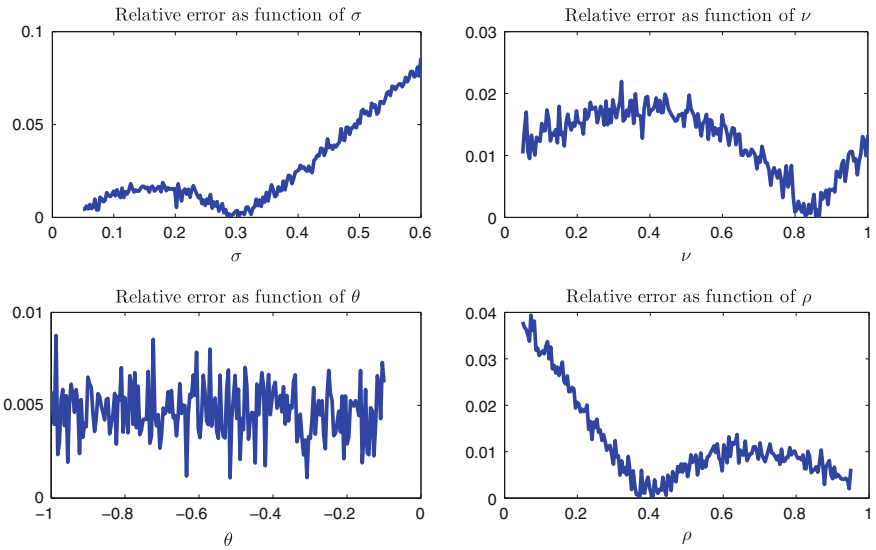

Fig. 2 Relative error in the one-factor VG model for the three-moments-matching approximation. The basket option consists of 3 stocks and $r = 0.05$, $\rho = 0.5$, $T = 1$, $\sigma_1 = 0.2$, $\sigma_2 = 0.4$, $\sigma_3 = 0.6$, $w_1 = w_2 = w_3 = \frac{1}{3}$. The strike price is $K = 105.13$. In the benchmark model, the VG parameters are $\sigma = 0.57$, $\nu = 0.75$, $\theta = -0.95$, $\mu = 0.95$

w_1 and w_2 are chosen such that the initial price $S(0)$ of the basket is equal to 100. The maturity of the basket option is equal to 30 days.

The S&P 500 and Nasdaq option curves are denoted by C_1 and C_2, respectively. These option curves are only partially known. The traded strikes for curve C_j are denoted by $K_{i,j}$, $i = 1, 2, \ldots, N_j$, where $N_j > 1$. If the volatilities σ_1 and σ_2

Table 3 Basket option prices in the one-factor VG model with $r = 0.05$, $w_1 = w_2 = 0.5$, $X_1(0) = X_2(0) = 100$ and $\sigma_1 = \sigma_2$

	T	ρ	σ_1	$C^{mc}[K,T]$	$C^{MM}[K,T]$	Length CI
$K = 115.64$	1	0.3	0.2	1.3995	1.3113	4.08E-03
			0.4	5.5724	5.6267	1.26E-02
		0.7	0.2	1.8963	1.8706	4.96E-03
			0.4	6.9451	7.0095	1.47E-02
$K = 127.80$	3	0.3	0.2	4.4427	4.4565	1.14E-02
			0.4	11.3138	11.5920	2.77E-02
		0.7	0.2	5.6002	5.6368	1.34E-02
			0.4	13.7444	13.9336	3.23E-02
$K = 105.13$	1	0.3	0.2	5.5312	5.5965	8.78E-03
			0.4	10.1471	10.3515	1.73E-02
		0.7	0.2	6.327	6.3731	9.74E-03
			0.4	11.7163	11.8379	1.95E-02
$K = 116.18$	3	0.3	0.2	8.9833	9.1489	1.66E-02
			0.4	15.8784	16.2498	3.27E-02
		0.7	0.2	10.3513	10.4528	1.86E-02
			0.4	18.4042	18.6214	3.73E-02
$K = 94.61$	1	0.3	0.2	12.3514	12.4371	1.29E-02
			0.4	16.213	16.4493	2.17E-02
		0.7	0.2	13.0696	13.1269	1.40E-02
			0.4	17.7431	17.8690	2.40E-02
$K = 104.57$	3	0.3	0.2	15.1888	15.3869	2.15E-02
			0.4	21.3994	21.7592	3.76E-02
		0.7	0.2	16.5069	16.6232	2.36E-02
			0.4	23.8489	24.0507	4.23E-02

and the characteristic function ϕ_L of the mother distribution L are known, we can determine the model price of an option on asset j with strike K and maturity T. This price is denoted by $C_j^{model}[K, T; \Theta, \sigma_j]$, where Θ denotes the vector containing the model parameters of L. Given the systematic component, the stocks are independent. Therefore, we can use the observed option curves C_1 and C_2 to calibrate the model parameters as follows:

Algorithm 1 (*Determining the parameters Θ and σ_j of the one-factor Lévy model*)

Step 1: Choose a parameter vector Θ.
Step 2: For each stock $j = 1, 2, \ldots, n$, determine the volatility σ_j as follows:

$$\sigma_j = \arg\min_{\sigma} \frac{1}{N_j} \sum_{i=1}^{N_j} \frac{\left| C_j^{model}[K_{i,j}, T; \Theta, \sigma] - C_j[K_{i,j}] \right|}{C_j[K_{i,j}]},$$

Table 4 The CPU time (in seconds) for the one-factor VG model for increasing basket dimension n

n	CPU TIMES
	Moment Matching
5	0.1991
10	0.1994
20	0.1922
30	0.2043
40	0.2335
50	0.2888
60	0.3705
70	0.4789
80	0.5909
90	0.6862
100	0.8680

The following parameters are used: $r = 0.05$, $T = 1$, $\rho = 0.5$, $w_j = \frac{1}{n}$, $\sigma_j = 0.4$, $q_j = 0$, $S_j(0) = 100$, for $j = 1, 2, \ldots, n$. The basket strike is $K = 105.13$

Table 5 Input data for the basket option

Date	Feb 19, 2009	
Maturity	March 21, 2009	
	S&P 500	Nasdaq
Forward	777.76	1116.72
Weights	0.06419	0.0428

Step 3: Determine the total error:

$$\text{error} = \sum_{j=1}^{n} \frac{1}{N_j} \sum_{i=1}^{N_j} \frac{\left| C_j^{model}[K_{i,j}, T; \Theta, \sigma_j] - C_j[K_{i,j}] \right|}{C_j[K_{i,j}]}.$$

Repeat these three steps until the parameter vector Θ is found for which the total error is minimal. The corresponding volatilities $\sigma_1, \sigma_2, \ldots, \sigma_n$ are called the implied Lévy volatilities.

Only a limited number of option quotes is required to calibrate the one-factor Lévy model. Indeed, the parameter vector Θ can be determined using all available option quotes. Additional, one volatility parameter has to be determined for each stock. However, other methodologies for determining Θ exist. For example, one can fix the parameter Θ upfront, as is shown in Sect. 5.2. In such a situation, only one implied Lévy volatility has to be calibrated for each stock.

The calibrated parameters together with the calibration error are listed in Table 6. Note that the relative error in the VG, Meixner, and NIG case is significantly smaller

Table 6 One-factor Lévy models: Calibrated model parameters

Model	Calibration error (%)	Model Parameters			Volatilities	
Normal	10.89	μ_{normal}	σ_{normal}		σ_1	σ_2
		0	1		0.2821	0.2734
VG	2.83	σ_{VG}	ν_{VG}	θ_{VG}		
		0.3477	0.49322	−0.3919	0.3716	0.3628
Meixner	2.81	$\alpha_{Meixner}$	$\beta_{Meixner}$			
		1.1689	−1.6761		0.3799	0.3709
NIG	2.89	α_{NIG}	β_{NIG}			
		2.2768	−1.4951		0.3863	0.3772

Table 7 Basket option prices for the basket given in Table 5

ρ	K	$C^{BLS}[K,T]$	$C^{VG}[K,T]$	$C^{Meixner}[K,T]$	$C^{NIG}[K,T]$
0.1	90	10.1783	10.7380	10.7893	10.8087
	95	5.9457	6.7092	6.7482	6.7418
	100	2.8401	3.4755	3.4843	3.4642
	105	1.0724	1.3375	1.3381	1.3374
	110	0.3158	0.3613	0.3690	0.3766
	120	0.0133	0.0198	0.0204	0.0197
0.5	90	10.3557	11.1445	11.2037	11.2169
	95	6.3160	7.2359	7.2754	7.2605
	100	3.3139	4.0376	4.0436	4.0154
	105	1.4699	1.7870	1.7798	1.7706
	110	0.5480	0.5857	0.5907	0.5980
	120	0.0461	0.0419	0.0421	0.0415
0.8	90	10.5000	11.4203	11.4837	11.4932
	95	6.5745	7.5877	7.6280	7.6091
	100	3.6292	4.4229	4.4287	4.3970
	105	1.7462	2.1247	2.1149	2.1010
	110	0.7301	0.7923	0.7954	0.8015
	120	0.0852	0.0726	0.0726	0.0723

The time to maturity is 30 days

than in the normal case. Using the calibrated parameters for the mother distribution L together with the volatility parameters σ_1 and σ_2, we can determine basket option prices in the different model settings. Note that here and in the sequel of the paper, we always use the three-moments-matching approximation for determining basket option prices. We put $T = 30$ days and consider the cases where the correlation parameter ρ is given by 0.1, 0.5, and 0.8. The corresponding basket option prices are listed in Table 7. One can observe from the table that each model generates a different basket option price, i.e. there is model risk. However, the difference between the

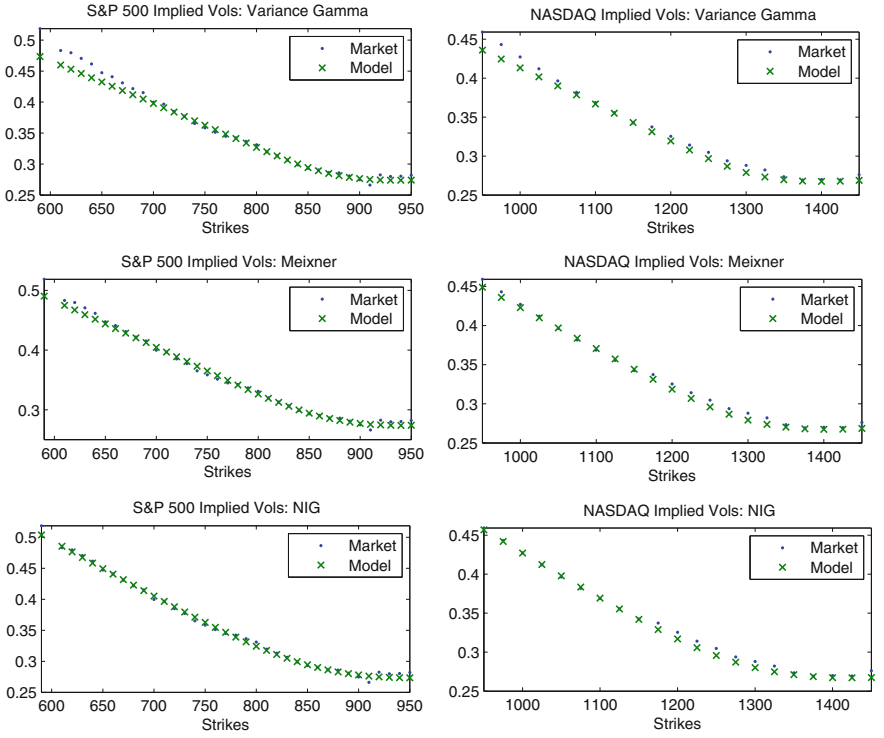

Fig. 3 Implied market and model volatilities for February 19, 2009 for the S&P 500 (*left*) and the Nasdaq (*right*), with time to maturity 30 days

Gaussian and the non-Gaussian models is much more pronounced than the difference within the non-Gaussian models. We also find that using normally distributed log returns, one underestimates the basket option prices. Indeed, the basket option prices $C^{VG}[K,T]$, $C^{Meixner}[K,T]$ and $C^{NIG}[K,T]$ are larger than $C^{BLS}[K,T]$. In the next section, however, we encounter situations where the Gaussian basket option price is larger than the corresponding VG price for out-of-the-money options. The reason for this behavior is that marginal log returns in the non-Gaussian situations are negatively skewed, whereas these distributions are symmetric in the Gaussian case. This skewness results in a lower probability of ending in the money for options with a sufficiently large strike (Fig. 3).

5 Implied Lévy Correlation

In Sect. 4.2 we showed how the basket option formulas can be used to obtain basket option prices in the Lévy copula model. The parameter vector Θ describing the

mother distribution L and the implied Lévy volatility parameters σ_j can be calibrated using the observed vanilla option curves $C_j[K, T]$ of the stocks composing the basket $S(T)$; see Algorithm 1. In this section we show how an implied Lévy correlation estimate ρ can be obtained if in addition to the vanilla options, market prices for a basket option are also available.

We assume that $S(T)$ represents the time-T price of a stock market index. Examples of such stock market indices are the Dow Jones, S&P 500, EUROSTOXX 50, and so on. Furthermore, options on $S(T)$ are traded and their prices are observable for a finite number of strikes. In this situation, pricing these index options is not a real issue; we denote the market price of an index option with maturity T and strike K by $C[K, T]$. Assume now that the stocks composing the index can be described by the one-factor Lévy model (6). If the parameter vector Θ and the marginal volatility vector $\underline{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ are determined using Algorithm 1, the model price $C^{model}[K, T; \underline{\sigma}, \Theta, \rho]$ for the basket option only depends on the choice of the correlation ρ. An *implied correlation* estimate for ρ arises when we match the model price with the observed index option price.

Definition 1 (*Implied Lévy correlation*) Consider the one-factor Lévy model defined in (6). The *implied Lévy correlation* of the index $S(T)$ with moneyness $\pi = S(T)/S(0)$, denoted by $\rho[\pi]$, is defined by the following equation:

$$C^{model}\left[K, T; \underline{\sigma}, \Theta, \rho[\pi]\right] = C[K, T], \tag{14}$$

where $\underline{\sigma}$ contains the marginal implied volatilities and Θ is the parameter vector of L.

Determining an implied correlation estimate $\rho[K/S(0)]$ requires an inversion of the pricing formula $\rho \to C^{model}[K, T; \underline{\sigma}, \Theta, \rho]$. However, the basket option price is not given in a closed form and determining this price using Monte Carlo simulation would result in a slow procedure. If we determine $C^{model}[K, T; \underline{\sigma}, \Theta, \rho]$ using the three-moments-matching approach, implied correlations can be determined in a fast and efficient way. The idea of determining implied correlation estimates based on an approximate basket option pricing formula was already proposed in Chicago Board Options Exchange [15], Cont and Deguest [16], Linders and Schoutens [30], and Linders and Stassen [31].

Note that in case we take L to be the standard normal distribution, $\rho[\pi]$ is an implied Gaussian correlation; see e.g. Chicago Board Options Exchange [15] and Skintzi and Refenes [45]. Equation (14) can be considered as a generalization of the implied Gaussian correlation. Indeed, instead of determining the single correlation parameter in a multivariate model with normal log returns and a Gaussian copula, we can now extend the model to the situation where the log returns follow a Lévy distribution. A similar idea was proposed in Garcia et al. [20] and further studied in Masol and Schoutens [37]. In these papers, Lévy base correlation is defined using CDS and CDO prices.

The proposed methodology for determining implied correlation estimates can also be applied to other multi-asset derivatives. For example, implied correlation

estimates can be extracted from traded spread options [46], best-of basket options [19], and quanto options [4]. Implied correlation estimates based on various multi-asset products are discussed in Austing [2].

5.1 Variance Gamma

In order to illustrate the proposed methodology for determining implied Lévy correlation estimates, we use the Dow Jones Industrial Average (DJ). The DJ is composed of 30 underlying stocks and for each underlying we have a finite number of option prices to which we can calibrate the parameter vector Θ and the Lévy volatility parameters σ_j. Using the available vanilla option data for June 20, 2008, we will work out the Gaussian and the Variance Gamma case.[4] Note that options on components of the Dow Jones are of American type. In the sequel, we assume that the American option price is a good proxy for the corresponding European option price. This assumption is justified because we use short term and out-of-the-money options.

The single volatility parameter σ_j is determined for stock j by minimizing the relative error between the model and the market vanilla option prices; see Algorithm 1. Assuming a normal distribution for L, this volatility parameter is denoted by σ_j^{BLS}, whereas the notation σ_j^{VG}, $j = 1, 2, \ldots, n$ is used for the VG model. For June 20, 2008, the parameter vector Θ for the VG copula model is given in Table 9 and the implied volatilities are listed in Table 8. Figure 4 shows the model (Gaussian and VG) and market prices for General Electric and IBM, both members of the Dow Jones, based on the implied volatility parameters listed in Table 8. We observe that the Variance Gamma copula model is more suitable in capturing the dynamics of the components of the Dow Jones than the Gaussian copula model.

Given the volatility parameters for the Variance Gamma case and the normal case, listed in Table 8, the implied correlation defined by Eq. (14) can be determined based on the available Dow Jones index options on June 20, 2008. For a given index strike K, the moneyness π is defined as $\pi = K/S(0)$. The implied Gaussian correlation (also called Black and Scholes correlation) is denoted by $\rho^{BLS}[\pi]$ and the corresponding implied Lévy correlation, based on a VG distribution, is denoted by $\rho^{VG}[\pi]$. In order to match the vanilla option curves more closely, we take into account the implied volatility smile and use a volatility parameter with moneyness π for each stock j, which we denote by $\sigma_j[\pi]$. For a detailed and step-by-step plan for the calculation of these volatility parameters, we refer to Linders and Schoutens [30].

Figure 5 shows that both the implied Black and Scholes and implied Lévy correlation depend on the moneyness π. However, for low strikes, we observe that $\rho^{VG}[\pi] < \rho^{BLS}[\pi]$, whereas the opposite inequality holds for large strikes, making the implied Lévy correlation curve less steep than its Black and Scholes counterpart. In Linders and Schoutens [30], the authors discuss the shortcomings of the implied Black and Scholes correlation and show that implied Black and Scholes correlations

[4] All data used for calibration are extracted from an internal database of the KU Leuven.

Table 8 Implied Variance Gamma volatilities σ_j^{VG} and implied Black and Scholes volatilities σ_j^{BLS} for June 20, 2008

Stock	σ_j^{VG}	σ_j^{BLS}
Alcoa Incorporated	0.6509	0.5743
American Express Company	0.4923	0.4477
American International Group	0.5488	0.4849
Bank of America	0.6003	0.5482
Boeing Corporation	0.3259	0.2927
Caterpillar	0.3009	0.2671
JP Morgan	0.5023	0.4448
Chevron	0.3252	0.3062
Citigroup	0.6429	0.5684
Coca Cola Company	0.2559	0.2343
Walt Disney Company	0.3157	0.2810
DuPont	0.2739	0.2438
Exxon Mobile	0.2938	0.2609
General Electric	0.3698	0.3300
General Motors	0.9148	0.8092
Hewlet–Packard	0.3035	0.2704
Home Depot	0.3604	0.3255
Intel	0.4281	0.3839
IBM	0.2874	0.2509
Johnson & Johnson	0.1741	0.1592
McDonald's	0.2508	0.2235
Merck & Company	0.3181	0.2896
Microsoft	0.3453	0.3068
3M	0.2435	0.2202
Pfizer	0.2779	0.2572
Procter & Gamble	0.1870	0.1671
AT&T	0.3013	0.2688
United Technologies	0.2721	0.2434
Verizon	0.3116	0.2847
Wal-Mart Stores	0.2701	0.2397

can become larger than one for low strike prices. Our more general approach and using the implied Lévy correlation solves this problem at least to some extent. Indeed, the region where the implied correlation stays below 1 is much larger for the flatter implied Lévy correlation curve than for its Black and Scholes counterpart. We also observe that near the at-the-money strikes, VG and Black and Scholes correlation estimates are comparable, which may be a sign that in this region, the use of implied Black and Scholes correlation (as defined in Linders and Schoutens [30]) is justi-

Table 9 Calibrated VG parameters for different trading days

			VG Parameters		
	$S(0)$	T (days)	σ	ν	θ
March 25, 2008	125.33	25	0.2981	0.5741	−0.1827
April 18, 2008	128.49	29	0.3606	0.5247	−0.2102
June 20, 2008	118.43	29	0.3587	0.4683	−0.1879
July 18, 2008	114.97	29	0.2639	0.5222	−0.1641
August 20, 2008	114.17	31	0.2467	0.3770	−0.1887

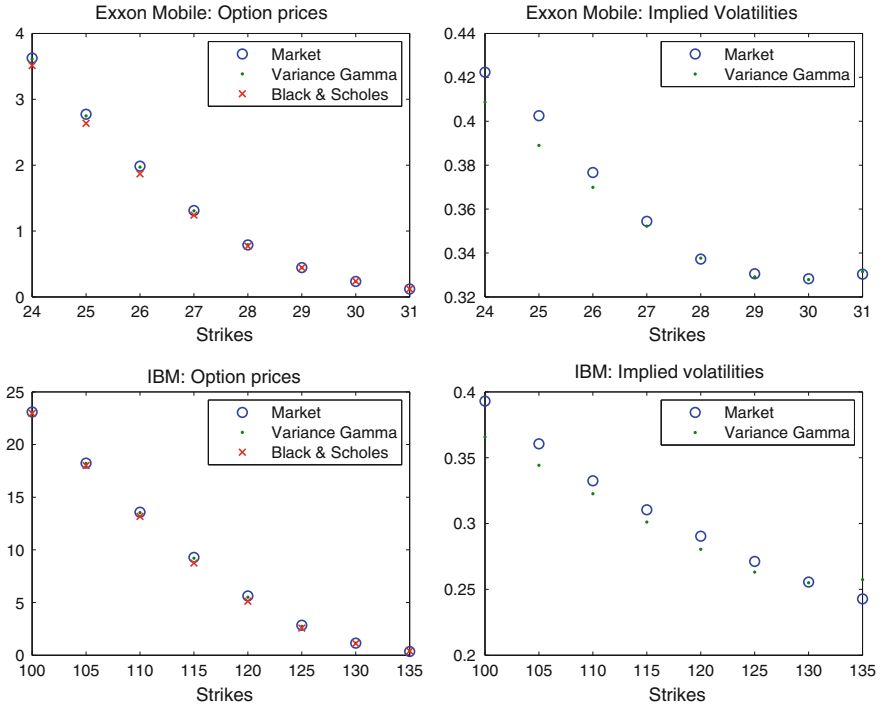

Fig. 4 Option prices and implied volatilities (model and market) for Exxon Mobile and IBM on June 20, 2008 based on the parameters listed in Table 8. The time to maturity is 30 days

fied. Figure 7 shows implied correlation curves for March, April, July and August, 2008. In all these situations, the time to maturity is close to 30 days. The calibrated parameters for each trading day are listed in Table 9.

We determine the implied correlation $\rho^{VG}[\pi]$ such that model and market quote for an index option with moneyness $\pi = K/S(0)$ coincide. However, the model price is determined using the three-moments-matching approximation and

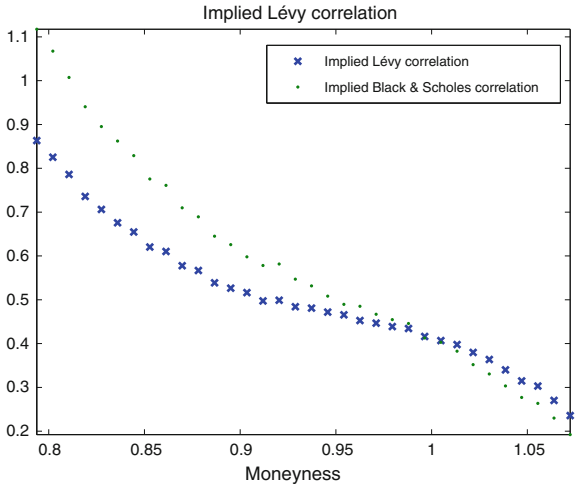

Fig. 5 Implied correlation smile for the Dow Jones, based on a Gaussian (*dots*) and a one-factor Variance Gamma model (*crosses*) for June 20, 2008

may deviate from the real model price. Indeed, we determine $\rho^{VG}[\pi]$ such that $C^{MM}\left[K, T; \underline{\sigma}, \Theta, \rho\,[\pi]\right] = C[K, T]$. In order to test if the implied correlation estimate obtained is accurate, we determine the model price $C^{mc}\left[K, T; \underline{\sigma}, \Theta, \rho\,[\pi]\right]$ using Monte Carlo simulation, where we plug in the volatility parameters and the implied correlation parameters. The results are listed in Table 10 and shown in Fig. 6. We observe that model and market prices are not exactly equal, but the error is still acceptable.

5.2 Double Exponential

In the previous subsection, we showed that the Lévy copula model allows for determining robust implied correlation estimates. However, calibrating this model can be a computational challenging task. Indeed, in case we deal with the Dow Jones Industrial Average, there are 30 underlying stocks and each stock has approximately 5 traded option prices. Calibrating the parameter vector Θ and the volatility parameters σ_j has to be done simultaneously. This contrasts sharply with the Gaussian copula model, where the calibration can be done stock per stock.

In this subsection we consider a model with the computational attractive calibration property of the Gaussian copula model, but without imposing any normality assumption on the marginal log returns. To be more precise, given the convincing arguments exposed in Fig. 7 we would like to keep L a $VG(\sigma, \nu, \theta, \mu)$ distribution. However, we do not calibrate the parameter vector $\Theta = (\sigma, \nu, \theta, \mu)$ to the vanilla option curves, but we fix these parameters upfront as follows

$$\mu = 0, \quad \theta = 0, \quad \nu = 1 \quad \text{and} \quad \sigma = 1.$$

Table 10 Market quotes for Dow Jones Index options for different basket strikes on June 20, 2008

Basket strikes	Market call prices	Implied VG correlation	VG call prices
94	24.45	0.8633	24.4608
95	23.45	0.8253	23.4794
96	22.475	0.786	22.4899
97	21.475	0.7358	21.4887
98	20.5	0.7062	20.5303
99	19.5	0.6757	19.5308
100	18.525	0.6546	18.5551
101	17.55	0.6203	17.5705
102	16.575	0.6101	16.6062
103	15.6	0.5778	15.6313
104	14.65	0.5668	14.6954
105	13.675	0.5386	13.7209
106	12.725	0.5266	12.7672
107	11.8	0.5164	11.8280
108	10.85	0.4973	10.8922
109	9.95	0.4989	9.9961
110	9.05	0.484	9.0813
111	8.2	0.4809	8.2202
112	7.35	0.4719	7.3519
113	6.525	0.4656	6.5193
114	5.7	0.4527	5.6755
115	4.95	0.4467	4.8908
116	4.225	0.4389	4.1554
117	3.575	0.4344	3.4788
118	2.935	0.4162	2.8118
119	2.375	0.4068	2.2337
120	1.88	0.3976	1.7227
121	1.435	0.3798	1.2977
122	1.065	0.3636	0.9549
123	0.765	0.3399	0.6906
124	0.52	0.3147	0.4793
125	0.36	0.3029	0.3517
126	0.22	0.2702	0.2321
127	0.125	0.2357	0.1479

For each price we find the corresponding implied correlation and the model price using a one-factor Variance Gamma model with parameters listed in Table 9

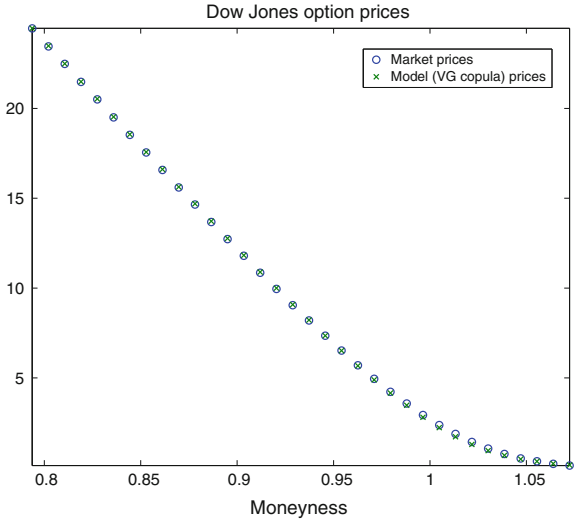

Fig. 6 Dow Jones option prices: Market prices (*circles*) and the model prices using a one-factor Variance Gamma model and the implied VG correlation smile (*crosses*) for June 20, 2008

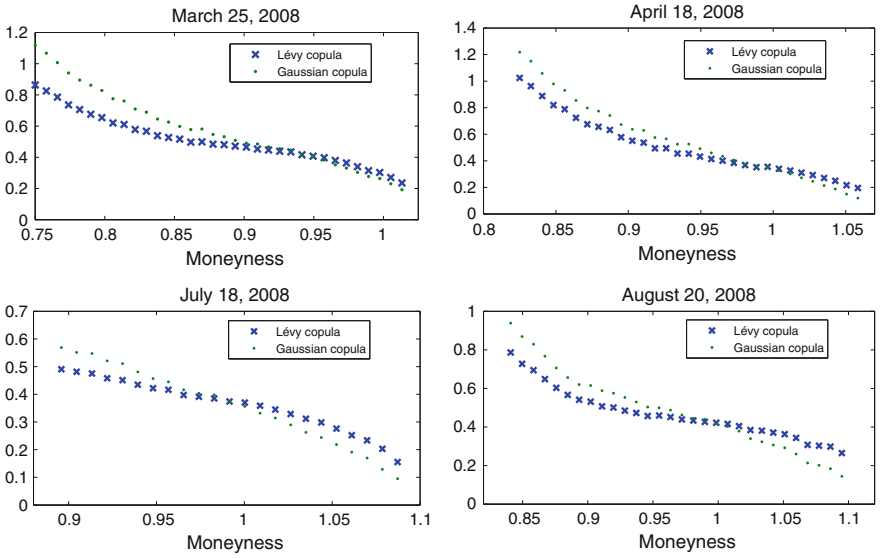

Fig. 7 Implied correlation smile for the Dow Jones, based on a Gaussian (*dots*) and a one-factor Variance Gamma model (*crosses*) for different trading days

In this setting, L is a standardized distribution and its characteristic function ϕ_L is given by

$$\phi_L(u) = \frac{1}{1 + \frac{u^2}{2}}, \quad u \in \mathbb{R}.$$

From its characteristic function, we see that L has a *Standard Double Exponential distribution*, also called Laplace distribution, and its pdf f_L is given by

$$f_L(u) = \frac{\sqrt{2}}{2} \mathrm{e}^{-\frac{|u|}{\sqrt{2}}}$$

The Standard Double Exponential distribution is symmetric and centered around zero, while it has variance 1. Note, however, that it is straightforward to generalize this distribution such that it has center μ and variance σ^2. Moreover, the kurtosis of this Double Exponential distribution is 6.

By using the Double Exponential distribution instead of the more general Variance Gamma distribution, some flexibility is lost for modeling the marginals. However, the Double Exponential distribution is still a much better distribution for modeling the stock returns than the normal distribution. Moreover, in this simplified setting, the only parameters to be calibrated are the marginal volatility parameters,

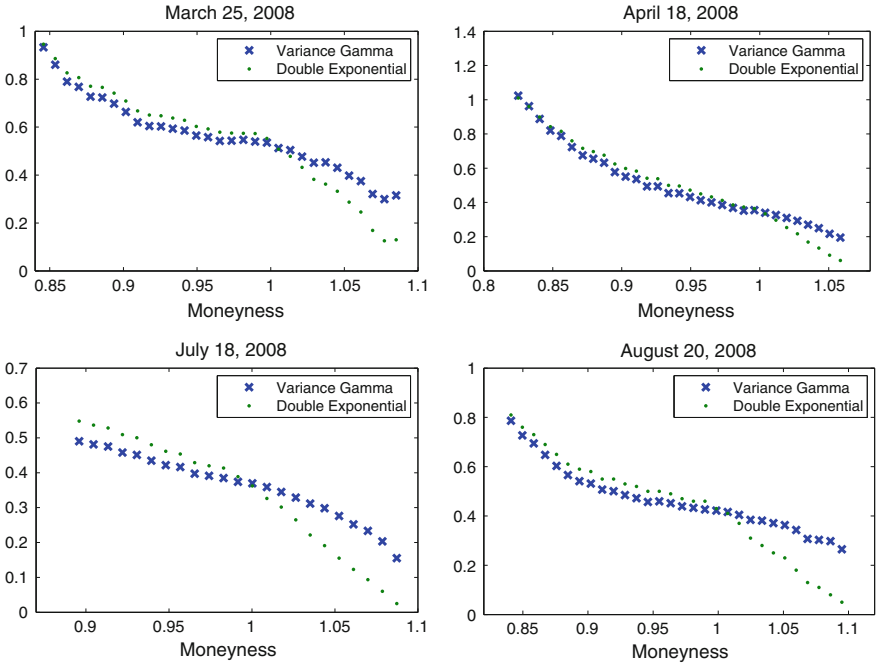

Fig. 8 Implied correlation smiles in the one-factor Variance Gamma and the Double Exponential model

which we denote by σ_j^{DE}, and the correlation parameter ρ^{DE}. Similar to the Gaussian copula model, calibrating the volatility parameter σ_j^{DE} only requires the option curve of stock j. As a result, the time to calibrate the Double Exponential copula model is comparable to its Gaussian counterpart and much shorter than the general Variance Gamma copula model.

Consider the DJ on March 25, 2008. The time to maturity is 25 days. We determine the implied marginal volatility parameter for each stock in a one-factor Variance Gamma model and a Double Exponential framework. Given this information, we can determine the prices $C^{VG}[K,T]$ and $C^{DE}[K,T]$ for a basket option in a Variance-Gamma and a Double Exponential model, respectively. Figure 8 shows the implied Variance Gamma and the Double Exponential correlations. We observe that the implied correlation based on a one-factor VG model is larger than its Double Exponential counterpart for a moneyness bigger than one, whereas both implied correlation estimates are relatively close to each other in the other situation.

6 Conclusion

In this paper we introduced a one-factor Lévy model and we proposed a three-moments-matching approximation for pricing basket options. Well-known distributions like the Normal, Variance Gamma, NIG, Meixner, etc., can be used in this one-factor Lévy model. We calibrate these different models to market data and determine basket option prices for the different model settings. Our newly designed (approximate) basket option pricing formula can be used to define implied Lévy correlation. The one-factor Lévy model provides a flexible framework for deriving implied correlation estimates in different model settings. Indeed, by employing a Brownian motion and a Variance Gamma process in our model, we can determine Gaussian and VG-implied correlation estimates, respectively. We observe that the VG implied correlation is an improvement of the Gaussian-implied correlation.

Acknowledgements The authors acknowledge the financial support of the Onderzoeksfonds KU Leuven (GOA/13/002: Management of Financial and Actuarial Risks: Modeling, Regulation, Disclosure and Market Effects). Daniël Linders also acknowledges the support of the AXA Research Fund (Measuring and managing herd behavior risk in stock markets). The authors also thank Prof. Jan Dhaene, Prof. Alexander Kukush, the anonymous referees and the editors for helpful comments.

The KPMG Center of Excellence in Risk Management is acknowledged for organizing the conference "Challenges in Derivatives Markets - Fixed Income Modeling, Valuation Adjustments, Risk Management, and Regulation".

Appendix: Proof of Lemma 1

The proof for expression (9) is straightforward.

Starting from the multinomial theorem, we can write the second moment m_2 as follows

$$\begin{aligned} m_2 &= \mathbb{E}\left[(w_1 S_1(T) + w_2 S_2(T) + \ldots w_n S_n(T))^2\right] \\ &= \mathbb{E}\left[\sum_{i_1+i_2+\ldots+i_n=2} \frac{2}{i_1! i_2! \ldots i_n!} \prod_{j=1}^{n} \left(w_j S_j(T)\right)^{i_j}\right]. \end{aligned}$$

Considering the cases $(i_n = 0)$, $(i_n = 1)$ and $(i_n = 2)$ separately, we find

$$m_2 = \mathbb{E}\left[\left(\sum_{j=1}^{n-1} w_j S_j(T)\right)^2 + 2 w_n S_n(T) \sum_{j=1}^{n-1} w_j S_j(T) + w_j^2 S_n^2(T)\right].$$

Continuing recursively gives

$$m_2 = \sum_{j=1}^{n}\sum_{k=1}^{n} w_j w_k \mathbb{E}\left[S_j(T) S_k(T)\right]. \tag{15}$$

We then find that

$$\begin{aligned} m_2 &= \sum_{j=1}^{n}\sum_{k=1}^{n} w_j w_k S_j(0) S_k(0) \\ &\quad \times \mathbb{E}\left[\exp\left\{(2r - q_j - q_k - \omega_j - \omega_k)T + (\sigma_j A_j + \sigma_k A_k)\sqrt{T}\right\}\right] \\ &= \sum_{j=1}^{n}\sum_{k=1}^{n} w_j w_k \frac{\mathbb{E}\left[S_j(T)\right]\mathbb{E}\left[S_k(T)\right]}{\phi_L\left(-\mathrm{i}\sigma_j\sqrt{T}\right)\phi_L\left(-\mathrm{i}\sigma_k\sqrt{T}\right)} \mathbb{E}\left[\exp\left\{(\sigma_j A_j + \sigma_k A_k)\sqrt{T}\right\}\right]. \end{aligned}$$

In the last step, we used the Expression $\omega_j = \log \phi_L\left(\mathrm{i}\sigma_j\sqrt{T}\right)/T$. If we use expression (1) to decompose A_j and A_k in the common component $X(\rho)$ and the independent components $X_j(1-\rho)$ and $X_k(1-\rho)$, we find the following expression for m_2

$$m_2 = \sum_{j=1}^{n}\sum_{k=1}^{n} w_j w_k \frac{\mathbb{E}\left[S_j(T)\right]\mathbb{E}\left[S_k(T)\right]}{\phi_L\left(-\mathrm{i}\sigma_j\sqrt{T}\right)\phi_L\left(-\mathrm{i}\sigma_k\sqrt{T}\right)} \mathbb{E}\left[\mathrm{e}^{(\sigma_j+\sigma_k)X(\rho)} \mathrm{e}^{\sigma_j\sqrt{T}X_j(1-\rho)} \mathrm{e}^{\sigma_k\sqrt{T}X_k(1-\rho)}\right].$$

The r.v. $X(\rho)$ is independent from $X_j(1-\rho)$ and $X_k(1-\rho)$. Furthermore, the characteristic function of $X(\rho)$ is ϕ_L^{ρ}, which results in

$$m_2 = \sum_{j=1}^{n}\sum_{k=1}^{n} w_j w_k \frac{\mathbb{E}\left[S_j(T)\right]\mathbb{E}\left[S_k(T)\right]}{\phi_L\left(-\mathrm{i}\sigma_j\sqrt{T}\right)\phi_L\left(-\mathrm{i}\sigma_k\sqrt{T}\right)}\phi_L\left(-\mathrm{i}(\sigma_j+\sigma_k)\sqrt{T}\right)^{\rho} \times \mathbb{E}\left[\mathrm{e}^{\sigma_j\sqrt{T}X_j(1-\rho)}\mathrm{e}^{\sigma_k\sqrt{T}X_k(1-\rho)}\right].$$

If $j \neq k$, $X_j(1-\rho)$ and $X_k(1-\rho)$ are i.i.d. with characteristic function $\phi_L^{1-\rho}$, which gives the following expression for m_2:

$$m_2 = \sum_{j=1}^{n}\sum_{k=1}^{n} w_j w_k \mathbb{E}\left[S_j(T)\right]\mathbb{E}\left[S_k(T)\right]\left(\frac{\phi_L\left(-\mathrm{i}(\sigma_j+\sigma_k)\sqrt{T}\right)}{\phi_L\left(-\mathrm{i}\sigma_j\sqrt{T}\right)\phi_L\left(-\mathrm{i}\sigma_k\sqrt{T}\right)}\right)^{\rho}.$$

If $j = k$, we find that

$$\mathbb{E}\left[\mathrm{e}^{\sigma_j\sqrt{T}X_j(1-\rho)}\mathrm{e}^{\sigma_k\sqrt{T}X_k(1-\rho)}\right] = \phi_L\left(-\mathrm{i}\left(\sigma_j+\sigma_k\right)\sqrt{T}\right),$$

which gives

$$m_2 = \sum_{j=1}^{n}\sum_{k=1}^{n} w_j w_k \mathbb{E}\left[S_j(T)\right]\mathbb{E}\left[S_k(T)\right]\frac{\phi_L\left(-\mathrm{i}(\sigma_j+\sigma_k)\sqrt{T}\right)}{\phi_L\left(-\mathrm{i}\sigma_j\sqrt{T}\right)\phi_L\left(-\mathrm{i}\sigma_k\sqrt{T}\right)}.$$

This proves expression (10) for m_2.

We can write m_3 as follows

$$\begin{aligned} m_3 &= \mathbb{E}\left[\left(\sum_{j=1}^{n} w_j S_j(T)\right)^3\right] \\ &= \mathbb{E}\left[\left(\sum_{j=1}^{n} w_j S_j(T)\right)^2 \sum_{l=1}^{n} w_l S_l(T)\right]. \end{aligned}$$

Using expression (15), we find the following Expression for m_3:

$$\begin{aligned} m_3 &= \mathbb{E}\left[\left(\sum_{j=1}^{n}\sum_{k=1}^{n} w_j w_k S_j(T)S_k(t)\right)\sum_{l=1}^{n} w_l S_l(T)\right] \\ &= \sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} w_j w_k w_l \mathbb{E}\left[S_j(T)S_k(T)S_l(T)\right]. \end{aligned}$$

Similar calculations as for m_2 result in

$$m_3 = \sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} w_j w_k w_l \mathbb{E}\left[S_j(T)\right]\mathbb{E}\left[S_k(T)\right]\mathbb{E}\left[S_l(T)\right] \times \frac{\phi_L\left(-\mathrm{i}(\sigma_j+\sigma_k+\sigma_l)\sqrt{T}\right)^{\rho}}{\phi_L\left(-\mathrm{i}\sigma_j\sqrt{T}\right)\phi_L\left(-\mathrm{i}\sigma_k\sqrt{T}\right)\phi_L\left(-\mathrm{i}\sigma_l\sqrt{T}\right)} A_{j,k,l},$$

where

$$A_{j,k,l} = \mathbb{E}\left[\mathrm{e}^{\sigma_j\sqrt{T}X_j(1-\rho)}\mathrm{e}^{\sigma_k\sqrt{T}X_k(1-\rho)}\mathrm{e}^{\sigma_l\sqrt{T}X_l(1-\rho)}\right].$$

Differentiating between the situations $(j=k=l)$, $(j=k, k\neq l)$, $(j\neq k, k=l)$, $(j\neq k, k\neq l, j=l)$ and $(j\neq k\neq l, j\neq l)$, we find expression (11).

References

1. Albrecher, H., Ladoucette, S., Schoutens, W.: A generic one-factor Lévy model for pricing synthetic CDOs. In: Fu, M., Jarrow, R., Yen, J.-Y., Elliott, R. (eds.) Advances in Mathematical Finance, pp. 259–277. Applied and Numerical Harmonic Analysis, Birkhäuser Boston (2007)
2. Austing, P.: Smile Pricing Explained. Financial Engineering Explained. Palgrave Macmillan (2014)
3. Ballotta, L., Bonfiglioli, E.: Multivariate asset models using Lévy processes and applications. Eur. J. Financ. http://dx.doi.org/10.1080/1351847X.2013.870917 (2014)
4. Ballotta, L., Deelstra, G., Rayée, G.: Extracting the implied correlation from quanto derivatives, Technical report. working paper (2014)
5. Baxter, M.: Lévy simple structural models. Int. J. Theor. Appl. Financ. (IJTAF) **10**(04), 593–606 (2007)
6. Black, F., Scholes, M.: The pricing of options and corporate liabilities. J. Polit. Econ. **81**(3), 637–654 (1973)
7. Brigo, D., Mercurio, F., Rapisarda, F., Scotti, R.: Approximated moment-matching dynamics for basket-options pricing. Quant. Financ. **4**(1), 1–16 (2004)
8. Brooks, R., Corson, J., Wales, J.D.: The pricing of index options when the underlying assets all follow a lognormal diffusion. Adv. Futures Options Res. **7** (1994)

9. Caldana, R., Fusai, G.: A general closed-form spread option pricing formula. J. Bank. Financ. **37**(12), 4893–4906 (2013)
10. Caldana, R., Fusai, G., Gnoatto, A., Graselli, M.: General close-form basket option pricing bounds. Quant. Financ. http://dx.doi.org/10.2139/ssrn.2376134 (2014)
11. Carmona, R., Durrleman, V.: Pricing and hedging spread options. SIAM Rev. **45**(4), 627–685 (2003)
12. Carmona, R., Durrleman, V.: Generalizing the Black–Scholes formula to multivariate contingent claims. J. Comput. Financ. **9**, 43–67 (2006)
13. Carr, P., Madan, D.B.: Option valuation using the Fast Fourier Transform. J. Comput. Financ. **2**, 61–73 (1999)
14. Cherubini, U., Luciano, E., Vecchiato, W.: Copula Methods in Finance. The Wiley Finance Series. Wiley (2004)
15. Chicago Board Options Exchange: CBOE S&P 500 implied correlation index. Working Paper (2009)
16. Cont, R., Deguest, R.: Equity correlations implied by index options: estimation and model uncertainty analysis. Math. Financ. **23**(3), 496–530 (2013)
17. Corcuera, J.M., Guillaume, F., Leoni, P., Schoutens, W.: Implied Lévy volatility. Quant. Financ. **9**(4), 383–393 (2009)
18. Deelstra, G., Liinev, J., Vanmaele, M.: Pricing of arithmetic basket options by conditioning. Insur. Math. Econ. **34**(1), 55–77 (2004)
19. Fonseca, J., Grasselli, M., Tebaldi, C.: Option pricing when correlations are stochastic: an analytical framework. Rev. Deriv. Res. **10**(2), 151–180 (2007)
20. Garcia, J., Goossens, S., Masol, V., Schoutens, W.: Lévy base correlation. Wilmott J. **1**, 95–100 (2009)
21. Guillaume, F.: The αVG model for multivariate asset pricing: calibration and extension. Rev. Deriv. Res. **16**(1), 25–52 (2013)
22. Guillaume, F., Jacobs, P., Schoutens, W.: Pricing and hedging of CDO-squared tranches by using a one factor Lévy model. Int. J. Theor. Appl. Financ. **12**(05), 663–685 (2009)
23. Hull, J., White, S.: Efficient procedures for valuing European and American path-dependent options. J. Deriv. **1**(1), 21–31 (1993)
24. Hurd, T.R., Zhou, Z.: A Fourier transform method for spread option pricing. SIAM J. Fin. Math. **1**(1), 142–157 (2010)
25. Kallsen, J., Tankov, P.: Characterization of dependence of multidimensional Lévy processes using Lévy copulas. J. Multivar. Anal. **97**(7), 1551–1572 (2006)
26. Kawai, R.: A multivariate Lévy process model with linear correlation. Quant. Financ. **9**(5), 597–606 (2009)
27. Korn, R., Zeytun, S.: Efficient basket Monte Carlo option pricing via a simple analytical approximation. J. Comput. Appl. Math. **243**(1), 48–59 (2013)
28. Leoni, P., Schoutens, W.: Multivariate smiling. Wilmott Magazin. 82–91 (2008)
29. Linders, D.: Pricing index options in a multivariate Black & Scholes model, Research report AFI-1383 FEB. KU Leuven—Faculty of Business and Economics, Leuven (2013)
30. Linders, D., Schoutens, W.: A framework for robust measurement of implied correlation. J. Comput. Appl. Math. **271**, 39–52 (2014)
31. Linders, D., Stassen, B.: The multivariate Variance Gamma model: basket option pricing and calibration. Quant. Financ. http://dx.doi.org/10.1080/14697688.2015.1043934 (2015)
32. Luciano, E., Schoutens, W.: A multivariate jump-driven financial asset model. Quant. Financ. **6**(5), 385–402 (2006)
33. Luciano, E., Semeraro, P.: Multivariate time changes for Lévy asset models: characterization and calibration. J. Comput. Appl. Math. **233**, 1937–1953 (2010)
34. Madan, D.B., Seneta, E.: The Variance Gamma (V.G.) model for share market returns. J. Bus. **63**(4), 511–524 (1990)
35. Madan, D.B., Carr, P., Chang, E.C.: The Variance Gamma process and option pricing. Eur. Financ. Rev. **2**, 79–105 (1998)

36. Mai, J.-F., Scherer, M., Zagst, R.: CIID frailty models and implied copulas. In: Proceedings of the workshop Copulae in Mathematical and Quantitative Finance, Cracow, 10-11 July 2012, Springer, pp. 201–230 (2012)
37. Masol, V., Schoutens, W.: Comparing alternative Lévy base correlation models for pricing and hedging CDO tranches. Quant. Financ. **11**(5), 763–773 (2011)
38. McWilliams, N.: Option pricing techniques understochastic delay models. Ph.D. thesis, University of Edinburgh (2011)
39. Milevsky, M., Posner, S.: A closed-form approximation for valuing basket options. J. Deriv. **5**(4), 54–61 (1998)
40. Moosbrucker, T.: Explaining the correlation smile using Variance Gamma distributions. J. Fixed Income **16**(1), 71–87 (2006)
41. Moosbrucker, T.: Pricing CDOs with correlated variance gamma distributions, Technical report, Centre for Financial Research, Univ. of Cologne. colloquium paper (2006)
42. Rubinstein, M.: Implied binomial trees. J. Financ. **49**(3), 771–818 (1994)
43. Schoutens, W.: Lévy Processes in Finance: Pricing Financial Derivatives. Wiley (2003)
44. Semeraro, P.: A multivariate Variance Gamma model for financial applications. Int. J. Theor. Appl. Financ. (IJTAF) **11**(01), 1–18 (2008)
45. Skintzi, V.D., Refenes, A.N.: Implied correlation index: a new measure of diversification. J. Futures Mark. **25**, 171–197 (2005). doi:10.1002/fut.20137
46. Tavin, B.: Hedging dependence risk with spread options via the power frank and power student t copulas, Technical report, Université Paris I Panthéon-Sorbonne. Available at SSRN: http://ssrn.com/abstract=2192430 (2013)
47. Vasicek, O.: Probability of loss on a loan portfolio. KMV Working Paper (1987)
48. Xu, G., Zheng, H.: Basket options valuation for a local volatility jump diffusion model with the asymptotic expansion method. Insur. Math. Econ. **47**(3), 415–422 (2010)
49. Xu, G., Zheng, H.: Lower bound approximation to basket option values for local volatility jump-diffusion models. Int. J. Theor. Appl. Financ. **17**(01), 1450007 (2014)

Pricing Shared-Loss Hedge Fund Fee Structures

Ben Djerroud, David Saunders, Luis Seco and Mohammad Shakourifar

Abstract The asset management business is driven by fee structures. In the context of hedge funds, fees have usually been a hybrid combination of two different types, which has coined a well-known business term of "2 and 20". As an attempt to provide better alignment with their investors, in a new context of low interest rates and lukewarm performance, a new type of fund fees has been introduced in the last few years that offers a more symmetric payment structure, which we will refer to as *shared loss*. In this framework, in return for receiving performance fees, the fund manager provides some downside protection against losses to the investors. We show that the position values of the investor and the hedge fund manager can be formulated as portfolios of options, and discuss issues regarding pricing and fairness of the fee rates, and incentives for both investors and hedge fund managers. In particular, we will be able to show that, from a present value perspective, these fee structures can be set up as being favorable either to the hedge fund manager or to the investor. The paper is based on an arbitrage-free pricing framework. However, if one is to take

This research was supported in part by the Natural Sciences and Engineering Research Council of Canada.

B. Djerroud · M. Shakourifar (✉)
Sigma Analysis & Management, Toronto, ON, Canada
e-mail: mohammad@sigmanalysis.com

B. Djerroud
e-mail: ben_d@sigmanalysis.com

D. Saunders
Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Canada
e-mail: dsaunders@uwaterloo.ca

L. Seco
Department of Mathematics, University of Toronto, Toronto, Canada
e-mail: seco@math.utoronto.ca

K. Glau et al. (eds.), *Innovations in Derivatives Markets*, Springer Proceedings in Mathematics & Statistics 165, DOI 10.1007/978-3-319-33446-2_17

into account the value to the business that investor capital brings to a fund, which is not part of our framework, it is possible to create a situation where both investors as well as asset managers win.

Keywords Hedge funds · Fee structures · First-loss · Shared-loss · Black-Scholes option pricing

1 Introduction

Hedge Funds are pooled investment vehicles overseen by a management company. They generally aim at absolute return portfolios and their success is usually linked to market inefficiencies, such as instrument mispricing, misguided market consensus or, in general terms, the manager's intelligence to anticipate market moves. The nature of these investments is that they exploit investment opportunities that are rare. This is a characteristic that they share with private equity investments, but they share with the mutual fund industry the fact that they often trade in liquid, marketable securities. Fund sizes are more in line with private equity investing than with the mammoth mutual fund industry. Their compensation structure, because of their limited access to opportunity, is also more in line with the private equity universe, and usually consists in a fixed, asset-based fee, and a variable, performance fee base. Because of market conditions that have been in place over the last several years, in particular the low interest rate environment, coupled with the lukewarm performance of the hedge fund sector in the recent years, investors have become increasingly more sensitive to fee structures. The traditional 2&20 fee structure, consisting of a flat fee of 2 % of assets under management together with a performance fee of 20 % of net profits is considered unfair on the basis of the asymmetry: the management company will always earn a fee, whereas the investor is only guaranteed to pay that fee. The advent of the 40-ACT funds[1] has, in particular, dispensed with the performance fee base in favor of a fixed management fee, which is more in line with the mutual fund industry than with the hedge fund industry. This compensation model essentially rewards funds for becoming asset gatherers instead of the alpha-seeking business the hedge fund was set out to be. In this paper we will examine, from a quantitative perspective, a suite of symmetric performance fee structures which are gaining traction with more sophisticated investors, known as first-loss (or shared-loss) fee structures. In this new framework, in return for receiving performance fees, the fund manager provides some downside protection against losses to the investors.

The issue of the incentives created by hedge fund fees bears much similarity to issues surrounding the structure of executive compensation. At first glance, the optionality inherent in both would seem to incentivize greater risk taking. However, the reality is more subtle. Carpenter [2] studies the case of executive compensation,

[1] Pooled investment vehicles, enforced and regulated by the Securities and Exchange Commission, that are packaged and sold to retail and institutional investors in the public markets.

when the manager cannot hedge options provided as compensation by trading the underlying. In certain conditions, a utility-maximizing manager may choose to reduce rather than increase the volatility of the underlying firm. Ross [9] gives necessary and sufficient conditions for a fee schedule to make a utility-maximizing manager more or less risk-averse. Hodder and Jackwerth [6] consider the effects of hedge fund fee incentives on a risk manager with power utility, and also in the presence of a liquidation barrier. They find that over a one-year horizon, risk-taking varies dramatically with fund value, but that this effect is moderated over longer time horizons. Kouwenberg and Ziemba [7] consider loss-averse hedge fund managers and find that higher incentive fees lead to riskier fund management strategies. However, this effect is reduced if a significant portion of the manager's own money is invested in the fund. They further provide empirical evidence showing that hedge funds with incentive fees have significantly lower mean returns (net of fees), and find a positive correlation between fee levels and downside risk. They find that risk is increasing with respect to the performance fee if the manager's objective function is based on cumulative prospect theory, rather than utility, and provide empirical evidence. Recent work on the analysis of hedge fund fee structures includes that of Goetzmann et al. [3], who value a fee structure with a highwater mark provision, using a PDE approach with a fixed investment portfolio, Panageas and Westerfield [8], who consider the portfolio selection decision of maximizing the present value of fees for a risk-neutral manager over an infinite horizon, and Guasoni and Obłój [4], who extend this work to managers with risk-averse power utility. Closest to the current work is He and Kou [5], who analyze shared-loss fee structures for hedge funds by looking at the portfolio selection decision of a hedge fund manager whose preferences are modeled using cumulative prospect theory. The problem is considered in the presence of a manager investing in the fund, and with a predetermined liquidation barrier. Analytical solutions of the portfolio selection problem are provided, and the result (cumulative prospect theory) for both the investor and the manager is examined. It is found that depending on the parameter values, either a traditional fee structure or a first-loss fee structure may result in a riskier investment strategy. While for some parameter values, the first-loss structure improves the utility of both the investor and the hedge fund manager, they find that for typical values, the manager is better off, while the investor is worse off. In this paper, we investigate the shared-loss fee structures from the perspective of risk-neutral valuation, with no further assumptions about investor preferences, while He and Kou [5] solve the stochastic control problem (under the real-world measure) corresponding to the manager maximizing the utility function from cumulative prospect theory, and also evaluate the investor's payoff using the same type of criterion.

The paper is organized as follows. First, we will review the traditional fee structures in some detail. Next, we will introduce the notion and mechanics of the first-loss structures, and a framework for a fee pricing based on the theory of option price valuation. After that, we will introduce the concept of net fee, a number that will allow us to determine whether the investor or the management company is the net winner in a given fee agreement. Finally, we will present a set of computational examples that will display the net fee as a function of the agreement and market variables.

2 Hedge Fund Fees

The hedge fund manager charges two types of fees to the fund investors:

- A fixed management fee, usually ranging from 1 % to 2 % of net asset values.
- A performance fee, most commonly equal to 20 % of net profits obtained by the fund.

In this paper we assume a single investor and a single share issued by the fund. The extension to the case of multiple investors and multiple shares is straightforward. Although fees are paid according to a determined schedule (usually monthly or quarterly for management fees and annual for performance fees), we will assume a single payment at the end of a fixed term T.

The fund value evolution and fee payment mechanics are denoted as follows: the initial fund supplied by the investor is X_0. The hedge fund manager then invests fund assets to create future gross values X_t, for $t > 0$. The gross fund value X_t is split between the investor's worth I_t (the net asset value) and the manager's fee M_t:

$$X_t = I_t + M_t.$$

At time 0, $X_0 = I_0$ and $M_0 = 0$.

There are countless variations to this basic framework, including hurdles, claw-backs, etc. (for more details on first-loss arrangements see Banzaca [1]). We will ignore those and assume the commonly used version of a management fee equal to $m \cdot X_0$ (m represents a fixed percentage of the initial investment by the investor), and a performance fee of

$$\alpha \cdot (X_T - (1+m)X_0))_+ ,$$

payable only when it is positive, and equal to zero when it is negative. Hence,

$$M_T = m \cdot X_0 + \alpha \cdot (X_T - (1+m)X_0)_+) \tag{1}$$

In other words, while the management fee is a fixed future liability to the investor, the performance fee is a contingent claim on the part of the manager. As a consequence, we will be pricing the management fee simply as a fixed guaranteed fee with a predetermined future cash value, and we will be valuing the performance fee as the value of a certain call option. In our setting, we will assume normally distributed log-returns for the invested assets X_t, which allows us to value the performance fee in the Black–Scholes framework. It is worth mentioning that hedge funds managers can speculate on volatility, credit risks, etc. and in contrast to the traditional money managers, they can go long and short. The diversity in investment styles and the different levels of gross and net exposure that they can employ could result in leptokurtic (non-normal) properties in their returns, which is revealed through frequent large negative

returns to the left of the return distribution. Generalization of the current framework to models that account for non-normality of the hedge fund returns, for example by employing generalized autoregressive conditional heteroskedasticity (GARCH) models, could be a subject for future research.

3 The First-Loss Model

Calpers announced in 2014 that they were exiting hedge fund investments WSJ [10]. While not the main stated reason for their decision, one they mentioned was high fees payable to their hedge fund managers, something that has caught the attention of investors worldwide in the contemporary context of a widely accepted notion that hedge fund fees nowadays are too high. Certain hedge funds are reacting to this shifting balance of power between the sell-side and the buy-side of the investment business with the creation of innovative fee structures which still reward the intellectual capital of the hedge fund manager and allow for business growth but at the same time offer the investor a more symmetric compensation structure.

An example of a first-loss structure is the following:

- The investor provides an investment of $100M to a fund.
- The fund manager will absorb the first-loss up to 10 % of the initial investment.
- The investor pays a management fee of 1 % to the manager, and performance fee of 50 %.

In our paper we will present a quantitative comparison of the fees payable to the manager and the risk-neutral valuation of the guarantee offered to the investor. We want to note, for the sake of completeness, that there are many other qualitative considerations which are relevant when analyzing both the fee structure as well as the business value offered to a management company by the investor, which are not the objective of this paper. In fact, hedge fund start-ups have become more difficult in recent times, increasing value to any investor action that allows a hedge fund business to succeed. That value is linked to a wide variety of fund characteristics, including the size of assets under management (AUM), the track record, or historical performance, and the reputation of its investor base, among others.

In addition to the initial investment X_0, the management fee m and the performance fee α, payable at a fixed time horizon T, we will now also consider a deposit amount c, as a percentage of the initial investment X_0, which the manager will provide as a guarantee for losses. Our objective is to analyze the relationship between all four variables to determine whether the investor, or the manager, is the net winner of value-add from a risk-neutral valuation perspective.

4 An Option Pricing Framework

The fund value X_t is split between the investors I_t and the manager M_t, where, $X_t = I_t + M_t$. In the following sub-sections we derive the payoff function of each player separately, and then price the positions accordingly.

4.1 Payoff to the Investor

The payoff to the investor at the terminal time T is:

$$I_T = \begin{cases} X_T - mX_0 - \alpha(X_T - mX_0 - X_0) & \text{when } X_T - mX_0 \geq X_0 \\ X_0 & \text{when } (1-c)X_0 \leq X_T - mX_0 \leq X_0 \\ X_T + (c-m)X_0 & \text{when } X_T - mX_0 \leq (1-c)X_0 \end{cases}$$

or, writing the payoff in a more compact form:

$$\begin{array}{rcl} I_T = & X_T - mX_0 & \text{(pays a management fee)} \\ & -\alpha(X_T - mX_0 - X_0)_+ & \text{(pays a performance fee)} \\ & +(X_0 - X_T + mX_0)_+ - ((1-c)X_0 - X_T + mX_0)_+ & \text{(receives a guarantee)} \end{array}$$

Thus, we see that the position of the investor is equivalent to the following portfolio:

- A position in the hedge fund assets, with initial investment X_0, less management fee, that is, $X_0 - mX_0$.
- A short position in α call options on the hedge fund assets, with strike price $X_0 + mX_0$ (the performance fee, or performance call option, given to the hedge fund manager).
- A long position in a put option on the fund assets, with the strike price $X_0 + mX_0$ (the insurance put option).
- A short position in a put option on the fund assets, with strike price $(1-c)X_0 + mX_0$ (yielding a cap on the insurance payment).

4.2 Payoff to the Manager

The payoff to the manager is $M_T = X_T - I_T$. In other words, the payoff to the hedge fund manager results from the manager having the opposite position in all of the options of the investor. More explicitly,

$$\begin{aligned} M_T = \quad & mX_0 && \text{(receives a management fee)} \\ & +\alpha(X_T - mX_0 - X_0)_+ && \text{(receives a performance fee)} \\ & -(X_0 - X_T + mX_0)_+ + ((1-c)X_0 - X_T + mX_0)_+ && \text{(provides a guarantee)} \end{aligned}$$

which implies that the hedge fund manager has a portfolio of options consisting of:

- A constant position in the fixed management fee of mX_0.
- A long position in α call options on the hedge fund assets, with strike price $X_0 + mX_0$.
- A short position in a put option on the fund assets, with the strike price $X_0 + mX_0$.
- A long position in a put option on the fund assets, with strike price $(1-c)X_0 + mX_0$.

Note that net income to the management company is now no longer guaranteed to be positive. In addition, since the options trades constitute a zero-sum game (the positions of the manager and the investor are opposite each other), the sum of the investor payoff and the manager payoff is equal to X_T.

4.3 Valuation: Pricing Fees as Derivatives

In this section, we will value the positions of the investor and the hedge fund manager using a simple Black–Scholes model for the underlying fund value process. In particular, we employ risk-neutral valuation, and assume that under the risk-neutral probabilities, the fund value process satisfies the stochastic differential equation:

$$dX_t = rX_t\,dt \;+\; \sigma\, X_t\, dW_t, \tag{2}$$

with solution:

$$X_t = X_0 \exp\left((r - \tfrac{\sigma^2}{2})t + \sigma W_t\right) \tag{3}$$

where W_t is a standard Brownian motion, and r and σ are positive constants, giving the continuously compounded risk-free interest rate and the volatility of the hedge fund assets respectively. It should be noted that the Black–Scholes framework is applicable to our context as the underlying, that is the fund value, can be dynamically traded. Moreover, in a managed account context, even the liquidity of the fund can be made to match the liquidity of the underlying traded securities.

The Black–Scholes formula can be used to derive the price of the investor's position under the Black–Scholes model:

$$\begin{aligned} V_I(0) = {} & X_0 - e^{-rT} mX_0 - \alpha C(X_0, T, X_0 + mX_0, r, \sigma) \\ & + P(X_0, T, X_0 + mX_0, r, \sigma) - P(X_0, T, (1-c)X_0 + mX_0, r, \sigma) \end{aligned} \tag{4}$$

where $C(X, T, K, r, \sigma)$ is the Black–Scholes price of a call option on a non-dividend paying asset with current value of the underlying X, time to expiration T, strike price

K, risk-free interest rate r and volatility σ, and $P(X, T, K, r, \sigma)$ is the Black–Scholes put option price with the same parameters as arguments.

5 Consequences of the Derivative Pricing Framework

5.1 Graphical Analysis

To compare and contrast the traditional and shared-loss fee structures, in our base case we take the investment horizon to be one month, that is $T = 1/12$, the performance fee $\alpha = 50\,\%$, the manager deposit $c = 10\,\%$, the risk-free interest rate $r = 2\,\%$, the volatility $\sigma = 15\,\%$, and the initial investment $X_0 = \$1$. For simplicity and without loss of generality we assume a *zero management fee* for our base case.

With our base case parameters, the total value of the investor's payoff is 1.0073, and the value of the manager's payoff is -0.0073. Notice that the value of the investor's payoff is greater than the initial investment of 1. In contrast, the price of the traditional investor payoff (without the insurance part of the payoff—i.e. removing both put options) is 0.9909, and the value of the manager's payoff in this instance is 0.0091.

5.1.1 Payoff Functions of the Investor and the Manager

The payoff functions of the investor under the shared-loss and the traditional fee structures are given in Fig. 1. The payoff to the hedge fund manager using the aforementioned benchmark values and under the shared-loss fee structure is also depicted in Fig. 2 along with the traditional payoff structure with only the performance fee $\alpha(X_T - X_0)_+$. Observe that since the options trades constitute a zero-sum game (the positions of the manager and the investor are opposite each other), the sum of the investor payoff and the manager payoff is equal to X_T.

Figure 3 illustrates the 'fair performance fee', where investor gets a payoff with present value equal to his initial cash injection, X_0, given volatility and manager's deposit levels, i.e. we set $V_I(0) = X_0$. The fair performance fee can be easily obtained from Eq. (4) as,

$$\alpha_{\text{fair}} = \frac{-e^{-rT} m X_0 + P(X_0, T, X_0 + m X_0, r, \sigma) - P(X_0, T, (1-c)X_0 + m X_0, r, \sigma)}{C(X_0, T, X_0 + m X_0, r, \sigma)}$$

Interested reader can derive explicit, well-known expressions for the sensitivities of the α_{fair} relative to different parameters in terms of the Greeks and Vega of the involving options. As can be seen from the figure, for small values of volatility, the fair performance fee is indifferent to the levels of manager's deposit; however, as volatility increases, a higher level of deposit by the manager translates into a higher

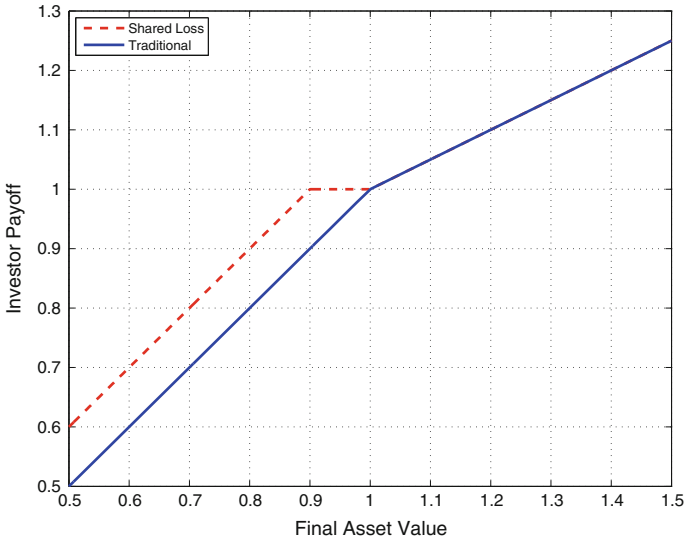

Fig. 1 Payoff for the hedge fund investor

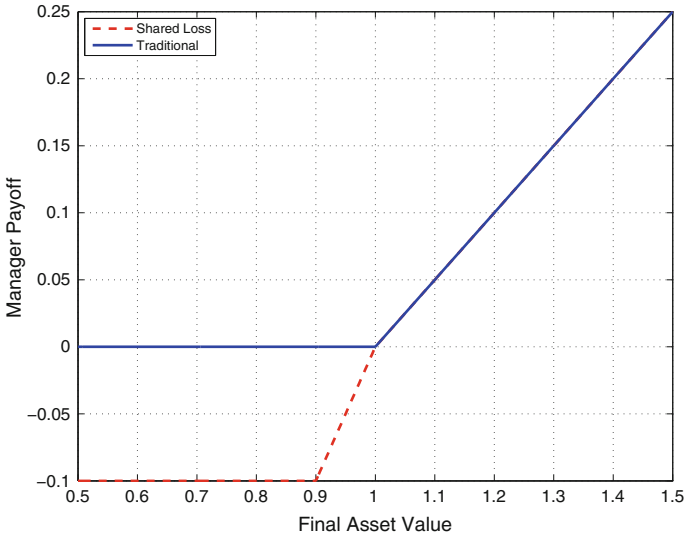

Fig. 2 Payoff for the hedge fund manager

performance fee paid by the investor to make the deal a fair one. In Fig. 4, we normalize the volatility on the horizontal axis by the manager's deposit defined *as a percentage of the initial investment* X_0. For a given level of deposit, the higher the volatility of the underlying investment, the higher the probability that the loss incurred by the manager exceeds the deposit. In other words, the probability that the manager

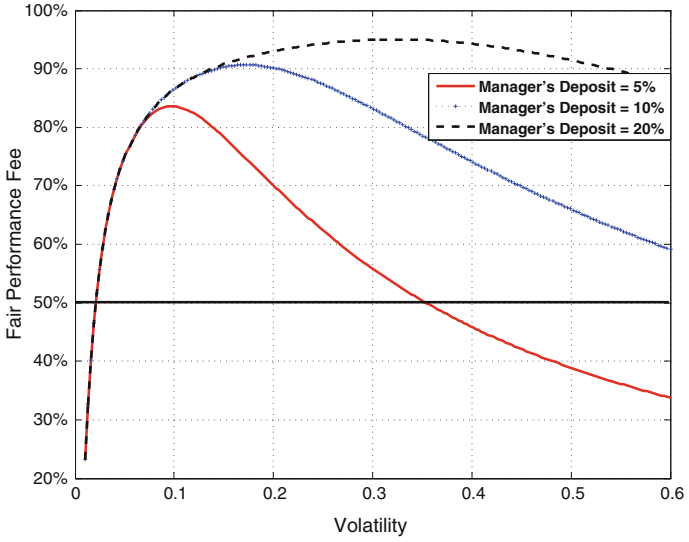

Fig. 3 Fair performance fee versus volatility

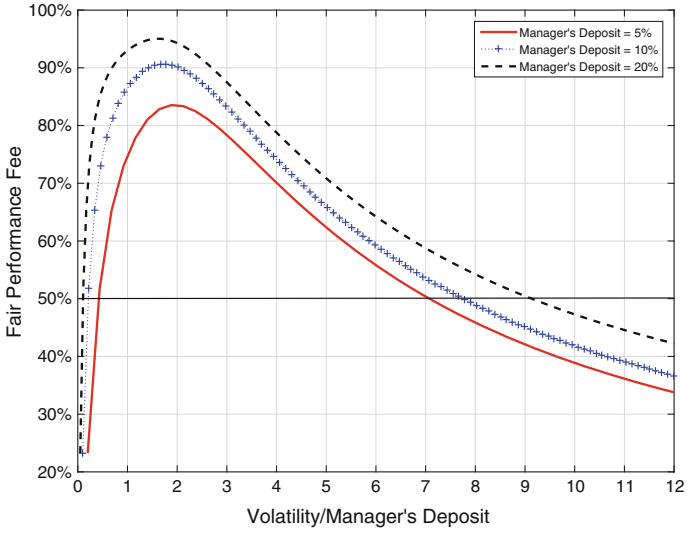

Fig. 4 Fair performance fee versus normalized (by deposit) volatility

exercises the put option offered by the investor increases, which results in a reversal in the fair performance fee for higher levels of volatility. This is clearly illustrated in Fig. 4 where volatility and deposit are combined in a single scaling variable, that is, volatility/deposit, *where the deposit is expressed as a percentage of the initial investment* X_0. The corresponding maximum value for the fair performance fee

increases with the size of the deposit; that's because for higher deposits, the manager will have to lose more and more before the investor starts bearing the residual loss, therefore his compensation should be higher accordingly. Note that the x-axis in Figs. 3 and 4 is incorporating the annual volatility of the fund assets; however, the performance fee is crystallized on a monthly basis which suggests a comparison between the deposit level and *monthly* volatility, as opposed to annual volatility. Since returns are assumed to follow a normal distribution in our Black–Scholes framework, one can explicitly calculate the probability of the returns falling into a certain interval, in particular, with about 68 % probability, the return falls within 1 standard deviation of the mean. This explains why the curves for various deposits reach a maximum roughly around the same level of (annual) volatility/deposit ratio, in the [1, 2] interval.

5.2 *Sensitivity Analysis*

In this section, we perform a sensitivity analysis of the prices of the investor's and manager's payouts, as a function of the different model parameters.

5.2.1 Volatility (σ)

Figure 5 shows the value of the investor's position as a function of the volatility parameter σ, as σ ranges from 5 % to 60 %.

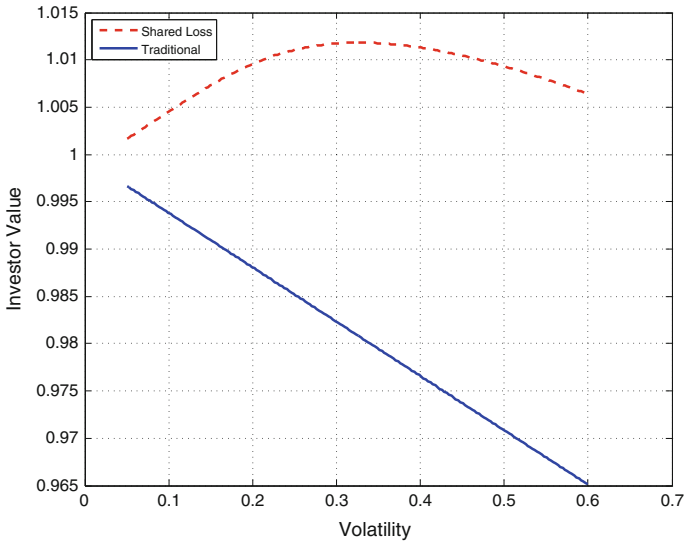

Fig. 5 Value of the investor's position versus volatility σ

We see that the position is initially an increasing function of the volatility, owing to the increasing value of the investor's put option as a function of σ. However, as the volatility becomes very large, the value of the investor's position starts to decline as the hedge fund's call option, as well as its put option, become more valuable. The maximum value for the investor occurs at a volatility around $\sigma = 32.5\,\%$. Observe however, that the value is relatively insensitive to the level of σ, with a minimum value of 1.0016, and a maximum value of 1.0118.

5.2.2 Manager Deposit (c)

We varied the manager deposit between 1 % and 25 %, while holding all other parameters at their base case values. The results of the sensitivity analysis are shown in Fig. 6.

As would be expected, the value of the investor's position is an increasing function of the manager's deposit. The value of the position is equal to one (break-even point, or 'fair fee point') at around $c = 0.0233$. Any deposit level less than $c = 0.0233$ puts the investor at a disadvantage, and the investor is indifferent to deposit levels higher than 10 %.

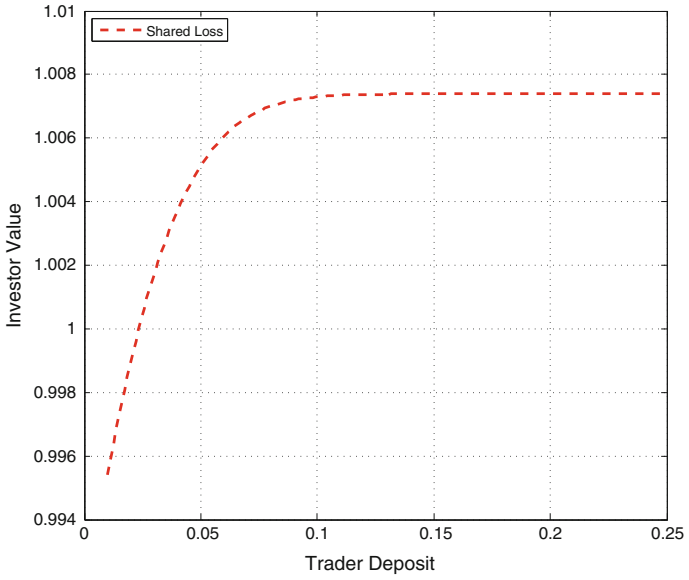

Fig. 6 Value of the investor's position versus manager's deposit c

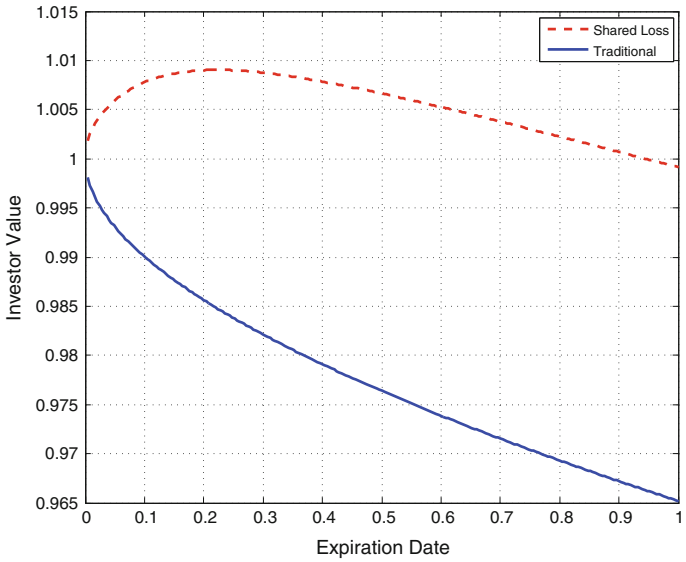

Fig. 7 Value of the investor's position versus the expiration date T

5.2.3 Maturity Date (T)

The dependence on the time to maturity is of interest specially when adapting the results of this paper to realistic situations. As we mentioned earlier, our mathematical assumption is that fees will be paid at a fixed time in the future. In practice, fees are payable according to calendars agreed between the investors and the manager. In the graphs that follow, we address this by varying the expiration date T from 1 day to 1 year. The results are shown in Fig. 7.

Initially, the value of the position is increasing in T, but eventually, it begins to decrease in T, as the options given to the hedge fund manager become more valuable. The maximum value of the investor's position occurs at T around one quarter of a year ($T \sim 0.22$).

6 Conclusion

The exchange of business value between the manager and the investor is always a complex one: beyond fees paid, there are intangibles the investor gives the manager. An asset management business is valued taking into account many factors, such as track records, years in business, assets under management, the reputation of its investors, and of course fees. In this paper we focus on first-loss fee structures, which are bringing novel points of attention between investors and hedge fund managers

in the historical discussions on fair compensation. We focus only on the fee payable by the investor and the guarantee offered by the manager, which is the main novelty in this set up. The main challenge in this new paradigm is to evaluate the value of the guarantee offered by the hedge fund manager in relation to the fee paid by the investor. In this paper, we developed a mathematical approach to compare the two features of guarantee and performance fee from an option pricing perspective. The framework is flexible and can be used for different specific investment settings and can account for slight variations from one fund to another. Our salient leitmotif is: fee agreements must be structured to be attractive to managers so they are willing to participate, and at the same time provide a cushion against losses to the investor. A significant contribution, that sheds light on the road-map and paves the way for deeper investigations, is to see, and more importantly formulate, the underlying fee structure from the lens of option valuation. By employing a risk-neutral framework and options pricing theory, one is able to not only price, but also analyze the sensitivity of the value of the investor's and manager's positions in reference to a set of influential parameters.

Acknowledgements We wish to express our gratitude to Sigma Analysis & Management Ltd., and especially to Dr. Ranjan Bhaduri and Mr. Kurt Henry for many endless valuable discussions.

The KPMG Center of Excellence in Risk Management is acknowledged for organizing the conference "Challenges in Derivatives Markets - Fixed Income Modeling, Valuation Adjustments, Risk Management, and Regulation".

References

1. Banzaca, J.: First loss capital arrangements for hedge fund managers: Structures, risks and the market for key terms. The Hedge Fund Law Report **5**(37), (2012)
2. Carpenter, J.: Does option compensation increase managerial risk appetite? J. Financ. **55**(5), 2311–2331 (2000)
3. Goetzmann, W., Ingersoll, J., Ross, S.: High-water marks and hedge fund management contracts. J. Financ. **58**(4), 1685–1717 (2003)
4. Guasoni, P., Obłój, J.: The incentives of hedge fund fees and high-water marks, forthcoming. Mathematical Finance (2013)
5. He, X., Kou, S.: Profit sharing in hedge funds. www.ssrn.com (2013)
6. Hodder, J., Jackwerth, J.: Incentive contracts and hedge fund management. J. Financ. Quant. Anal. **42**(4), 811–826 (2007)

7. Kouwenberg, R., Ziemba, W.: Incentives and risk taking in hedge funds. J. Bank Financ. **31**, 3291–3310 (2007)
8. Panageas, S., Westerfield, M.: High-water marks: high risk appetites? Convex compensation, long horizons, and portfolio choice. J. Financ. **64**(1), 1–36 (2009)
9. Ross, S.: Compensation, incentives, and the duality of risk aversion and riskiness. J. Financ. **59**(1), 207–225 (2004)
10. WSJ: Calpers to Exit Hedge Funds. http://www.wsj.com/articles/calpers-to-exit-hedge-funds-1410821083?alg=y (2014)

Negative Basis Measurement: Finding the Holy Scale

German Bernhart and Jan-Frederik Mai

Abstract Investing into a bond and at the same time buying CDS protection on the same bond is known as buying a basis package. Loosely speaking, if the bond pays more than the CDS protection costs, the position has an allegedly risk-free positive payoff known as "negative basis". However, several different mathematical definitions of the negative basis are present in the literature. The present article introduces an innovative measurement, which is demonstrated to fit better into arbitrage pricing theory than existing approaches. This topic is not only interesting for negative basis investors. It also affects derivative pricing in general, since the negative basis might act as a liquidity spread that contributes as a net funding cost to the value of a transaction; see Morini and Parampolini (Risk, 58–63, 2011, [23]).

Keywords Negative basis measurement · Bond-CDS basis · Hidden yield

1 Introduction

On first glimpse, it is surprising that investing into a bond and buying CDS protection on that underlying bond, henceforth called a basis package, can earn an attractive spread on top of the risk-free rate of return, as it appears to be free of default risk. This excess return over the risk-free rate is informally called *negative basis*[1]; more formal definitions are given in the main body of this article. [8] has even devoted an entire book to the topic. If, conversely, the cost of CDS protection exceeds the bond earnings, one speaks of a positive basis. In this article, we only speak of negative bases, as fundamentally the concepts of positive and negative basis are simply inverse.

[1] Sometimes also called *bond-CDS basis*.;

G. Bernhart · J.-F. Mai (✉)
XAIA Investment, Sonnenstraße 19, 80331 München, Germany
e-mail: jan-frederik.mai@xaia.com

G. Bernhart
e-mail: german.bernhart@xaia.com

K. Glau et al. (eds.), *Innovations in Derivatives Markets*, Springer Proceedings in Mathematics & Statistics 165, DOI 10.1007/978-3-319-33446-2_18

The appropriate measurement of negative basis plays an important role with regard to the cost of funding literature, which has become of paramount interest in the financial industry since the recent liquidity crisis. Generally speaking, this stream of literature reconsiders the pricing of derivatives under the new post-crisis fundamentals regarding funding, liquidity, and credit risk issues. Substantial contributions have been made, among others, by [5, 7, 12, 13, 23, 27, 29]. Loosely speaking, most references agree upon the fact that, at least under certain simplifying assumptions (full, bilateral, and continuous collateralization), derivative contracts can be evaluated in the traditional way, only the involved discount factors have to be adjusted by means of a spread accounting for funding and liquidity charges. In particular, [23] show in a simple, theoretical framework that the negative basis is a spread which plays an essential role in this regard. In order to set these theoretical findings into action in the industry's pricing machinery, it is therefore an essential task to establish viable and reasonable measurements for the negative basis. The present article shows that this topic is not only important but also challenging, and contributes a careful comparison of three different measurement methods. In particular, we point out why the most common measurement approaches (denoted by (Z) and (PE) below) are not recommended, and propose a decent alternative.

In the present article, we take the point of view of a negative basis investor whose goal is to detect interesting negative basis positions and to monitor the evolution of such investments over time. Alternatively, consider a bank which has to evaluate its derivative book. As the aforementioned references show that the required discount factors for the pricing algorithms might have to be adjusted by means of the negative basis, one faces the task of measuring this negative basis appropriately. For the effective implementation of these tasks, it is crucial to come up with a reliable and viable, yet reasonable mathematical definition of what the negative basis actually is. Specific focus is put on simple-to-implement approaches that rely on commonly applied pricing methodologies for bonds and CDS, described in, e.g., [18, 25]. In total, we discuss three different measurements (two traditional and one innovative):

- Difference between Z-spread of the bond and CDS running spread, as presented, e.g., in [8], and defined by Bloomberg on the screen YAS.
- Par-equivalent CDS-methodology, as described in the Appendix of [2], who apply this definition for an empirical study, see also [3].
- A hidden yield approach that assumes the risk-free discounting curve to be a reference interest rate curve shifted by the (initially unknown) negative basis.

Important to note is that, according to all these definitions, a negative basis is assigned to a bond, not to an issuer. This means that two different bonds issued by the same company are allowed to have two different negative bases. This viewpoint stands in glaring contrast to some of the more macro-economic considerations carried out in references cited in the next section. CDS protection typically refers to a whole battery of eligible bonds by a reference issuer, and normally the major driver for CDS spreads is considered to be the issuer's default risk. However, some of the deliverable bonds might trade at diverse yields for reasons other than the issuer's

default risk—for instance legal issues, liquidity issues, or funding issues, cf. [21] and Sect. 2.

The rest of this article is organized as follows. Section 2 recalls reasons for the existence of negative basis. Section 3 introduces general notations, which are used throughout the remaining sections. Section 4 reviews the traditional methods (Z) and (PE), Sect. 5 discusses the innovative method (HY), and Sect. 6 concludes.

2 Why Does Negative Basis Exist?

There are a couple of intuitive explanations for the existence of negative basis, see, e.g., [1, 2, 4, 6, 10, 19, 24, 26, 30]. For the convenience of the reader, we briefly recall some of them in the sequel.

- **Liquidity issues**: Some bond issues are distributed only among a few investors. If one of these investors has to sell her bonds, for instance due to regulatory requirements or demand for liquidity, supply may exceed demand and thus the price of the bond must drop significantly in order for the bond to be sold. At the same time the CDS price might remain unaffected.
- **Funding costs**: From a pure credit risk perspective, selling CDS protection economically is the same risk as buying the underlying bond. However, buying a bond requires an initial investment that must be funded, whereas selling CDS protection typically requires much less initial funding (unless the CDS upfront exceeds the bond price). Therefore, in times of high funding costs there is an incentive to sell CDS rather than to buy bonds, which might lead to an increase in supply of CDS protection, making it cheap relative to bond prices.
- **Market segmentation**: Empirical observations suggest that bond trades sometimes have larger volumes and might be motivated much less by quantitative aspects than CDS trades. Arguing similarly, [6, p. 5, l. 5–7] conjecture that *"market-implied [risk] measures have a stronger impact on the CDS market, while the more easily available rating information affects the bond market more strongly"*. Such instrument-specific differences might contribute to the existence of negative basis.
- **Legal risk**: The bond of the negative basis position might bear certain risks that cannot be protected against by means of a CDS. Examples are certain collective action clauses, debt restructuring events, or call rights for the bond issuer. Such "legal gaps" explain parts of the negative basis.
- **Counterparty credit risk**: A joint default of both the CDS counterparty and the issuer of the bond could lead to a loss for the basis position.[2] These potential losses imply that CDS protection is not 100 % and consequently might contribute to the negative basis, see, e.g., [5, 22].
- **Mark-to-market risk**: The negative basis might further increase after one has entered into the position, due to one of the aforementioned reasons. In this case, one

[2]However, counterparty credit risk can be reduced significantly by a negative basis investor when the CDS is collateralized, which is the usual case.

loses money due to mark-to-market balancing. In theory, one gets this money back eventually, but it might occur that mark-to-market losses exceed one's personal tolerance level during the bond's lifetime. In this case, one has to exit the position and realize the loss. This risk is especially significant if the negative basis position is levered (which has happened heavily during the financial crisis). Part of the negative basis might be viewed as a risk premium for taking this mark-to-market-risk.

Basis "arbitrageurs" are investors that try to earn the negative basis by investing into basis packages. This means that they consider the negative basis an adequate compensation for taking the aforementioned risks. In classical arbitrage theory, their appearance improves trading liquidity. Counterintuitively, however, [9] argue that the advent of CDS was detrimental to bond markets and [20] find some evidence that basis arbitrageurs bring new risks into the corporate bond markets.

3 General Notations

All definitions to follow rely on the pricing of CDS and a plain vanilla coupon bond according to the most simple mathematical setup we can think of. This is in order to make the article as reader-friendly as possible; furthermore, we think the setup is already rich enough in order to convey the main ideas. The only randomness considered in the present article is the default time of the bond issuer, which is formally defined on a probability space $(\Omega, \mathscr{F}, \mathbb{Q})$, with state space Ω, σ-algebra $\mathscr{F}$, and probability measure $\mathbb{Q}$. Expected values with respect to the pricing measure $\mathbb{Q}$ are denoted by $\mathbb{E}$. The default intensity $\lambda(.)$ of the issuer's default time τ is assumed to be deterministic, i.e. $\mathbb{Q}(\tau > t) = \exp(-\int_0^t \lambda(s)\,\mathrm{d}s)$. Sometimes the function $\lambda(.)$ is constant, sometimes piecewise constant, depending on our application. For example, the computation of a so-called Z-spread requires $\lambda(.)$ to be constant,[3] whereas the joint consistent pricing of several CDS quotes with different maturities requires $\lambda(.)$ to be piecewise constant.

Generally speaking, it is our understanding that a negative basis is a measure for the mispricing between CDS and bonds with respect to default risk alone. This explains why considering the default time as the sole stochastic object corresponds to the most minimal modeling approach possible. Besides the non-randomness of the default intensity, the following further simplifying assumptions are taken for granted throughout:

- We ignore recovery risk: Upon default, the bond holder receives the constant proportion $R \in [0, 1]$ of her nominal. Default is assumed to instantaneously trigger a credit event of the CDS. The bond is assumed to be a deliverable security in the auction following the CDS trigger event, and the auction process is assumed to yield the same recovery rate R. Although this is an unrealistic assumption in

[3] See below in Step 3 of Definition 1.

principle (see, e.g., [17]), a negative basis investor can always eliminate recovery risk by delivering his bonds into the auction (physical settlement), in which case he gets compensated by the (nominal-matched) CDS for the nominal loss of the bond.[4] Consequently, our assumption is not severe for the present purpose.

- We ignore interest rate risk: The discounting curve is deterministic and the discount factors are denoted by $DF(t) := \exp(-\int_0^t r(s)\,\mathrm{d}s)$ with some given deterministic short rate function $r(.)$. All presented negative basis figures are measurements relative to the applied short rate function $r(.)$.

Under these assumptions we introduce the following notations:

- $t_j^{(B)}$ denotes the coupon payment dates of the bond.
- The bond's lifetime is denoted by T, i.e. T denotes the last coupon payment date, which at the same time is the redemption date. Moreover, the bond is assumed to pay a constant coupon rate C at each coupon payment date.
- $t_i^{(C)}$ denotes the payment dates of the considered CDS contracts, which typically are quarterly on the 20th of March, June, September, and December, respectively, according to the terms and conditions of ISDA standard contracts.[5]
- For a CDS with maturity T, the (usually standardized) running coupon is denoted by $s(T)$ and the upfront payment to be made at CDS settlement by $\mathrm{upf}(T)$.
- The expected discounted value of the sum of all premium payments to be made by the CDS protection buyer (the premium leg) is denoted by[6]

$$\begin{aligned} &EDPL(\lambda(.), r(.), s(T), \mathrm{upf}(T), T) \\ &\quad := \mathrm{upf}(T) + s(T) \sum_{0<t_i^{(C)}\le T} \left(t_i^{(C)} - t_{i-1}^{(C)}\right) DF\left(t_i^{(C)}\right) \mathbb{Q}\left(\tau > t_i^{(C)}\right) \\ &\quad = \mathrm{upf}(T) + s(T) \sum_{0<t_i^{(C)}\le T} \left(t_i^{(C)} - t_{i-1}^{(C)}\right) DF\left(t_i^{(C)}\right) e^{-\int_0^{t_i^{(C)}} \lambda(s)\,\mathrm{d}s}. \end{aligned}$$

- The expected discounted value of the sum of all default compensation payments to be made by the CDS protection seller (the default/protection leg) is denoted by

$$\begin{aligned} EDDL(\lambda(.), r(.), R, T) &:= (1-R)\,\mathbb{E}[1_{\{\tau\le T\}}\, DF(\tau)] \\ &= (1-R) \int_0^T DF(y)\,\lambda(y)\, e^{-\int_0^y \lambda(s)\,\mathrm{d}s}\,\mathrm{d}y. \end{aligned}$$

[4] Interestingly, a mismatch between bond and CDS recovery is often favorable for the negative basis investor, since the CDS recovery rate tends to be lower than the bond recovery, see, e.g., [14]. Thus, it might make sense for a negative basis investor to opt for cash settlement of the CDS and sell his bonds in the marketplace, speculating on a favorable recovery mismatch.

[5] See http://www2.isda.org/asset-classes/credit-derivatives/.

[6] For the sake of notational convenience we ignore accrued interest upon default, which can, of course, be incorporated easily.

- The model price of the bond is given by

$$\begin{aligned}\text{Bond}(\lambda(.), r(.), R, C, T) &:= C \sum_{0<t_j^{(B)}\le T} \left(t_j^{(B)} - t_{j-1}^{(B)}\right) DF\left(t_j^{(B)}\right) \mathbb{Q}\left(\tau > t_j^{(B)}\right) \\ &\quad + DF(T)\, \mathbb{Q}(\tau > T) + R\, \mathbb{E}[1_{\{\tau \le T\}}\, DF(\tau)] \\ &= C \sum_{0<t_j^{(B)}\le T} \left(t_j^{(B)} - t_{j-1}^{(B)}\right) DF\left(t_j^{(B)}\right) e^{-\int_0^{t_j^{(B)}} \lambda(s)\,\mathrm{d}s} \\ &\quad + DF(T)\, e^{-\int_0^T \lambda(s)\,ds} + R \int_0^T DF(y)\, \lambda(y)\, e^{-\int_0^y \lambda(s)\,\mathrm{d}s}\,\mathrm{d}y.\end{aligned}$$

4 Traditional Measurements

4.1 The Z-Spread Methodology

The main idea of the *Z-spread methodology* is to define the negative basis as the difference between (expected) annualized bond earnings and annualized protection costs. This method is described, e.g., in [8]. The negative basis $NB^{(Z)}$ is computed by the following algorithm.

Definition 1 (*Negative Basis (Z)*)

1. A reference discounting curve, resp. the associated short rate $r(.)$, is chosen and used in all subsequent steps, e.g. bootstrapped from quoted prices for interest rate derivatives according to one of the methods described in [15, 16].
2. From a term structure of quoted CDS with different maturities, piecewise constant intensities $\lambda(.)$ are bootstrapped, as described, e.g., in [25]. For this, a recovery assumption is made, i.e. R is model input.[7]
3. Denoting by B the quoted market price of the bond, the bond's **Z-spread** z is defined as the root of the function[8]

$$x \mapsto \text{Bond}(x, r(.), 0, C, T) - B, \tag{1}$$

[7] If CDS prices are quoted in running spreads with zero upfronts, then these quotes typically come naturally equipped with a recovery assumption that is required in order to convert the running spreads into actually tradable standardized coupon and upfront payments. However, after this conversion the recovery rate is a free model parameter.

[8] For a reader-friendly explanation of the Z-spread see [28]. In particular, it is useful to observe that $\text{Bond}(x, r(.), R, C, T) = \text{Bond}(0, r(.) + x, R, C, T)$ for $R = 0$, implying that the Z-spread equals a constant default intensity under a zero recovery assumption.

if existent. In words, the Z-spread is the amount by which the reference short rate $r(.)$ needs to be shifted parallelly in order for the discounted bond cash flows to match the market quote. The root, whenever existing at all, is unique.

4. The (zero-upfront) running CDS spread $s(T)$ for a CDS contract, whose maturity matches the bond's maturity, is defined as

$$s(T) := \frac{EDDL(\lambda(.), r(.), R, T)}{EDPL(\lambda(.), r(.), 1, 0, T)},$$

i.e. the fair running spread when no upfront payment is present.

5. $NB^{(Z)} := z - s(T)$.

Intuitively, the Z-spread z is a measure of the annualized excess return of the bond on top of the "risk-free" rate $r(.)$, whereas $s(T)$ is the annualized CDS protection cost. Hence, $NB^{(Z)}$ equals the difference between earnings and costs (expected in case of survival). If the function (1) does not have a root in $(0, \infty)$, this means that the bond is less risky than the default risk intrinsic in the chosen discounting curve $r(.)$. Especially since the liquidity crisis, when the interbank money transfer ran dry, significant spreads between discounting curves obtained from overnight rates and LIBOR-based swap rates are observed. Consequently, one could recognize, e.g., German government bonds with a "negative Z-spread" with respect to the interest rate curve $r(.)$, which was obtained from 6-month EURIBOR swap rates. For such reasons it has become market standard to extract the "risk-free" discounting curve from overnight rates rather than from LIBOR-based swap rates. Moreover, [19] point out that the difference between bond yields and CDS spreads can depend on whether treasury rates or swap rates are used for discounting. Since negative basis investors are typically trading in the high yield sector, the function (1) normally does have a root in $(0, \infty)$ for several canonical choices of $r(.)$, be it extracted from swap rates with overnight tenor, 3-month tenor, or 6-month tenor. But it is important to stress that all presented negative basis measurements are always relative measures depending on the applied interest rate curve $r(.)$.

The Z-spread methodology has some drawbacks:

- **Imprecision:** Earnings and costs are not measured accurately, but only approximately. The Z-spread is only a rough estimate for the expected annualized earnings, and the zero-upfront running CDS spread is also not really tradable, but only a fictitious quantity. Furthermore, the Z-spread is earned on the bond value, whereas the CDS spread is paid on the (bond and) CDS nominal, which may result in a nonsense measurement for bonds trading away from par, see Example 1 below. To this end, [10] proposes to replace the Z-spread by an asset swap spread. It is possible to define more accurate measurements of earnings and costs taking into account actual cash flows. However, in the present article we do not elaborate on these fine-tunings, since the "earnings and costs"-perspective in general suffers from the following second difficulty.
- **Inaccurate hedge:** The measurement assumes that bond and CDS have the same maturity and nominals and furthermore implicitly assumes a survival

until maturity. Upon a default event the PnL of the position might be considerably different, depending on the timing of the default, see Fig. 1 in Example 1 below. Hence, the assumed CDS hedge cannot really be considered to be default-risk eliminating (it might either profit from or lose on a default event), and consequently the number $NB^{(Z)}$ does not deserve to be called a return figure after elimination of default risk, which the negative basis should be in our opinion.

4.2 The Par-Equivalent CDS Methodology

The *par-equivalent CDS methodology* is described in the Appendix of [2]. A similar idea is also outlined in [8, p. 101 ff] and [3]. The negative basis $NB^{(PE)}$ is computed along the steps of the following algorithm.

Definition 2 (*Negative Basis (PE)*)

1. A reference discounting curve, resp. the associated short rate $r(.)$, is chosen and used in all subsequent steps, e.g. bootstrapped from quoted prices for interest rate derivatives according to one of the methods described in [15, 16].
2. From a term structure of CDS contracts on the reference entity, piecewise constant intensities $\lambda(.)$ are bootstrapped, as described, e.g., in [25]. For this a recovery assumption is made, i.e. R is model input.
3. The (zero-upfront) running CDS spread $s(T)$ for a CDS contract, whose maturity matches the bond's maturity, is defined as

$$s(T) := \frac{EDDL(\lambda(.), r(.), R, T)}{EDPL(\lambda(.), r(.), 1, 0, T)},$$

   i.e. the fair running spread when no upfront payment is present.
4. Denoting by B the quoted market price of the bond, a shift $\tilde{z}$ is defined as the root of the function

$$x \mapsto \text{Bond}(\lambda(.) + x, r(.), R, C, T) - B,$$

   if existent. In words, the bond is priced with the default intensities $\lambda(.)$ that are consistent with CDS quotes, which are then shifted parallelly until the bond's market quote is matched.
5. A second (zero-upfront) running CDS spread $\tilde{s}(T)$ for a CDS contract, whose maturity matches the bond's maturity, is defined as

$$\tilde{s}(T) := \frac{EDDL(\lambda(.) + \tilde{z}, r(.), R, T)}{EDPL(\lambda(.) + \tilde{z}, r(.), 1, 0, T)},$$

   i.e. the fair spread when no upfront payment is present, but now with the shifted intensity rates $\lambda(.) + \tilde{z}$, which are required in order to price the bond correctly.
6. $NB^{(PE)} := \tilde{s}(T) - s(T)$.

The main idea of (PE) is to question the default probabilities bootstrapped from the given CDS quotes, and to adjust them in order to match the bond quote. On a high level, this negative basis measurement is based on the difference between default probabilities that are required in order to match the bond price and default probabilities that are required in order to fit the CDS quotes.

The methodology (PE) has some drawbacks:

- **No link to arbitrage pricing theory:** In our view, there is no convincing economic argument as to why two different survival functions for the same default time should be used. In particular, the method provides no joint pricing model for bond and CDS that explains the negative basis as one of its parameters. The method is "decoupled" from arbitrage pricing theory.
- **No link to "earnings and costs"-perspective:** Unlike the method (Z), the method (PE) does not have a clear link to an earnings measure above a reference rate, which is what the negative basis is informally thought of.

5 An Innovative Methodology

In our opinion, the negative basis should be a spread on top of a reference discounting curve which can be earned without exposure to default risk. This means we question the usual assumption that the applied discounting curve $r(.)$ is the appropriate risk-free rate to be used, because there is actually a higher rate that can be earned "risk-free" (recalling that default risk is the only risk within our tiny model). This motivates what we call the *hidden yield approach*. The negative basis $NB^{(HY)}$ is computed along the steps of the following algorithm.

Definition 3 (*Negative Basis (HY)*)

1. A reference discounting curve, resp. the associated short rate $r(.)$, is chosen and used in all subsequent steps, e.g. bootstrapped from quoted prices for interest rate derivatives according to one of the methods described in [15, 16].
2. Denote by $\lambda_x(.)$ the piecewise constant intensity rates that are bootstrapped from CDS market quotes, when the assumed discounting curve is $r(.) + x$, as described, e.g., in [25]. The recovery rate R is fixed and chosen as model input.
3. The negative basis $NB^{(HY)}$ is defined as the root[9] of the function

$$x \mapsto \text{Bond}(\lambda_x(.), r(.) + x, R, C, T) - B.$$

In words, $NB^{(HY)}$ is precisely the parallel shift of the reference short rate $r(.)$ which allows for a calibration such that the model prices of bond and CDS match the observed market quotes for bond and CDS.

[9] Lemma A.1 in the Appendix guarantees that this root typically exists and is unique.

The idea of method (HY) can also be summarized as follows: If the risk-free interest rate curve is assumed to be $r(.) + NB^{(HY)}$, then the market quotes for bond and CDS are arbitrage-free (as we have found a corresponding pricing measure). It allows for the intuitive interpretation of the negative basis as a spread earned on top of a reference discounting rate after elimination of default risk. Abstractly speaking, assuming no transaction costs and availability of CDS protection at all maturities $T > 0$ (= perfect market conditions), arbitrage pricing theory suggests the existence of a trading strategy which buys the bond and hedges it via CDS, and which earns[10] precisely the rate $r(.) + NB^{(HY)}$ until the minimum of default time τ and bond maturity T. Since this way of thinking about $NB^{(HY)}$ is its distinctive property and highlights its intrinsic coherence with arbitrage pricing theory, the following lemma demonstrates by a heuristic argument how the rate $r(.) + NB^{(HY)}$ can be earned in a risk-free way.

Lemma 1 (The rate $r(.) + NB^{(HY)}$ can be earned without default risk) *Assuming perfect market conditions, there exists a (static) portfolio, which is long the bond and invested in several CDS, which earns the rate* $r(.) + NB^{(HY)}$ *until* $\min\{\tau, T\}$.

Proof (heuristic) We denote by $\mathbb{Q}$ the probability measure under which τ has piecewise constant default intensity $\lambda_{NB^{(HY)}}(.)$. We discretize the time interval $[0, T]$ into m buckets $0 =: t_0 < t_1 < \ldots < t_m := T$, but m may be chosen arbitrarily large such that the mesh of the discrete-time grid tends to zero as m tends to infinity. We introduce the following $m + 1$ probabilities:

$$w_j^{(m)} := \mathbb{Q}\big(\tau \in (t_{j-1}, t_j]\big), \quad j = 1, \ldots, m, \quad w_{m+1}^{(m)} := \mathbb{Q}(\tau > t_m).$$

Now let $\tau^{(m)}$ denote a random variable with distribution

$$\mathbb{Q}\Big(\tau^{(m)} = \bar{t}_j\Big) = w_j^{(m)}, \quad \bar{t}_j := \frac{t_{j-1} + t_j}{2}, \quad j = 1, \ldots, m,$$
$$\mathbb{Q}\Big(\tau^{(m)} > t\Big) = \mathbb{Q}(\tau > t), \quad t \geq t_m, \Big(\text{in particular, } \mathbb{Q}(\tau^{(m)} > t_m) = w_{m+1}^{(m)}\Big).$$

Notice that $\tau^{(m)} \approx \tau$ in distribution, with the approximation improving with increasing m. In the sequel, we work with $\tau^{(m)}$, assuming that default during $[0, T]$ can only take place at the possible realizations $\bar{t}_1, \ldots, \bar{t}_m$ of $\tau^{(m)}$ in $[0, T]$. We now consider a portfolio of $m + 1$ instruments, namely the bond and one CDS for each maturity $t_1, \ldots, t_m$. We assume that the bond nominal is given by N_0. Furthermore, $N_i \in \mathbb{R}$ denotes the nominal of the CDS with maturity t_i. Negative nominal means that we sell the bond or sell CDS protection. Let's have a look at the following random variables, which are functions of $\tau^{(m)}$:

[10] By "earning" $r(.) + NB^{(HY)}$ we mean that the internal rate of return of the position is the reference rate $r(.)$ plus a spread $NB^{(HY)}$.

$$V^{(0)}\big(\tau^{(m)}\big) := \Big(r(.) + NB^{(HY)}\Big)\text{-discounted value of all cash flows from the bond, when default takes place at } \tau^{(m)},$$

$$V^{(i)}\big(\tau^{(m)}\big) := \Big(r(.) + NB^{(HY)}\Big)\text{-discounted value of all cash flows from the CDS with maturity } t_i\text{, when default takes place at } \tau^{(m)}, \quad i = 1, \ldots, m.$$

All random variables $V^{(i)}\big(\tau^{(m)}\big)$ take on only $m+1$ possible values, since their value on the event $\{\tau^{(m)} > t_m\}$ does not depend on $\tau^{(m)}$ (as there are no cash flows after t_m). So without loss of generality we may write $V^{(i)}\big(\tau^{(m)}\big) = V^{(i)}\big(\bar{t}_{m+1}\big)$ for some arbitrary $\bar{t}_{m+1} > t_m$ on the event $\{\tau^{(m)} > t_m\}$. Our goal is to show that it is possible to find a non-zero vector $(N_0, \ldots, N_m) \in \mathbb{R}^{m+1}$ such that

$$\underbrace{N_0\, V^{(0)}\big(\tau^{(m)}\big) + \sum_{i=1}^{m} N_i\, V^{(i)}\big(\tau^{(m)}\big)}_{\left(r(.)+NB^{(HY)}\right)\text{-discounted value of outcome}} \equiv \underbrace{N_0\, B + \sum_{i=1}^{m} N_i\, \mathrm{upf}(t_i)}_{\text{initial investment amount}}, \tag{2}$$

where B denotes the market bond price and $\mathrm{upf}(t_i)$ the market upfront of the CDS with maturity t_i. This mathematical statement intuitively means that the considered portfolio of bond and CDS earns the rate $r(.) + NB^{(HY)}$ until $\min\{\tau^{(m)}, T\}$ in a risk-free manner, regardless of the actual timing of the default. Now why is this possible? Considering the randomness on the left-hand side of Eq. (2), we actually have $m+1$ equations for the $m+1$ unknowns $N_0, N_1, \ldots, N_m$. Rewriting Eq. (2) in terms of linear algebra, we obtain

$$\begin{pmatrix} V^{(0)}\big(\bar{t}_1\big) - B & V^{(1)}\big(\bar{t}_1\big) - \mathrm{upf}(t_1) & \ldots & V^{(m)}\big(\bar{t}_1\big) - \mathrm{upf}(t_m) \\ \vdots & & \ddots & \vdots \\ \vdots & & \ddots & \vdots \\ V^{(0)}\big(\bar{t}_{m+1}\big) - B & V^{(1)}\big(\bar{t}_{m+1}\big) - \mathrm{upf}(t_1) & \ldots & V^{(m)}\big(\bar{t}_{m+1}\big) - \mathrm{upf}(t_m) \end{pmatrix} \begin{pmatrix} N_0 \\ N_1 \\ \vdots \\ N_m \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{3}$$

In order to prove the existence of a non-trivial solution $(N_0, \ldots, N_m)$ to Eq. (3), it suffices to verify that the associated $(m+1) \times (m+1)$-matrix does not have full rank. Now here enters the essential heuristic argument: it follows from the definition of $NB^{(HY)}$ that

$$\sum_{j=1}^{m+1} w_j^{(m)}\, V^{(0)}\big(\bar{t}_j\big) \approx B = \sum_{j=1}^{m+1} w_j^{(m)}\, B,$$

$$\sum_{j=1}^{m+1} w_j^{(m)}\, V^{(i)}\big(\bar{t}_j\big) \approx \mathrm{upf}(t_i) = \sum_{j=1}^{m+1} w_j^{(m)}\, \mathrm{upf}(t_i), \quad i = 1, \ldots, m,$$

with the approximations becoming equalities as $m \to \infty$. In other words, this means that the rows of the equation system (3) are linearly dependent. Consequently, the associated matrix cannot have full rank and the columns must also be linearly dependent, i.e. there exists a non-zero solution $(N_0, \ldots, N_m)$ of Eq. (3), and hence (2), as desired. Finally, taking a close look at the structure of the involved cash flows, it is obvious that a solution must satisfy $N_0 \neq 0$. Without loss of generality we may hence set $N_0 = 1$ (because if $(N_0, \ldots, N_m)$ is a solution, so is $\alpha\ (N_0, \ldots, N_m)$ for arbitrary $\alpha \in \mathbb{R}$). Concluding, the portfolio we have found is long the bond. □

We present an example that demonstrates how different the three presented measurements of negative basis can be in practice. The specifications are inspired by a real-world case.

Example 1 We consider a bond with maturity $T = 3.5$ years paying a semi-annual coupon rate of $C = 8.25\,\%$. It trades far below par value, namely at $B = 46.5\,\%$. An almost maturity-matched CDS contract is available at an upfront value of $\text{upf}(T) = 53\,\%$ with a running coupon of $s(T) = 5\,\%$, payed quarterly. This means a nominal-matched negative basis investment comes at a package price of $46.5 + 53 = 99.5\,\%$, and pays a coupon rate of $8.25 - 5 = 3.25\,\%$ until default (however, the bond and CDS coupon payments have different frequencies and payment dates). In the sequel we assume a recovery rate of $R = 20\,\%$, and the reference rate $r(.)$ is bootstrapped from 3-month tenor-based interest rate swaps according to the raw interpolation method described in [15, 16]. Because the bond trades far below par, the measurement (Z) is highly questionable and returns $NB^{(Z)} = -0.42\,\%$, which is clearly not an appropriate measurement. As indicated earlier, improved versions of earnings and costs-measurements must be used in order to deal with such extreme situations of highly distressed bonds, but this lies outside the scope of the present article. The par-equivalent CDS methodology returns the measurement $NB^{(PE)} = 2.29\,\%$, whereas the hidden yield methodology returns the significantly lower number $NB^{(HY)} = 1.18\,\%$. While the authors are not aware of a strategy how to monetize the (PE)-measurement 2.29 %, Lemma 1 provides a clear interpretation for the (HY)-measurement 1.18 % in terms of an internal rate of return that can be earned on top of the risk-free rate, when the negative basis investment is structured as indicated in the proof of Lemma 1.

Now if the described nominal-matched investment seems to earn a rate of 3.25 %, which equals a spread of around 1.75 % above the chosen reference rate $r(.)$ in the present example, why is the measurement $NB^{(HY)}$ so low? Fig. 1 visualizes the discounted value of the sum over all cash flows from the nominal-matched investment in dependence of the default time. For instance, in case of survival until maturity, this value equals approximately 104 %, yielding a return (after discounting) of 5.61 % on the initial investment of 98.39 % (which equals the package price minus accrued CDS coupon, the bond accrued equals zero). Distributed on the 3.5-year investment horizon, this corresponds to a rate of approximately 1.6 % per annum. However, in case of a default just before the first or second bond coupon payment date the

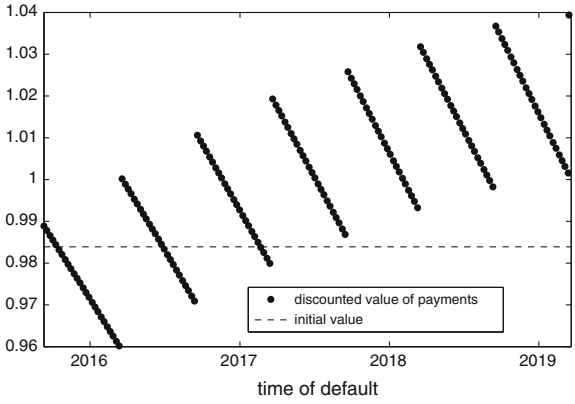

Fig. 1 Sum over all discounted cash flows arising from the described nominal-matched negative basis investment are depicted, in dependence on the time of default

described negative basis investment faces a loss. The additional short-dated CDS-protection required in order to hedge these potential losses decreases the earnings potential of the investment, which is accounted for in the (HY)-methodology, as explained in the proof of Lemma 1.

6 Conclusion

We proposed an innovative measurement for the negative basis, denoted $NB^{(HY)}$. Compared to traditional approaches, it is based on an arbitrage-free pricing model for the simultaneous pricing of the bond and the CDS, which provides a sound economic interpretation. Within a simple model with only default risk being present, the negative basis is perfectly explained as the spread on top of a reference interest rate curve $r(.)$. It was pointed out how the rate $r(.) + NB^{(HY)}$ can be earned without exposure to default risk.

Acknowledgements The KPMG Center of Excellence in Risk Management is acknowledged for organizing the conference "Challenges in Derivatives Markets - Fixed Income Modeling, Valuation Adjustments, Risk Management, and Regulation".

Appendix: The algorithm in Definition 3 is well-defined

The following technical lemma guarantees that Step 3 in Definition 3 admits a unique solution that can be found efficiently by means of a bisection routine.

Lemma A.1 (Method (HY) is well-defined)

(a) The function $x \mapsto Bond(\lambda_x(.), r(.) + x, R, C, T)$ *is continuous.*

(b) The function $x \mapsto Bond(\lambda_x(.), r(.) + x, R, C, T)$ *is decreasing on the interval* $[-\inf\{r(t) : t \geq 0\}, \infty)$.

(c) We have the lower bound $Bond(\lambda_x(.), r(.) + x, R, C, T) \geq \frac{R}{1-R}\, upf(T)$.

Proof We prove parts (a), (b), and (c) separately.

(a) For fixed x, the function $\lambda_x(.)$ is piecewise constant, so actually we only deal with a finite vector of values of the default intensity, depending on x. For the remainder of the proof we denote these values by $(y_1(x), \ldots, y_m(x))$. In other words, we observe m CDS maturities $T_1, \ldots, T_m$ and the value $y_k(x)$ is the level of the default intensity on the piece $(T_{k-1}, T_k]$, for $k = 1, \ldots, m$, with $T_0 := 0$. Obviously, the bond price then equals a concatenation of continuous functions if each $y_k(x)$ is continuous in x. However, this is guaranteed by the implicit function theorem since $y_k(x)$ is defined as the implicit function yielding the root of a smooth function. Concluding, continuity of the bond price is clear.

(b) In order to see that the bond price is decreasing in x, we first re-write it as

$$\begin{aligned} \text{Bond}(\lambda_x(.), r(.) + x, R, C, T) = {} & e^{-\int_0^T \lambda_x(s)+r(s)+x\,\mathrm{d}s} \\ & + C \sum_{0<t_j^{(B)} \leq T} \left(t_j^{(B)} - t_{j-1}^{(B)}\right) e^{-\int_0^{t_j^{(B)}} \lambda_x(s)+r(s)+x\,\mathrm{d}s} \\ & + \frac{R}{1-R}\, EDPL(\lambda_x(.), r(.) + x, \text{upf}(T), T), \end{aligned}$$

where we have used $EDPL = EDDL$ from the CDS boostrap. This shows that it suffices to check that the function

$$x \mapsto \lambda_x(t) + x$$

is increasing for each fixed t, because all summands in the above bond formula are then obviously decreasing.

We proceed with an auxiliary observation. If τ_1 and τ_2 are two positive random variables with distribution functions F_1 and F_2, satisfying $F_1 \geq F_2$ pointwise on an interval (T, ∞) and $F_1 \equiv F_2$ on $[0, T]$, then $\mathbb{E}[g(\tau_1)] \geq \mathbb{E}[g(\tau_2)]$ for any bounded function $g : (0, \infty) \to [0, K]$, which is non-increasing on (T, ∞). To verify this,[11] define the non-decreasing function $h := -g$ and use integration by parts:

[11] One says that τ_1 is less than τ_2 in the usual stochastic order, and the following computation is standard in the respective theory.

$$\begin{aligned}\mathbb{E}[g(\tau_1)] &= -\int h\,\mathrm{d}F_1 = -\int_{(0,T]} h\,\mathrm{d}F_1 - \int_{(T,\infty)} h\,\mathrm{d}F_1 \\ &= -\int_{(0,T]} h\,\mathrm{d}F_2 - \Big(h(\infty)\underbrace{F_1(\infty)}_{=1=F_2(\infty)} - h(T)\underbrace{F_1(T)}_{=F_2(T)} - \int_{(T,\infty)} F_1\,\mathrm{d}h\Big) \\ &\geq -\int_{(0,T]} h\,\mathrm{d}F_2 - \Big(h(\infty)\,F_2(\infty) - h(T)\,F_2(T) - \int_{(T,\infty)} F_2\,\mathrm{d}h\Big) \\ &= -\int_{(0,T]} h\,\mathrm{d}F_2 - \int_{(T,\infty)} h\,\mathrm{d}F_2 = -\int h\,\mathrm{d}F_2 = \mathbb{E}[g(\tau_2)].\end{aligned}$$

Now we proceed inductively over $k = 1, \ldots, m$ by showing that $x \mapsto \lambda_x(t) + x$ is non-decreasing for all fixed $t \in (T_{k-1}, T_k]$, i.e. that $x \mapsto y_k(x) + x$ is non-decreasing. We start the induction for $k = 1$. To this end, recall that $y_1(x)$ is the unique root of the equation

$$EDPL(y_1(x), r(.) + x, s(T_1), \mathrm{upf}(T_1), T_1) = EDDL(y_1(x), r(.) + x, R, T_1).$$

For the sake of a more compact notation we denote the left-hand side of the last equation by $LHS(x, y_1(x))$ and the right-hand side by $RHS(x, y_1(x))$. Furthermore, we denote the value of both sides by $V(x) := LHS(x, y_1(x)) = RHS(x, y_1(x))$. Since all the summands of LHS depend on the function $x \mapsto x + y_1(x)$ in a monotonic way, it is obvious that $V(x)$ is non-increasing in x if and only if the function $x \mapsto x + y_1(x)$ is non-decreasing. Hence, it suffices to prove that $V(x)$ is non-increasing in x. To this end, we (obviously) observe with $\varepsilon > 0$ that

$$LHS(x + \varepsilon, y_1(x)) \leq LHS(x, y_1(x)) = V(x), \tag{4}$$
$$RHS(x + \varepsilon, y_1(x)) \leq RHS(x, y_1(x)) = V(x). \tag{5}$$

Furthermore, the function $y \mapsto LHS(x + \varepsilon, y)$ is obviously strictly decreasing. Concerning the right-hand side, we denote by $\mathbb{E}_y[f(\tau)]$ the expectation over $f(\tau)$ when the default time τ has an exponential distribution with parameter y. The function

$$y \mapsto RHS(x + \varepsilon, y) = (1 - R)\,\mathbb{E}_y\Big[e^{-\int_0^\tau r(s)+x+\varepsilon\,\mathrm{d}s}\,1_{\{\tau \leq T_1\}}\Big]$$

is non-decreasing on the claimed interval by the auxiliary observation we have derived above (increasing y corresponds to increasing the distribution function of the default time τ pointwise[12]). We now distinguish two cases:

[12]Here, we have used that the function $\tau \mapsto \exp(-\int_0^\tau r(s) + x + \varepsilon\,\mathrm{d}s)\,1_{\{\tau \leq T_1\}}$ is non-increasing if $x \geq -\inf\{r(t) : t \geq 0\}$.

(i) $LHS(x+\varepsilon, y_1(x)) \leq RHS(x+\varepsilon, y_1(x))$:
In this case $y_1(x+\varepsilon) \leq y_1(x)$, because otherwise we would observe the following contradiction:

$$\begin{aligned} LHS(x+\varepsilon, y_1(x+\varepsilon)) &< LHS(x+\varepsilon, y_1(x)) \leq RHS(x+\varepsilon, y_1(x)) \\ &\leq RHS(x+\varepsilon, y_1(x+\varepsilon)). \end{aligned}$$

This implies that

$$V(x+\varepsilon) = RHS(x+\varepsilon, y_1(x+\varepsilon)) \leq RHS(x+\varepsilon, y_1(x)) \overset{(5)}{\leq} V(x).$$

(ii) $LHS(x+\varepsilon, y_1(x)) > RHS(x+\varepsilon, y_1(x))$:
In this case $y_1(x+\varepsilon) \geq y_1(x)$, because otherwise we would observe the following contradiction:

$$\begin{aligned} RHS(x+\varepsilon, y_1(x+\varepsilon)) &\leq RHS(x+\varepsilon, y_1(x)) < LHS(x+\varepsilon, y_1(x)) \\ &\leq LHS(x+\varepsilon, y_1(x+\varepsilon)). \end{aligned}$$

This implies that

$$V(x+\varepsilon) = LHS(x+\varepsilon, y_1(x+\varepsilon)) \leq LHS(x+\varepsilon, y_1(x)) \overset{(4)}{\leq} V(x).$$

Concluding, $V(x)$ is non-increasing in x and the induction start is finished.
We proceed with the induction step, assuming that we already know that $x + \lambda_x(t)$ is non-decreasing in x for each fixed $t \leq T_{k-1}$. To this end, recall that $y_k(x)$ is the unique root of the equation

$$EDPL(\lambda_x(.), r(.)+x, s(T_k), \mathrm{upf}(T_k), T_k) = EDDL(\lambda_x(.), r(.)+x, R, T_k),$$

where $y_k(x)$ enters the equation as the function value of $\lambda_x(.)$ on the interval $(T_{k-1}, T_k]$. The left-hand side of the last equation can be rewritten as follows, using the standard market convention of standardized CDS strike rates $s(T_{k-1}) = s(T_k) =: s$:

$$\begin{aligned} &EDPL(\lambda_x(.), r(.)+x, s, \mathrm{upf}(T_k), T_k) \\ &\quad = EDPL(\lambda_x(.), r(.)+x, s, upf(T_{k-1}), T_{k-1}) + \mathrm{upf}(T_k) - \mathrm{upf}(T_{k-1}) \\ &\qquad + s \sum_{T_{k-1} < t_i^{(C)} \leq T_k} \left(t_i^{(C)} - t_{i-1}^{(C)}\right) e^{-\int_0^{t_i^{(C)}} \lambda_x(s) + r(s) + x \, \mathrm{d}s}. \end{aligned}$$

Similarly, the right-hand side can be rewritten as follows:

$$\begin{aligned} EDDL(\lambda_x(.), r(.)+x, R, T_k) &= EDDL(\lambda_x(.), r(.)+x, R, T_{k-1}) \\ &\quad + (1-R)\, y_k(x) \int_{T_{k-1}}^{T_k} e^{-\int_0^t r(s)+x+\lambda_x(s) \, \mathrm{d}s} \, \mathrm{d}t. \end{aligned}$$

Since the values $(y_1(x), \ldots, y_{k-1}(x))$ have been determined before, we may subtract the *EDDL* and *EDPL* with maturity T_{k-1} on both sides of the defining equation for $y_k(x)$, simplifying the latter to

$$\begin{aligned}&\mathrm{upf}(T_k) - \mathrm{upf}(T_{k-1}) + s \sum_{T_{k-1} < t_i^{(C)} \le T_k} \left(t_i^{(C)} - t_{i-1}^{(C)}\right) e^{-\int_0^{t_i^{(C)}} \lambda_x(s) + r(s) + x \,\mathrm{d}s} \\ &\quad = (1-R)\, y_k(x) \int_{T_{k-1}}^{T_k} e^{-\int_0^t r(s) + x + \lambda_x(s)\,\mathrm{d}s}\,\mathrm{d}t.\end{aligned}$$

Again, we denote the left-hand side of the last equation by $LHS(x, y_1(x), \ldots, y_k(x))$, and the right-hand side is denoted $RHS(x, y_1(x), \ldots, y_k(x))$. Furthermore, we denote the value of both sides by

$$V(x) := LHS(x, y_1(x), \ldots, y_k(x)) = RHS(x, y_1(x), \ldots, y_k(x)).$$

By induction hypothesis, the function $x \mapsto x + \lambda_x(t)$ is non-decreasing for each $t \le T_{k-1}$. With $\varepsilon > 0$ this obviously implies that

$$\begin{aligned}&LHS(x+\varepsilon, y_1(x+\varepsilon), \ldots, y_{k-1}(x+\varepsilon), y_k(x)) \\ &\quad \le LHS(x, y_1(x), \ldots, y_k(x)) = V(x), \qquad (6) \\ &RHS(x+\varepsilon, y_1(x+\varepsilon), \ldots, y_{k-1}(x+\varepsilon), y_k(x)) \\ &\quad \le RHS(x, y_1(x), \ldots, y_k(x)) = V(x). \qquad (7)\end{aligned}$$

Also, the function $y \mapsto LHS(x+\varepsilon, y_1(x+\varepsilon), \ldots, y_{k-1}(x+\varepsilon), y)$ is obviously non-increasing, whereas the function $y \mapsto RHS(x+\varepsilon, y_1(x+\varepsilon), \ldots, y_{k-1}(x+\varepsilon), y)$ is non-decreasing by a similar argument as in the induction start, namely: the right-hand side has the form[13]

$$\begin{aligned}&RHS(x+\varepsilon, y_1(x+\varepsilon), \ldots, y_{k-1}(x+\varepsilon), y) \\ &\quad = (1-R)\, \mathbb{E}_y\left[e^{-\int_0^\tau r(s) + x + \varepsilon \,\mathrm{d}s}\, 1_{\{\tau \in (T_{k-1}, T_k]\}}\right],\end{aligned}$$

which is non-decreasing in y. Why? Because an increase of y increases the distribution function of τ pointwise on $[T_{k-1}, \infty)$ but leaves it unchanged on $[0, T_{k-1}]$, and the function $\tau \mapsto \exp(-\int_0^\tau r(s) + x\,\mathrm{d}s)\, 1_{\{\tau \in (T_{k-1}, T_k]\}}$ is clearly non-increasing on (T_{k-1}, ∞) (so that our auxiliary observation above applies). Like in the induction start, showing that $x \mapsto x + y_k(x)$ is non-decreasing in x is equivalent to showing that $V(x)$ is non-increasing in x. The remaining proof is now completely analogous to the induction start (this is an exercise we leave to the reader).

[13] Similar as in the induction start, we denote by $\mathbb{E}_y[f(\tau)]$ the expectation over $f(\tau)$ when the default time has piecewise constant intensity with the level y on the piece $(T_{k-1}, T_k]$.

(c) Denoting by $\mathbb{Q}_x$ the probability measure in dependence of the default intensities $\lambda_x(.)$, we have

$$\begin{aligned}\text{Bond}(\lambda_x(.), r(.)+x, R, C, T) := C \sum_{0<t_j^{(B)} \le T} DF\big(t_j^{(B)}\big)\,\big(t_j^{(B)} - t_{j-1}^{(B)}\big)\,\mathbb{Q}_x\big(\tau > t_j^{(B)}\big) \\ + DF(T)\,\mathbb{Q}_x(\tau > T) + R\,\mathbb{E}_x[1_{\{\tau \le T\}}\,DF(\tau)].\end{aligned}$$

We know from the consistent CDS pricing that the appearing expectation can be replaced by the premium leg of the CDS, which allows to be estimated by the upfront, i.e.

$$\begin{aligned}\mathbb{E}_x[1_{\{\tau \le T\}}\,DF(\tau)] &= \frac{R}{1-R}\,EDPL(\lambda_x(.), r(.)+x, s(T), \text{upf}(T), T) \\ &\ge \frac{R}{1-R}\,\text{upf}(T),\end{aligned}$$

which in turn implies the claim.

References

1. Andritzky, J., Singh, M.: The pricing of credit default swaps during distress. IMF Working Paper 06/254 (2006)
2. Bai, J., Collin-Dufresne, P.: The determinants of the CDS-Bond basis during the financial crises of 2007–2009. Working Paper (2011)
3. Beinstein, E., Scott, A., Graves, B., Sbityakov, A., Le, K., Goulden, J., Muench, D., Doctor, S., Granger, A., Saltuk, Y., Allen, P.: Credit Derivatives Handbook. J.P. Morgan Corporate Quantitative Research (2006)
4. Blanco, R., Brennan, S., Marsh, I.W.: An empirical analysis of the dynamic relation between investment-grade bonds and credit default swaps. J. Financ. **60**(5), 2255–2281 (2005)
5. Brigo, D., Capponi, A., Pallavicini, A.: Arbitrage-free bilateral counterparty risk valuation under collateralization and application to Credit Default Swaps. Math. Financ. **24**(1), 125–146 (2014)
6. Bühler, W., Trapp, M.: Explaining the bond-CDS basis-the role of credit risk and liquidity. In: Risikomanagement und kapitalmarktorientierte Finanzierung: Festschrift für Bernd Rudolph zum 65. Geburtstag, Knapp-Verlag, Frankfurt a. Main, pp. 375–397 (2009)
7. Burgard, C., Kjaer, M.: In the balance. Risk pp. 72–75 (2011)

8. Choudhry, M.: The Credit Default Swap Basis. Bloomberg Press, New York (2006)
9. Das, S., Kalimipalli, M.: Did CDS trading improve the market for corporate bonds? J. Financ. Econ. **111**(2), 495–525 (2014)
10. De Wit, J.: Exploring the CDS-Bond basis. National Bank of Belgium working paper No. 104 (2006)
11. Doctor, S., White, D., Elizalde, A., Goulden, J., Toublan, D.D.: Differential discounting for CDS: J.P. Morgan Europe Credit Research (2012)
12. Fries, C.: Discounting Revisited: Valuation Under Funding, Counterparty Risk and Collateralization. Working paper (2010), available at SSRN: http://ssrn.com/abstract=1609587
13. Fujii, M., Shimada, Y., Takahashi, A.: Collateral Posting and Choice of Collateral Currency. CIRJE Discussion Papers (2010)
14. Gupta, S., Sundaram, R.K.: Mispricing and arbitrage in CDS auctions. J. Deriv. **22**(4), 79–91 (2015)
15. Hagan, P.S., West, G.: Interpolation methods for curve construction. Appl. Math. Financ. **13**(2), 89–129 (2006)
16. Hagan, P.S., West, G.: Methods for constructing a yield curve. Wilmott magazine pp.70–81 (2008)
17. Höcht, S., Kunze, M., Scherer, M.: Implied recovery rates-auction and models. In: Innovations in Quantitative Risk Management. Springer, Berlin pp. 147–162 (2015)
18. Hull, J., White, A.: Valuing credit default swaps I: no counterparty default risk. J. Deriv. **8**(1), 29–40 (2000)
19. Hull, J., Predescu, M., White, A.: The relationship between credit default swap spreads, bond yields, and credit rating announcements. J. Bank. Financ. **28**, 2789–2811 (2004)
20. Li, H., Zhang, W., Kim, G.H.: The CDS-Bond basis and the cross section of corporate bond returns. Working paper (2011)
21. Longstaff, F.A., Neis, E., Mithal, S.: Corporate yield spreads: default risk or liquidity? New evidence from the credit-default swap market. J. Financ. **60**(5), 2213–2253 (2005)
22. Mai, J.-F., Scherer, M.: Simulating from the copula that generates the maximal probability for a joint default under given (inhomogeneous) marginals. In: Melas, V.B. et al.: Topics in Statistical Simulation: Springer Proceedings in Mathematics and Statistics, vol. 114, pp. 333–341, Springer, Heidelberg (2014)
23. Morini, M., Prampolini, A.: Risky funding with counterparty and liquidity charges. Risk pp. 58–63 (2011)
24. O'Kane, D.: The link between Eurozone sovereign debt and CDS prices. EDHEC-Risk Institute Working Paper (2012)
25. O'Kane, D., Turnbull, S.: Valuation of credit default swaps. Fixed Income Quantitative Research Lehman Brothers (2003)
26. Palladini, G., Portes, R.: Sovereign CDS and bond pricing dynamics in the Euro-area. Centre for Economic Policy Research Discussion Paper No. 8651 (2011)
27. Pallavicini, A., Perini, D., Brigo, D.: Funding Valuation Adjustment: A Consistent Framework Including CVA, DVA, Collateral, Netting Rules and Re-Hypothecation. Working paper (2011), available at SSRN: http://ssrn.com/abstract=1969114
28. Pedersen, C.M.: Explaining the Lehman Brothers option adjusted spread of a corporate bond. Fixed Income Quantitative Credit Research, Lehman Brothers (2006)
29. Piterbarg, V.: Funding beyond discounting: collateral agreements and derivatives pricing. Risk, pp. 97–102 (2010)
30. Zhu, H.: An empirical comparison of credit spreads between the bond market and the credit default swap market. BIS Working Paper No. 160 (2004)

The Impact of a New CoCo Issuance on the Price Performance of Outstanding CoCos

Jan De Spiegeleer, Stephan Höcht, Ine Marquet and Wim Schoutens

Abstract Contingent convertible bonds (CoCos) are new hybrid capital instruments that have a loss absorbing capacity which is enforced either automatically via the breaching of a particular CET1 level or via a regulatory trigger. The price performance of outstanding CoCos, after a new CoCo issue is announced by the same issuer, is investigated in this paper via two methods. The first method compares the returns of the outstanding CoCos after an announcement of a new issue with some overall CoCo indices. This method does not take into account idiosyncratic movements and basically compares with the general trend. A second model-based method compares the actual market performance of the outstanding CoCos with a theoretical model. The main conclusion of the investigation of 24 cases of new CoCo bond issues is a moderated negative effect on the outstanding CoCos.

Keywords Contingent convertibles · CoCo bonds · New issuance

1 Introduction

Contingent convertible bonds or CoCo bonds are new hybrid capital instruments that have a loss absorbing capacity which is enforced either automatically via the breaching of a particular CET1 level or via a regulatory trigger. CoCos either convert

J. De Spiegeleer (✉) · I. Marquet · W. Schoutens
Department of Mathematics, KU Leuven, Celestijnenlaan 200B,
Box 2400, 3001 Leuven, Belgium
e-mail: Jan.Spiegeleer@mac.com

I. Marquet
e-mail: Ine.Marquet@wis.kuleuven.be

W. Schoutens
e-mail: wim@schoutens.be

S. Höcht
Assenagon GmbH, Prannerstraße 8, 80333 München, Germany
e-mail: Stephan.Hoecht@assenagon.com

K. Glau et al. (eds.), *Innovations in Derivatives Markets*, Springer Proceedings
in Mathematics & Statistics 165, DOI 10.1007/978-3-319-33446-2_19

into equity or suffer a write-down of the face value upon the appearance of such a trigger event.

The financial crisis of 2007–2008 triggered an avalanche of financial worries for financial institutions around the globe. After the collapse of Lehman Brothers, governments intervened and bailed out banks using tax-payers money. Preventing such bail-outs in the future, and designing a more stable banking sector in general, requires both higher capital levels and regulatory capital of a higher quality. The implementation under the new regulatory frameworks like Basel III and Capital Requirement Directive IV (CRD IV) tries to achieve this in various ways, i.e. with the use of CoCo bonds (Basel Committee on Banking Supervision [1], European Commision [2]). CoCo bonds are allowed as new capital instruments by the Basel III guidelines. The Swiss regulators have forced their systemic important banks to issue large amounts of these instruments. Further, the European CRD IV which entered into force on 17 July 2013 enforces all new additional Tier 1 instruments to have CoCo features.

The specific design of a CoCo bond enhances the capital of a bank when it is in trouble in an automatic way. Hence, a loss-absorbing cushion is created with the aim to avoid or at least to reduce potential interventions using tax-payers' money.

The first CoCos have been issued in the aftermath of the credit crisis. In December 2009 Lloyds exchanged some of their old hybrid instruments into this new type of bonds in order to strengthen their capital position after the bank had been hit very hard due to the financial crisis of 2008. Since then a lot of other banks have been issuing CoCos and one is expecting that many will follow in the next years. The market of CoCos is currently above USD 100 bn and is expanding very rapidly.[1]

When an issuer has already some CoCos outstanding and is announcing the issuance of a new CoCo bond, there are at least two opposite forces at work. On one hand, a new issue means that the capital of the issuing institute is strengthened (at the additional Tier 1 or Tier 2 level). Due to the new issue, the losses in case of a future trigger event will be shared over a larger bucket and hence recovery rates are expected to be higher. On the other hand, there are the market dynamics and investors who often prefer to invest rather in the new CoCo than in the older ones. This can be just due to the fact that one prefers new things above old stuff, but also because one believes there is a basis spread to be earned on a new issuance. Some believe a new issuance is brought to the market with a certain discount, to attract investors and to make the whole capital raising exercise a success. Investors then will move out of the old bonds and ask for allocation in the new issue.

In this paper, we estimate the price impact on the outstanding CoCos via two methods. The first method compares the returns of the outstanding CoCo bonds after an announcement of a new issue with some overall CoCo indices. More precisely, we compare the performance with CS Contingent Convertible Euro Total Return index and the BofA Merrill Lynch Contingent Capital index. Here we basically compare the performance of the outstanding CoCos with the general market performance. However such a comparison does not take into account idiosyncratic movements; it

[1] Source: Bloomberg.

basically compares with the general market trend. The issuing company is nevertheless exposed to market dynamics. Its stock price, its credit worthiness etc. can exhibit different timely evolutions compared with the respective quantities of their competitors. This can be especially the case around capital raising announcements since then financial details of the company are published and discussed at, for example roadshows around the new issuance. Therefore, we also deploy a second methodology taking into account idiosyncratic movements. Using an equity derivatives model, we compare the actual market performance of the outstanding CoCo bonds, with a theoretical model performance taking into account idiosyncratic effects, like movements in the underlying stock, credit default spreads or volatilities. The model is derivatives based and is taking as such forward-looking expectations into account.

In total, we investigate 24 cases of new CoCo bond issues. The main conclusion of the investigation is that there is a moderated negative effect on outstanding CoCo bonds. This is confirmed by both methodologies and the impact is an underperformance of about 25–50 bps on average in between the announcement date and the issue date. An extra negative impact of 40 bps was observed in the 10 trading days after the issue.

The analysis in this paper is constrained to CoCo bonds only, but a similar study could be done for other types of bonds as well. A comparative study for corporate bonds was, e.g. done in Akigbe et al. [3], where the authors investigate the impact of 574 outstanding debt issues. The investigation was divided by different reasons of a new debt issue. A significant negative impact on the price of the outstanding debt and equity was observed in case the public debt securities were issued to finance unexpected cash flow shortfalls. No significant reaction was observed when the new debt issues were motivated by unexpected increase in capital expenditures, unexpected increase in leverage or expected refinancing of outstanding debt.

This paper is organized as follows. We first provide in the next section the details of the equity derivatives model. In Sect. 3, we provide details on the data set used and in particular overview the new issuances of a whole battery of issuers that are part of our study. Next, we report on the exact methodology and results of our comparison with other CoCo indices. The final part of that section reports and discusses the results of the Equity Derivatives model. The final section concludes.

2 The Equity Derivatives Model

CoCos are hybrid instruments, with characteristics of both debt and equity. This gives rise to different approaches for pricing CoCos. Without considering the heuristic models, two main schools of thoughts exist, namely the structural models and market-implied models. Structural models are based on the theory of Merton and can be found in Pennacchi [4] and Pennacchi et al. [5]. We will apply a market-implied model where the derivation is based on market data such as share prices, credit default spreads and volatilities. The models were introduced in a Black–Scholes framework in De Spiegeleer and Schoutens [6] and De Spiegeleer et al. [7]. Pricing CoCos under

smile conform models can be found in Corcuera et al. [8]. Based on the Heston model, the impact of skew is discussed in De Spiegeleer et al. [9]. In De Spiegeleer et al. [10] the implied CET1 volatility is derived from the market price of a CoCo bond. Further extensions and variations can be found in De Spiegeleer and Schoutens [11, 12], Corcuera et al. [13], De Spiegeleer and Schoutens [14], Cheridito and Zhikai [15], Madan and Schoutens [16].

The actual valuation of a CoCo incorporates the modelling of both the trigger probability and the expected loss for the investor. Notice that the trigger is defined by a particular CET1 level or decided upon a regulator's decision. Since these trigger mechanisms are hard to model or even quantify, we project the trigger into the stock price framework as considered in the equity derivatives model of De Spiegeleer and Schoutens [6]. This means that the CoCo will be triggered under the model once the share price drops below a specified barrier level, denoted by $S^\star$. We infer from existing CoCo market data the share price at the moment the CoCo bond gets triggered and we will call this the (implied) trigger level. As a result the valuation of a CoCo bond is transformed into a barrier-pricing exercise in an equity setting.

Under such a framework the CoCo bond can be broken down to several different derivative instruments. In first place, the CoCo behaves like a standard (non-defaultable) corporate bond where the holder will receive coupons c_i on regular time points t_i together with the principal N at maturity T. However, in case the share price drops below the trigger level $S^\star$, the investor will lose his initial investment and all future coupons. This will be modelled by short positions in binary down-and-in (BIDINO) options with maturities t_i for each coupon c_i and a BIDINO with maturity T to model the cancelling of the initial value. After the trigger event has occurred, the investor of a conversion CoCo will receive C_r shares. We can model this with C_r down-and-in asset-(at hit)-or-nothing options on the stock. For a write-down CoCo, the investor does not receive any shares and we can just set C_r equal to zero in this case. Therefore, the price of a CoCo can be calculated with the following formula:

$$
\begin{aligned}
P = {} & \text{Corporate bond} \\
& - N \times \text{binary down-and-in option} \\
& - \sum_i c_i \times \text{binary down-and-in option} \\
& + C_r \times \text{down-and-in asset-(at hit)-or-nothing option on the stock}
\end{aligned}
$$

Under the Black–Scholes model, we can find an explicit formula for the price of the CoCo at time t:

$$
\begin{aligned}
P = {} & N \exp(-r(T-t)) + \sum_{i=1}^{k} c_i \exp(-r(t_i - t)) \\
& - N \times \exp(-r(T-t))[\Phi(-x_1 + \sigma\sqrt{T-t}) + (S^\star/S)^{2\lambda-2}\Phi(y_1 - \sigma\sqrt{T-t})] \\
& - \sum_i c_i \times \exp(-r(t_i - t))[\Phi(-x_{1i} + \sigma\sqrt{t_i - t}) + (S^\star/S)^{2\lambda-2}\Phi(y_{1i} - \sigma\sqrt{t_i - t})] \\
& + C_r \times S^\star \left[\left(\frac{S^\star}{S}\right)^{a+b} \Phi(z) + \left(\frac{S^\star}{S}\right)^{a-b} \Phi(z - 2b\sigma\sqrt{T-t}) \right] \qquad (1)
\end{aligned}
$$

with

$$z = \frac{\log(S^\star/S)}{\sigma\sqrt{T-t}} + b\sigma\sqrt{T-t} \qquad x_1 = \frac{\log(S/S^\star)}{\sigma\sqrt{T-t}} + \lambda\sigma\sqrt{T-t}$$

$$a = \frac{r - q - \frac{1}{2}\sigma^2}{\sigma^2} \qquad y_1 = \frac{\log(S^\star/S)}{\sigma\sqrt{T-t}} + \lambda\sigma\sqrt{T-t}$$

$$b = \frac{\sqrt{(r - q - \frac{1}{2}\sigma^2)^2 + 2r\sigma^2}}{\sigma^2} \qquad x_{1i} = \frac{\log(S/S^\star)}{\sigma\sqrt{t_i-t}} + \lambda\sigma\sqrt{t_i-t}$$

$$\lambda = \frac{r - q + \sigma^2/2}{\sigma^2} \qquad y_{1i} = \frac{\log(S^\star/S)}{\sigma\sqrt{t_i-t}} + \lambda\sigma\sqrt{t_i-t}$$

where Φ is the cdf of a standard normal distribution, r is the risk free rate, q the dividend yield and σ the volatility.

Applying this equity derivatives pricing model, a CoCo price can be found for a trigger level $S^\star$. However, the other way around is often more interesting. Knowing the market CoCo price, we can filter out an implied trigger $\widehat{S^\star}$ in such a way that market and model price match. Since CoCos of one financial institution with the same contractual trigger should trigger at the same time, their implied trigger levels should theoretically also be the same. Hence the implied barriers give us a way to compare different CoCos in order to detect over- or undervaluation, irrespectively of different currencies and maturities.

Our goal is to compare the actual market performance of the outstanding CoCo bonds with the theoretical model performance. This theoretical price takes idiosyncratic effects into account. Any changes in the actual market performance compared to the theoretical model performance will be described to the effect of the announcement of a new CoCo issuance. The research can also be translated in terms of implied trigger levels. In case the new CoCo does not influence the outstanding CoCo, the implied barrier of the outstanding CoCo should remain constant. Whereas if its implied barrier derived from the market will change, this change will be caused by the new CoCo issuance.

3 Measuring the Price Performance of the Outstanding CoCos

3.1 New Issuances

The impact of a new CoCo issuance is investigated on the outstanding CoCos of the same issuing company. The issuers in our study contain UBS, Barclays, Crédit Agricole, Sociéty Général, Deutsche Bank, UniCredit, Credit Suisse, Santander, Rabobank, Danske and BBVA. The effect on the outstanding CoCos is investigated in the period between announcement and issuance of the new CoCo, which are summarised in Table 1. Notice that UBS, Barclays and Crédit Agricole all have

Table 1 Announcement date, issue date and issue size (in bn) of the new CoCos

Name	ISIN	Announc.	Issue	Size	Curr.
ACAFP 6 5/8 09/29/49	USF22797YK86	11/09/2014	18/09/2014	1,250	USD
ACAFP 6 1/2 04/29/49	XS1055037177[a]	01/04/2014	08/04/2014	1,000	EUR
ACAFP 7 7/8 01/29/49	USF22797RT78	15/01/2014	23/01/2014	1,750	USD
BACR 7 06/15/49	XS1068561098[b]	13/06/2014	17/06/2014	697.60	GBP
BACR 8 12/15/49	XS1002801758	03/12/2013	10/12/2013	1,000	EUR
BACR 8 1/4 12/29/49	US06738EAA38	13/11/2013	20/11/2013	2,000	USD
BBVASM 6 3/4 12/29/49	XS1190663952	10/02/2015	18/02/2015	1,500	EUR
CS 6 1/4 12/29/49	XS1076957700	10/06/2014	18/06/2014	2,500	USD
CS 7 1/2 12/29/49	XS0989394589	04/12/2013	11/12/2013	2,250	USD
CS 5 3/4 09/18/25	XS0972523947	11/09/2013	18/09/2013	1,250	EUR
CS 6 09/29/49	CH0221803791	20/08/2013	04/09/2013	290	CHF
DANBNK 5 7/8 04/29/49	XS1190987427	11/02/2015	18/02/2015	750	EUR
DB 7 1/2 12/29/49	US251525AN16	18/11/2014	21/11/2014	1,500	USD
RABOBK 5 1/2 01/22/49	XS1171914515	15/01/2015	22/01/2015	1,500	EUR
SANTAN 6 1/4 09/11/49	XS1107291541	02/09/2014	11/09/2014	1,500	EUR
SANTAN 6 3/8 05/29/49	XS1066553329	08/05/2014	19/05/2014	1,500	USD
SOCGEN 6 10/27/49	USF8586CXG25	19/06/2014	25/06/2014	1,500	USD
SOCGEN 6 3/4 04/07/49	XS0867620725	28/03/2014	07/04/2014	1,000	EUR
SOCGEN 7 7/8 12/29/49	USF8586CRW49	11/12/2013	18/12/2013	1,750	USD
UBS 7 1/8 12/29/49	CH0271428317[c]	13/02/2015	19/02/2015	1,250	USD

(continued)

Table 1 (continued)

Name	ISIN	Announc.	Issue	Size	Curr.
UBS 5 1/8 05/15/24	CH0244100266	08/05/2014	15/05/2014	2,500	USD
UBS 4 3/4 02/12/26	CH0236733827	06/02/2014	13/02/2014	2,000	EUR
UBS 4 3/4 05/22/23	CH0214139930	15/05/2013	22/05/2013	1,500	USD
UCGIM 6 3/4 09/29/49	XS1107890847	03/09/2014	10/09/2014	1,000	EUR

[a]Incl. XS1055037920
[b]Incl. US06738EAB11 and XS1068574828
[c]Incl. CH0271428333 and CH0271428309
Source Bloomberg/own calculations

issued different CoCos on the same day. Since it is not possible to distinguish their influence from each other, these new CoCos are assumed to have one general impact on all the outstanding CoCos of the same issuing company.

3.2 CoCo Index Comparison

The first analysis is based on indices as a benchmark to observe a certain impact. It basically compares the returns of the outstanding CoCo bonds after an announcement of a new issue with some overall CoCo indices. More precisely, we compare the performance with the CS Contingent Convertible Euro Total Return index and the BofA Merrill Lynch Contingent Capital Index (whenever the data is available). The methods are explained for one particular new CoCo issuance, namely the USF22797YK86 CoCo of Crédit Agricole. In the end, the overall results and conclusions are shown.

3.2.1 Method

In a first step, we analyse the impact of each new CoCo separately on all the outstanding CoCos of the same issuer. The simple returns are derived for the outstanding CoCos during the period between the announcement date and the issue date of the new CoCo. In a second step, we accumulate these simple returns and obtain the returns between announcement and issue date. As an example, the first steps are shown for two outstanding CoCos of Crédit Agricole in Fig. 1.

On each day, we calculate the (equally weighted) average of the cumulative simple returns of all outstanding CoCos. In a last step, we take the difference between these averages and the cumulative returns of the CoCo index on each day between the announcement date and issue date of the new CoCo.

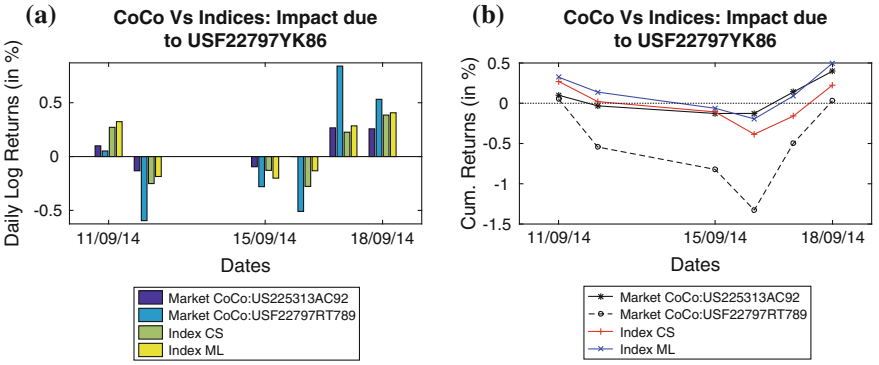

Fig. 1 Impact of USF22797YK86. **a** Daily returns. **b** Cumulative returns

3.2.2 Results

In Table 2, the difference in cumulative returns over the observation period, meaning the period between announcement and issuance, is shown. For some observation periods, the Merrill Lynch index did not yet exist. When the CoCo does show a significant change compared to the global index, we can assume that this change is due to the new CoCo issuance. The averaged difference over all new CoCos analysed is shown in Fig. 2. These averaged differences in cumulative returns are shown for one day until five days after the announcement of the new CoCo and also over the full period as was given in Table 2. As a conclusion, we see that on average the outstanding CoCos get a negative impact of around 25 bps on their return between announcement and issuance due to a new CoCo.

Multiple CoCo indices can be used for this analysis but CoCo indices are relatively new on the market. As such we are obliged to restrict our analysis to indices already available during the period of each analysis. Remark also that we need to handle these indices with care, in the sense that the indices are applied to give a global market view on the CoCos. A point of criticism to this approach can be that the indices are not that representative for the true market. There is also high concentration on some issuers in the indices, e.g. for the ML index the top 5 issuers almost make 50 % of the index (as of December 2014).

Furthermore, this comparison with the general market performance does not take into account idiosyncratic movements but compares with the general market trend. The issuing company is nevertheless exposed to individual dynamics. Its stock price, its credit worthiness, etc. can change differently from their competitors. This can be especially the case around capital raising announcements since then financial details of the company are published and discussed at, for example road-shows around the new issuance. Therefore, we move on to a second methodology taking into account idiosyncratic movements.

Table 2 Averaged difference in cumulative returns (in %) between the outstanding CoCos and the Credit Suisse CoCo index (left column) and Merrill Lynch CoCo index (right column) over the observation period of the new CoCo

Issuer	ISIN	CS index	ML index
ACAFP	USF22797YK86	0.06	−0.21
	XS1055037177	−0.06	0.07
	USF22797RT78	−0.21	−0.47
BACR	XS1068561098	−0.25	−0.12
	XS1002801758	0.08	/
	US06738EAA38	0.56	/
BBVA	XS1190663952	−0.51	−1.18
CS	XS1076957700	0.31	0.60
	XS0989394589	−0.28	/
	XS0972523947	−0.50	/
	CH0221803791	1.09	/
DANBNK	XS1190987427	−1.32	−2.27
DB	US251525AN16	−0.13	−0.29
RABOBK	XS1171914515	−0.36	−0.41
SANTAN	XS1107291541	−1.38	−1.19
	XS1066553329	−0.49	−0.06
SOCGEN	USF8586CXG25	0.23	0.17
	XS0867620725	−0.27	−0.02
	USF8586CRW49	1.24	/
UBS	CH0271428317	−0.46	−1.01
	CH0244100266	−0.12	0.42
	CH0236733827	0.37	0.31
	CH0214139930	−1.15	/
UCGIM	XS1107890847	−1.75	−1.48
Mean		**−0.22**	**−0.42**
Std Dev		**0.71**	**0.77**

Source Bloomberg/own calculations

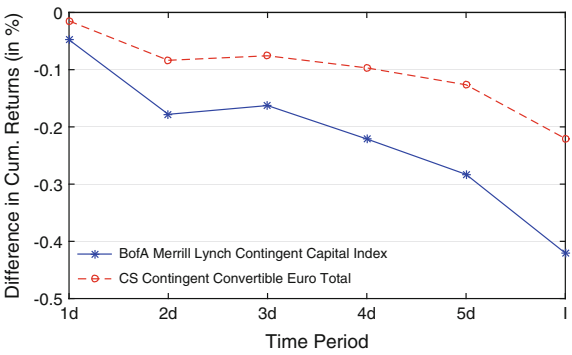

Fig. 2 Averaged difference in cumulative returns between the outstanding CoCos and the Credit Suisse and Merrill Lynch CoCo index

3.3 Model-Based Performance

As experienced by all CoCo investors, the difficulty in these financial products lies in their different characteristics which are hard to compare like the trigger type, conversion type, maturity, coupon cancellation, and so on. However, the implied barrier methodology can be used as a tool to compare CoCos with different characteristics. In this second approach, we will use the implied barrier to derive theoretical values for the outstanding CoCos under the assumption of no impact by the new CoCo issuance and compare them with the actual market values.

3.3.1 Method

The implied barrier can be interpreted as the stock price level (assumed by the market) that is hit (for the first time) when the CoCo gets converted or written down. If nothing changes, the market will keep the same idea about the implied barrier level and hence result in a constant implied barrier over time. In other words, when there is no impact due to this new CoCo, no change could theoretically be observed in

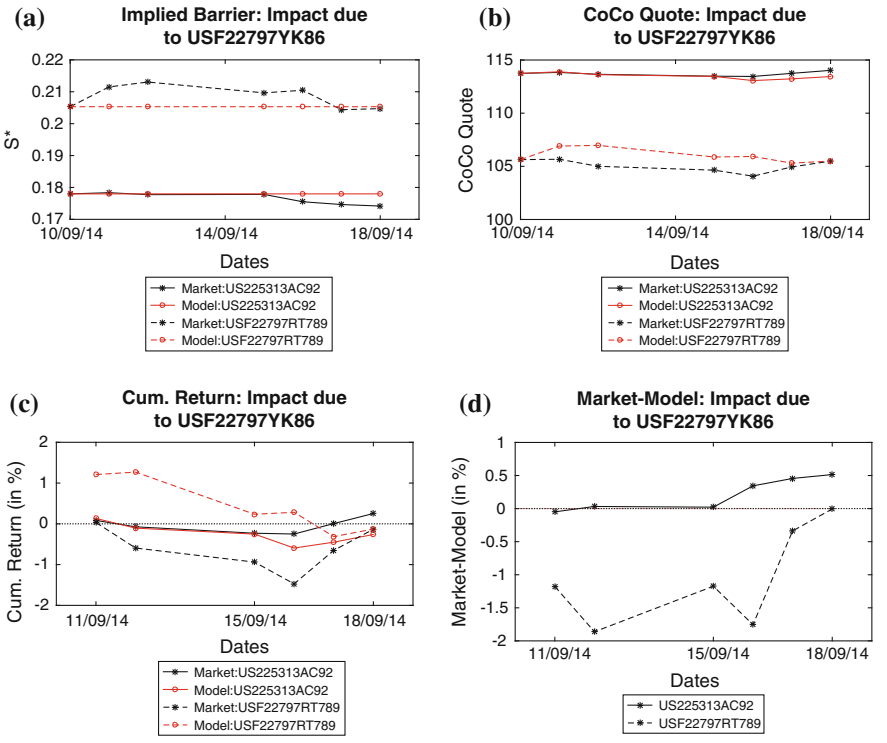

Fig. 3 Impact of USF22797YK86. **a** Implied barriers. **b** CoCo quotes. **c** Returns compared with announcement. **d** Difference in returns

the implied barrier. As such we can see in the levels of the implied barrier if there is an impact due to the announced new CoCo. This leads us easily to the second approach of our impact analysis. As an example, we show the implied barriers of the two outstanding CoCos of ACAFP from the previous section in Fig. 3a.

As from the previous section, the implied barriers can be translated into CoCo quotes. The theoretical CoCo price does not take any information of a new CoCo issuance into account by assuming a constant implied barrier. These values can be used as our reference. Any change in the market compared with this reference, is then due to the impact of the announcement of a new CoCo issuance. As such we can calculate the theoretical CoCo prices from a constant implied barrier and compare them with the market values. The results of our CoCo examples are shown in Fig. 3b. As a last step, we define cheapness as the difference between the market CoCo return and the theoretical CoCo return until the announcement date. In Fig. 3, the cheapness of the two outstanding CoCos of Crédit Agricole is shown.

Table 3 Averaged difference between the market and model CoCo cumulative returns (in %)

Issuer	ISIN	Cheapness
ACAFP	USF22797YK86	0.14
	XS1055037177	−1.45
	USF22797RT78	−0.19
BACR	XS1068561098	0.17
	XS1002801758	−0.05
	US06738EAA38	−0.38
BBVA	XS1190663952	−1.91
CS	XS1076957700	0.12
	XS0989394589	0.53
	XS0972523947	−2.63
	CH0221803791	1.44
DANBNK	XS1190987427	−3.25
DB	US251525AN16	0.65
RABOBK	XS1171914515	0.31
SANTAN	XS1107291541	−2.24
	XS1066553329	0.07
SOCGEN	USF8586CXG25	1.26
	XS0867620725	−2.31
	USF8586CRW49	1.67
UBS	CH0271428317	0.71
	CH0244100266	−1.23
	CH0236733827	1.03
	CH0214139930	−0.89
UCGIM	XS1107890847	−1.60
Mean		**−0.42**
Std Dev		**1.37**

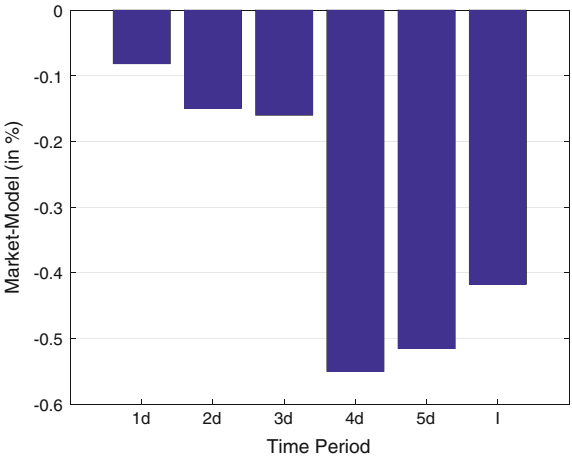

Fig. 4 Overall difference between the market and model CoCo cumulative returns

3.3.2 Results

An overall view is derived for the cheapness by averaging the differences in theoretical and market CoCo prices for each outstanding CoCo during the observation period of the new CoCo. We averaged the differences of all the CoCos on one day until five days after the announcement and also on the issue date of the new CoCo (Table 3).

Clearly, from Fig. 4, on average the cheapness on each day of our observation period is negative, meaning that market price is below the theoretical price assuming no impact. As such we conclude also from this approach that there is a negative impact of about 42 bps on average on the outstanding CoCos when a new CoCo issuance is announced.

4 Impact After Issue Date

At this point, we investigated the impact of a new CoCo issuance between the announcement and issue date. In this section, we show the results for a longer observation period. More concrete, both analyses are extended to 10 trading days after the issue date.

From our first analysis, where we compare the outstanding CoCos with the CoCo indices, a downward trending impact is observed in Fig. 5a. The second analysis which compares the market and model prices of the outstanding CoCos is shown in Fig. 5b. In both analyses, the negative impact gets more significant after the issue dates. Hence until 10 trading days after the issue date, there is still a negative impact observable.

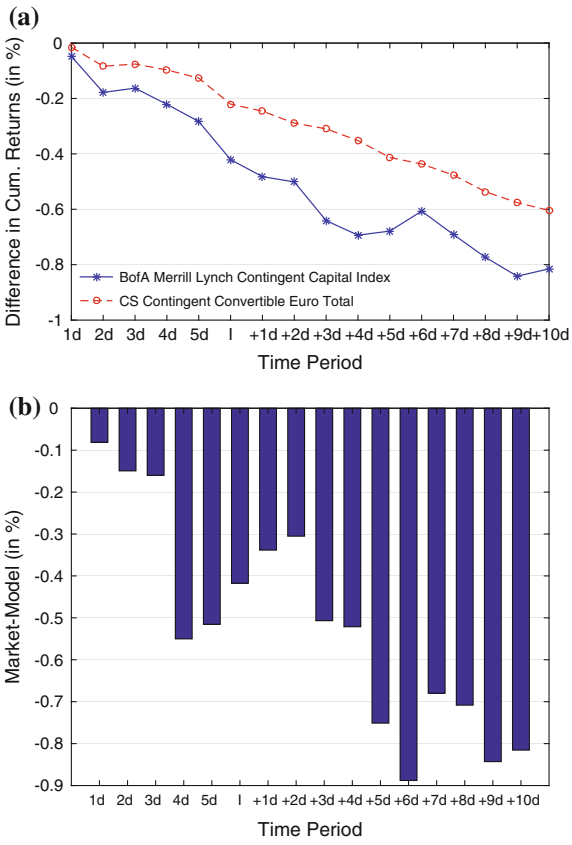

Fig. 5 Impact from the announcement date until 10 days after the issue date. **a** Method 1. **b** Method 2

5 Conclusion

The price performance of outstanding CoCos was investigated after a new CoCo issue is announced by the same issuer. Based on two approaches, we estimated the price impact on the outstanding CoCos. The first method compared returns of the outstanding CoCos with some overall CoCo indices. As a conclusion, we found that the return of the outstanding CoCos, during the period between announcement and issuance, was slightly lower than the returns of the CoCo indices. There was an underperformance of about 22 bps compared with the Credit Suisse index and about 42 bps with the Merrill Lynch index (although with relative high standard deviations). Since this first study did not take idiosyncratic movements into account, we used also a second method based on the equity derivatives model for CoCos. In this method we compared the actual market performance of the outstanding CoCo bonds with a theoretical model performance taking into account idiosyncratic effects,

like movements in the underlying stock, credit default spreads and volatilities. This second approach also concludes that the averaged market returns of the outstanding CoCos were about 42 bps lower than one would expect in case of no influence.

In total, we investigate 24 cases of new CoCo bond issues. The main conclusion of the investigation is that there is a moderated negative effect on outstanding CoCo bonds. This is confirmed by both methodologies and the impact is an underperformance of about 20–40 bps on average in between the announcement date and the issue date. During the period of 10 trading days after the issue date, an extra decrease of 40 bps was observed.

Acknowledgements The authors would like to thank Robert Van Kleeck and Michael Hünseler for useful discussions on the topic.

The KPMG Center of Excellence in Risk Management is acknowledged for organizing the conference "Challenges in Derivatives Markets - Fixed Income Modeling, Valuation Adjustments, Risk Management, and Regulation".

References

1. Basel Committee on Banking Supervision: Basel III: A global regulatory framework for more resilient banks and banking systems. Bank for International Settlements, Dec. 2010
2. European Commision: Possible Further Changes to the Capital Requirements Directive. Working Document. Commission Services Staff (2010)
3. Akigbe, A., Easterwood, J.C., Pettit, R.R.: Wealth effects of corporate debt issues: the impact of issuer motivations. Financ. Manag. **26**(1), 32–47 (1997)
4. Pennacchi, G.: A Structural Model of Contingent Bank Capital (2010)
5. Pennacchi, G., Vermaelen, T., Wolff, C.: Contingent capital: the case for COERCs. INSEAD. Working paper (2010)
6. De Spiegeleer, J., Schoutens, W.: Contingent convertible contingent convertibles: a derivative approach. J. Deriv. (2012)
7. De Spiegeleer, J., Schoutens, W., Van Hulle, C.: The Handbook of Hybrid Securities: Convertible Bonds, CoCo Bonds and Bail-in. Wiley, New York (2014), Chap. 3
8. Corcuera, J.M., De Spiegeleer, J., Ferreiro-Castilla, A., Kyprianou, A.E., Madan, D.B., Schoutens, W.: Pricing of contingent convertibles under smile conform models. J. Credit Risk **9**(3), 121–140 (2013)
9. De Spiegeleer, J., Forys, M., Marquet, I., Schoutens, W.: The Impact of Skew on the Pricing of CoCo Bonds. Available on SSRN: http://ssrn.com/abstract=2418150 (2014)
10. De Spiegeleer, J., Marquet, I., Schoutens, W.: CoCo Bonds and Implied CET1 Volatility. Available on SSRN: http://ssrn.com/abstract=2575558 (2015)

11. De Spiegeleer, J., Schoutens, W.: Steering a bank around a death spiral: multiple Trigger CoCos. Wilmott Mag. **2012**(59), 62–69 (2012)
12. De Spiegeleer, J., Schoutens, W.: CoCo Bonds with Extension Risk (2014)
13. Corcuera, J.M., De Spiegeleer, J., Fajardo, J., Jönsson, H., Schoutens, W., Valdivia, A.: Close form pricing formulas for CoCa CoCos. J. Bank. Financ. (2014)
14. De Spiegeleer, J., Schoutens, W.: Multiple Trigger CoCos: contingent debt without death spiral risk. Financ. Mark. Inst. Instr. **22**(2) (2013)
15. Cheridito, P., Zhikai, X.: A reduced form CoCo model with deterministic conversion intensity. Available on SSRN: http://ssrn.com/abstract=2254403 (2013)
16. Madan, D.B., Schoutens, W.: Conic coconuts: the pricing of contingent capital notes using conic finance. Math. Financ. Econ. **4**(2), 87–106 (2011)

The Impact of Cointegration on Commodity Spread Options

Walter Farkas, Elise Gourier, Robert Huitema and Ciprian Necula

Abstract In this work we explore the implications of cointegration in a system of commodity prices on the premiums of options written on various spreads on the futures prices of these commodities. We employ a parsimonious, yet comprehensive model for cointegration in a system of commodity prices. The model has an exponential affine structure and is flexible enough to allow for an arbitrary number of cointegration relationships. We conduct an extensive simulation study on pricing spread options. We argue that cointegration creates an upward sloping term structure of correlation, that in turn lowers the volatility of spreads and consequently the price of options on them.

Keywords Cointegration · Futures prices · Commodities · Spread options · Simulation

1 Introduction

A distinctive feature of commodity markets is the existence of long-run equilibrium relationships that exist between the levels of various commodity prices, such as the one between the price of crude oil and the price of heating oil. These long-run

W. Farkas (✉) · R. Huitema · C. Necula
Department of Banking and Finance, University of Zurich,
Plattenstrasse 14, 8032 Zurich, Switzerland
e-mail: walter.farkas@bf.uzh.ch

W. Farkas
Department of Mathematics, ETH Zurich, Rämistrasse 101, 8092 Zurich, Switzerland

W. Farkas
Swiss Finance Institute, Zurich, Switzerland

E. Gourier
School of Economics and Finance, Queen Mary, University of London, London, UK

C. Necula
Department of Money and Banking, Bucharest University of Economic Studies,
Bucharest, Romania

K. Glau et al. (eds.), *Innovations in Derivatives Markets*, Springer Proceedings
in Mathematics & Statistics 165, DOI 10.1007/978-3-319-33446-2_20

equilibrium relations can be captured in economic models by so-called *cointegration* relations.

In this work we employ the continuous time model of cointegrated commodity prices developed by the authors in Farkas et al. [6] in order to conduct a simulation study for assessing the impact of cointegration on spread options. In our model, commodity prices are non-stationary and several cointegration relations are allowed amongst them, capturing long-run equilibrium relationships. Cointegration (Engle and Granger [5]) is the property of two or more non-stationary time series of having at least one linear combination that is stationary.

There is a vast literature on modeling the price of a single commodity as a non-stationary process (see Back and Prokopczuk [1] for a comprehensive recent review). For example, Schwartz and Smith [13] assume the log price of a commodity to be the sum of two latent factors: the long-term equilibrium level, modeled as a geometric Brownian motion, and a short-term deviation from the equilibrium, modeled as a zero mean Ornstein–Uhlenbeck (OU) process. More recently, Paschke and Prokopczuk [11] propose to model these deviations as a more general CARMA process and Cortazar and Naranjo [3] generalize the Schwartz and Smith [13] model in a multi-factor framework.

However, the literature on modeling a system of commodity prices is still quite scarce. Two fairly recent models are proposed in Cortazar et al. [4] and Paschke and Prokopczuk [10], both of which account for cointegration by incorporating common and commodity-specific factors into their modeling framework. Amongst the common factors, only one is assumed non-stationary. Although they explicitly take into account cointegration between prices, the cointegrated systems generated by these two models are not covering the whole range of possible number of cointegration relations, but allow for none or for exactly $n-1$ relations to exist between the n prices. In Farkas et al. [6] we propose an easy-to-use, yet comprehensive, model for a system of cointegrated commodity prices that retains the exponential affine structure of previous approaches and allows, in the same time, for an arbitrary number of cointegration relationships.

The rest of the work is organized as follows. In Sect. 2 we briefly describe the model proposed in Farkas et al. [6] and point out some qualitative aspects regarding the dynamics of the system. Section 3 is devoted to an extensive simulation study focused on computing spread options prices and on assessing the impact of cointegration on pricing spread options. Section 4 is reserved for concluding remarks.

2 Outline of the Model

Before proceeding to the simulation study, in this section we present for the sake of completeness, a short description of the model developed in Farkas et al. [6].

Consider n commodities with spot prices $\mathbf{S}(t) = (S_1(t), \ldots, S_n(t))^\top$.

First it is assumed that the spot log-prices $\mathbf{X}(t) = \log \mathbf{S}(t)$ can be decomposed into three components:

$$\mathbf{X}(t) = \mathbf{Y}(t) + \boldsymbol{\varepsilon}(t) + \boldsymbol{\phi}(t), \tag{1}$$

where $\mathbf{Y}$(t) signifies the long-run levels, $\boldsymbol{\varepsilon}(t)$ is an n-dimensional stationary process capturing short-term deviations, and $\boldsymbol{\phi}(t) = \boldsymbol{\chi}_1 \cos(2\pi t) + \boldsymbol{\chi}_2 \sin(2\pi t)$ controls for seasonal effects with $\boldsymbol{\chi}_1$ and $\boldsymbol{\chi}_2$ being n-dimensional vectors of constants.

The notion of cointegration (Engle and Granger [5], Johansen [7], Phillips [12]) refers to the property of two or more non-stationary time series of having a linear combination that is stationary. For example, if $X_1(t)$ and $X_2(t)$ are two non-stationary processes, one says that they are cointegrated if there is a linear combination of them, $X_1(t) - \alpha X_2(t)$, that is stationary for some positive real α. Intuitively, cointegration occurs when two or more non-stationary variables are linked in a long-run equilibrium relationship from which they might depart only temporarily.

Regarding cointegration in the model, we stress that n cointegration relationships are implicitly assumed by (1): the n seasonally adjusted spot log-prices $\mathbf{X}(t) - \boldsymbol{\phi}(t)$ are cointegrated with their corresponding long-run levels, $\mathbf{Y}(t)$, since the linear combination $\mathbf{X}(t) - \boldsymbol{\phi}(t) - \mathbf{Y}(t)$ is stationary.

Secondly, cointegration is allowed to exist between the variables in $\mathbf{Y}(t)$ as well. We denote the number of cointegration relationships between them by h, where $h \geq 0$ and $h < n$. The corresponding cointegration matrix is symbolized by Θ, an $n \times n$ matrix with the last $n - h$ rows equal to zero vectors. Each of the h non-zero rows of Θ encodes a stationary (i.e., cointegrating) combination of the variables in $\mathbf{Y}(t)$, normalized such that $\Theta_{ii} = 1,\ i \leq h$. The total $n + h$ cointegration relationships between the variables in the vector $\mathbf{Z}(t) := (\mathbf{X}(t) - \boldsymbol{\phi}(t), \mathbf{Y}(t))^\top$ can be characterized by the $(2n \times 2n)$-matrix $\begin{bmatrix} \mathbf{I}_n & -\mathbf{I}_n \\ \mathbf{O}_n & \Theta \end{bmatrix}$ where $\mathbf{O}_n$ denotes the zero-matrix with dimension $n \times n$.

The dynamics of $\mathbf{X}(t)$ and $\mathbf{Y}(t)$ under the real-world probability measure is assumed to be given by:

$$d\begin{bmatrix} \mathbf{X}(t) - \phi(t) \\ \mathbf{Y}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{0}_n \\ \mu_y \end{bmatrix} dt + \begin{bmatrix} -K_x & \mathbf{O}_n \\ \mathbf{O}_n & -K_y \end{bmatrix} \begin{bmatrix} \mathbf{I}_n & -\mathbf{I}_n \\ \mathbf{O}_n & \Theta \end{bmatrix} \begin{bmatrix} \mathbf{X}(t) - \phi(t) \\ \mathbf{Y}(t) \end{bmatrix} dt + \begin{bmatrix} \Sigma_x^{\frac{1}{2}} & \mathbf{O}_n \\ \Sigma_{xy}^{\frac{1}{2}} & \Sigma_y^{\frac{1}{2}} \end{bmatrix} d\begin{bmatrix} \mathbf{W}_x(t) \\ \mathbf{W}_y(t) \end{bmatrix}, \tag{2}$$

where $\mathbf{0}_n$ is an n-dimensional vector of zeros, and $\mathbf{W} := (\mathbf{W}_x, \mathbf{W}_y)^\top$ is a $2n$-dimensional standard Brownian motion. Furthermore, the matrix $\begin{bmatrix} -K_x & \mathbf{O}_n \\ \mathbf{O}_n & -K_y \end{bmatrix}$ measures the speed by which $\mathbf{Z}(t)$ reverts to its long-run (cointegration) equilibrium level. More specifically, K_x quantifies the speed of mean reversion of the elements in $\mathbf{X}$ around the long term levels in $\mathbf{Y}$. The matrix K_y is an $n \times n$ matrix with the last $n - h$ columns equal to zero vector, such that $K_y\Theta$ is an $n \times n$ matrix of rank h. Each of the h non-zero columns in K_y quantifies the speed of adjustment of each element in $\mathbf{Y}$ to the corresponding cointegration relation. The dynamics given by Eq. (2) is "error-correcting" in that a deviation from a given cointegration relation induces an appropriate change in variables in the direction of correcting the deviation.

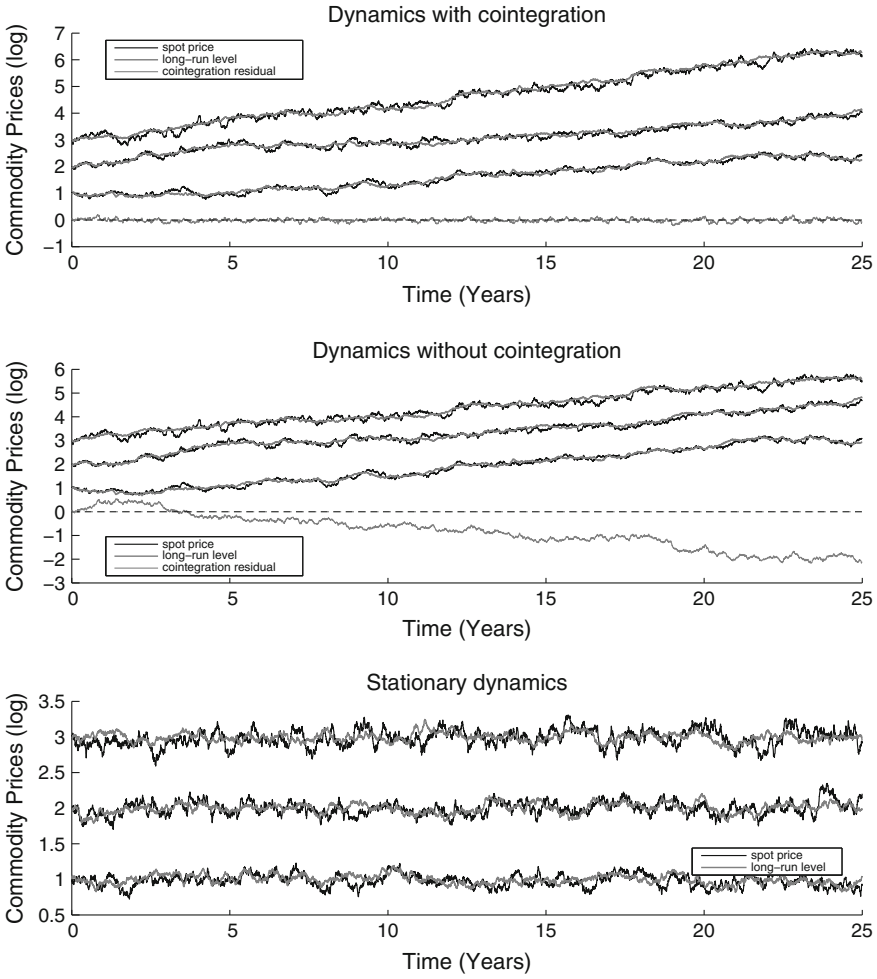

Fig. 1 Simulated price paths for various choices of the Θ matrix. *Top panel* prices are non-stationary and there is one cointegration relation. *Middle panel* prices are non-stationary and there is no cointegration. *Bottom panel* prices are stationary

In order to assess qualitatively the role of the cointegration matrix Θ on the properties of the dynamics of the system, Fig. 1 depicts the results of a simulation of a system of three variables for various choices of the Θ matrix.

In the top panel of Fig. 1 we assume that there is a cointegration relation and the first line of the Θ matrix is $\begin{bmatrix} 1 & 1 & -1 \end{bmatrix}$ and, therefore, the residual of the cointegration relation, $Y_1(t) + Y_2(t) - Y_3(t)$, is stationary. On the other hand, in the middle panel, depicts the case when Θ is the null matrix and, therefore, the prices are non-stationary and not cointegrated. For example, the residual of the cointegration relation from the previous case, $Y_1(t) + Y_2(t) - Y_3(t)$, is no longer stationary. In fact, there is no

stationary linear combination of the long run levels in this case. Moreover, as depicted in the bottom panel, the model also allows for stationary prices, if the Θ matrix is of full rank.

The characteristic functions of $\mathbf{X}$ and $\mathbf{Y}$ can be readily computed analytically given they are normally distributed since the dynamics of $\mathbf{Z}(t) = (\mathbf{X}(t) - \boldsymbol{\phi}(t), \mathbf{Y}(t))^\top$ is in fact given by a multivariate Ornstein–Uhlenbeck (OU) process:

$$d\mathbf{Z}(t) = [\boldsymbol{\mu} - K\mathbf{Z}(t)]\,dt + \Sigma^{\frac{1}{2}} d\mathbf{W}(t), \tag{3}$$

with $\boldsymbol{\mu} := \begin{bmatrix} \mathbf{0}_n \\ \boldsymbol{\mu}_y \end{bmatrix}$, $K := \begin{bmatrix} K_x & -K_x \\ \mathbf{O}_n & K_y\Theta \end{bmatrix}$, $\Sigma^{\frac{1}{2}} := \begin{bmatrix} \Sigma_x^{\frac{1}{2}} & \mathbf{O}_n \\ \Sigma_{xy}^{\frac{1}{2}} & \Sigma_y^{\frac{1}{2}} \end{bmatrix}$, $\mathbf{W}(t) := \begin{bmatrix} \mathbf{W}_x(t) \\ \mathbf{W}_y(t) \end{bmatrix}$.

At the same time, the vector of spot prices $\mathbf{S}(T)$ can be written as an exponential function of $\mathbf{X}(t)$ and $\mathbf{Y}(t)$:

$$\begin{aligned} \mathbf{S}(T) = \exp\Bigg\{ & e^{-K_x(T-t)}\mathbf{X}(t) + \psi(T-t)\mathbf{Y}(t) + \left[\phi(T) - e^{-K_x(T-t)}\phi(t)\right] \\ & + \left[\int_t^T \psi(T-u)du\right]\mu_y + \int_t^T \left[e^{-K_x(T-u)}\Sigma_x^{\frac{1}{2}} + \psi(T-u)\Sigma_{xy}^{\frac{1}{2}}\right] d\mathbf{W}_x(u) \\ & + \int_t^T \psi(T-u)\Sigma_y^{\frac{1}{2}} d\mathbf{W}_y(u)\Bigg\}. \end{aligned} \tag{4}$$

where

$$\psi(\tau) := K_x \left[\int_0^\tau e^{-K_x(\tau-u)} e^{-K_y\Theta u} du\right].$$

Given the affine structure of the model, futures prices can also be obtained in closed form. Under the simplifying assumption of constant market prices of risk, one has that $d\begin{bmatrix} \mathbf{W}_x^*(t) \\ \mathbf{W}_y^*(t) \end{bmatrix} = d\begin{bmatrix} \mathbf{W}_x(t) \\ \mathbf{W}_y(t) \end{bmatrix} + \begin{bmatrix} \boldsymbol{\lambda}_x \\ \boldsymbol{\lambda}_y \end{bmatrix} dt$ where $\mathbf{W}_x^*(t)$ and $\mathbf{W}_y^*(t)$ are standard Brownian motions under the risk-neutral measure, and $\boldsymbol{\lambda}_x$, $\boldsymbol{\lambda}_y$ are the market prices of $\mathbf{W}_x(t)$ and $\mathbf{W}_y(t)$ risks, respectively.

Under these circumstances it can be shown that at time t the vector of futures prices for the contracts with maturity T is given by

$$\mathbf{F}(t,T) = \exp\left\{\alpha(t,T) + \beta(T-t)\mathbf{X}(t) + \psi(T-t)\mathbf{Y}(t)\right\}, \tag{5}$$

with $\beta(\tau) := e^{-K_x\tau}$ and with $\alpha(t,T)$ defined by

$$\begin{aligned} \alpha(t,t+\tau) := & \left[\phi(t+\tau) - e^{-K_x\tau}\phi(t)\right] - \left(\mathbf{I}_n - e^{-K_x\tau}\right) K_x^{-1}\boldsymbol{\mu}_x^* + \left(\int_0^\tau \psi(\tau-u)du\right)\boldsymbol{\mu}_y^* \\ & + \operatorname{diag}\left\{\frac{1}{2}\left[\mathbf{I}_n\ \mathbf{O}_n\right]\left[e^{-K\tau}\left(\int_0^\tau e^{Ku}\Sigma e^{Ku}du\right)e^{-K\tau}\right]\begin{bmatrix}\mathbf{I}_n \\ \mathbf{O}_n\end{bmatrix}\right\}, \end{aligned} \tag{6}$$

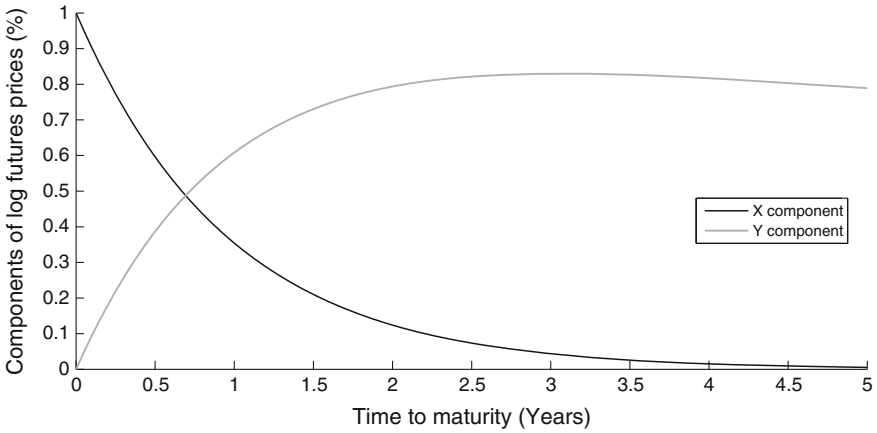

Fig. 2 The relative contribution of various components to log futures prices

where $\mathrm{diag}(A)$ returns the vector with diagonal elements of A, and where

$$\boldsymbol{\mu}^* = \begin{bmatrix} \boldsymbol{\mu}_x^* \\ \boldsymbol{\mu}_y^* \end{bmatrix} = \boldsymbol{\mu} + \Sigma^{\frac{1}{2}} \begin{bmatrix} \boldsymbol{\lambda}_x \\ \boldsymbol{\lambda}_y \end{bmatrix}.$$

To better assess qualitatively the impact of the two factors, $\mathbf{X}(t)$ and $\mathbf{Y}(t)$, on the term structure of futures prices, we depict in Fig. 2 the relative contribution of the corresponding two terms in Eq. (5) to the logarithm of the futures prices on one of the commodities in a cointegrated system.

The contribution of the $\mathbf{X}(t)$ component decreases exponentially as a function of time to maturity. On the other hand, the $\mathbf{Y}(t)$ component contributes significantly for higher maturities. Therefore, the two factors capture the short-end and, respectively, the long-end of the term-structure of futures prices.

By Itô's lemma, the risk-neutral dynamics of $\mathbf{F}(t, T)$ is given by

$$\frac{d\mathbf{F}(t,T)}{\mathbf{F}(t,T)} = \left[e^{-K_x(T-t)} \Sigma_x^{\frac{1}{2}} + \psi(T-t) \Sigma_{xy}^{\frac{1}{2}} \right] d\mathbf{W}_x^*(t) + \psi(T-t) \Sigma_y^{\frac{1}{2}} d\mathbf{W}_y^*(t), \tag{7}$$

and it follows immediately that the variance–covariance matrix of returns on futures prices is given by:

$$\begin{aligned} \Xi(\tau) = {} & e^{-K_x\tau} \Sigma_x e^{-K_x^\top \tau} + \psi(\tau) \Sigma_{xy}^{\frac{1}{2}} (\Sigma_x^{\frac{1}{2}})^\top e^{-K_x^\top \tau} + e^{-K_x\tau} \Sigma_x^{\frac{1}{2}} (\Sigma_{xy}^{\frac{1}{2}})^\top \psi^\top(\tau) \\ & + \psi(\tau) \Sigma_{xy} \psi^\top(\tau) + \psi(\tau) \Sigma_y \psi^\top(\tau) \end{aligned} \tag{8}$$

where $\tau = T - t$.

Since the term structure of correlation of futures prices returns plays an important role in the results of the simulations performed in the following section, it is worthwhile to point out some qualitative results about this term structure.

First, Eq. (8) shows that unless $K_x = \mathbf{O}_n$, the variance–covariance matrix $\Xi(\tau)$ depends on τ.

Second, let us consider the case that there is no instantaneous correlation between the shocks driving the dynamics, meaning that Σ_x and Σ_y are diagonal matrices and Σ_{xy} is the null matrix. Moreover, let us assume that K_x is diagonal, meaning that the spot price of a commodity reacts only to its deviation from the long run level and not to deviations of the other commodities. It follows that the first term in Eq. (8) is a diagonal matrix and the next three terms are null matrices. If, in addition, there is no cointegration in the system, meaning that Θ is the null matrix, then the last term in Eq. (8) is a diagonal matrix since $\psi(\tau)$ is also a diagonal matrix. So, in this case, the variance–covariance matrix $\Xi(\tau)$ is diagonal and, therefore, there is no correlation at any maturity. However, if there is at least one cointegration relation in the system, then the last term in Eq. (8) is no longer a diagonal matrix since $\psi(\tau)$ is not diagonal. Therefore, cointegration induces correlation at various maturities although it was assumed there is no instantaneous correlation between the Brownian motions in the model.

3 Spread Option Prices

In this section, we focus on futures prices and prices of European-style options written on the *spread* between two or more commodities, such as the difference between the price of electric power and the cost of the natural gas needed to produce it, or the price difference between crude oil and a basket of various refined products, known as the crack spread. The crack spread is in fact related to the profit margin that an oil refiner realizes when "cracking" crude oil while simultaneously selling the refined products in the wholesale market. The oil refiner can hedge the risk of losing profits by buying an appropriate number of futures contract on the crack spread or, alternatively, by buying call options of the crack spread. Since spread options have become regularly and widely used instruments in financial markets for hedging purposes, there is a growing need for a better understanding of the effects of cointegration on their prices.

There is extensive literature on approximation methods for spread and basket options on two (e.g. Kirk [8]) or more than two commodities, with recent contributions from Li et al. [9] and Caldana and Fusai [2]. However, mostly for simplicity, we relay in this chapter on the Monte-Carlo simulation method for pricing spread options written on two or more than two commodities.

From Eq. (7), it follows that $\mathbf{F}(t, T)$ (conditional on information available up to time $s \leq t \leq T$) is distributed as follows:

$$\mathbf{F}(t,T) \sim \log \mathcal{N}\left(\log \mathbf{F}(s,T) - \frac{1}{2}\int_s^t \operatorname{diag}(\Xi(T-u))du, \int_s^t \Xi(T-u)du\right), \tag{9}$$

where $\operatorname{diag}(X)$ denotes the vector containing the diagonal elements of the matrix X. Note that $\mathbf{F}(s,T)$ can be either computed from (5) or observed from data.

The fact that the distribution function of $\mathbf{F}(t,T)$ is known in an easy-to-use and analytic form is one of the key features of the model we employ. It allows us to simulate futures price curves at any time t in the future based on today's curves (time s) almost effortlessly. Hence, the price of a call option on the time-T value of a certain spread can be simply obtained by carrying out the following steps:

(i) compute or observe today's futures price curves $\mathbf{F}(s,T)$;
(ii) compute M realizations $\mathbf{F}^{(m)}$ $(m = 1,\ldots,M)$ of $\mathbf{F}(T,T)$ by sampling from (9) as follows:

$$\mathbf{F}^{(m)} = \mathbf{F}(s,T)\exp\left\{\boldsymbol{\varepsilon}^{(m)}\right\},$$

where $\boldsymbol{\varepsilon}^{(m)}$ is generated from a multivariate normal distribution with mean $-\frac{1}{2}\int_s^T \operatorname{diag}(\Xi(T-u))du$ and variance–covariance matrix $\int_s^T \Xi(T-u)du$[1];
(iii) compute the Monte-Carlo estimate of a call with strike k on the spread

$$\sum_{n=1}^{N}\omega_n S_n(T) \qquad \left(=\sum_{n=1}^{N}\omega_n F_n(T,T)\right),$$

with $\omega_n, n = 1,\ldots,N$ the weights of each component in the spread, as follows:

$$\frac{1}{M}\sum_{m=1}^{M}\max\left\{\left[\sum_{n=1}^{N}\omega_n F_n^{(m)}\right] - k, 0\right\}. \tag{10}$$

For the sake of clarity we have set the risk-free rate curve equal to zero. We note that the random variables $\boldsymbol{\varepsilon}^{(m)}$ can be simply re-used for pricing spread options with different maturity dates.

In the following we consider a system of three commodities[2] characterized by one cointegration relation with $\Theta = \begin{bmatrix} 1 & -0.4 & -0.6 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. The rest of the parameters describing the dynamics are $K_x = \begin{bmatrix} 1.5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0.5 \end{bmatrix}$, $\Sigma_x = \begin{bmatrix} 0.0625 & 0.0562 & 0.0437 \\ 0.0562 & 0.0900 & 0.0262 \\ 0.0437 & 0.0262 & 0.1225 \end{bmatrix}$, $\mu_y =$

[1] Here the technique of antithetic variables is used to reduce the number of random samples needed for a given level of accuracy.

[2] The structure of the parameters is chosen, in a parsimonious manner, taking into consideration the key facts of the empirical study conducted in Farkas et al. [6], where the results provide compelling evidence of cointegration between various commodities.

$\begin{bmatrix} 0.025 \\ 0.025 \\ 0.025 \end{bmatrix}$, $K_y = \begin{bmatrix} 1.5 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $\Sigma_y = \begin{bmatrix} 0.0225 & 0 & 0 \\ 0 & 0.0225 & 0 \\ 0 & 0 & 0.0225 \end{bmatrix}$, $\Sigma_{xy} = \mathbf{O}_3$. Since K_x is diagonal, each spot price is error-corrected only with respect to deviations from its own long-run level. Moreover, given the specific form of the K_y matrix, deviations from the cointegration relationships between the long-run levels influence only the dynamics of the first spot price. In this respect, the second and third commodities are "exogenous" in that their dynamics is not influenced by the variables characterizing the other commodities. Regarding instantaneous dependence, the shocks driving the dynamics of the long-run factors are not correlated, whereas we imposed positive correlations between the shocks driving the dynamics of the $\mathbf{X}(t)$. More specifically, the instantaneous variance–covariance matrix Σ_y for long-run shocks corresponds to an annual volatility of 0.15 for all three commodities. At the same time, the instantaneous variance–covariance matrix Σ_x for short-run shocks corresponds to an annual volatility of 0.25 for the first commodity, of 0.30 for the second and of 0.35 for the third and to a correlation coefficient of 0.75 between the first and the second commodities, of 0.50 between the first and the last and of 0.25 between the second and the third. For simplicity, we also assume there is no correlation between the two categories of shocks. Since we focus on the impact of cointegration on spread options, in the following simulations we have set, for illustration purposes, the vector of risk premiums $\boldsymbol{\lambda}_x$ and $\boldsymbol{\lambda}_y$ and the risk-free rate curve equal to zero.[3]

Figure 3 depicts the term structure of correlation, over a period of 5 years, between the returns of futures prices of the three commodities in the system in two cases: the one when the cointegration relation is taken into account and, respectively, the one where the cointegration relation is abstracted from (i.e. $\Theta = \mathbf{O}_3$).

One can observe that, regarding the correlation term structure between commodities 2 and 3, the two curves are identical (Fig. 3, bottom panel). This is not surprising since these two commodities are "exogenous" as explained above and their dynamics is not influenced by the cointegration relation. However, cointegration induces additional correlation when it comes to the commodities 1 and 2 and commodities 1 and 3, as also pointed out at the end of the previous section. In the absence of cointegration, the correlation vanishes after 2–3 years, whereas when the cointegration relation is taken into account the correlation exists also in the long run.

Next, we consider three spreads on two commodities, respectively $S_1(t) - S_2(t)$, $S_1(t) - S_3(t)$, $S_2(t) - S_3(t)$, and one spread on all the three commodities in the system $S_1(t) - 0.5(S_2(t) + S_3(t))$. We assume that at time 0, the two factors are such that $\mathbf{X}(0) = \mathbf{Y}(0) = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$ and, therefore, the current spot prices of all four spreads equal 0. We focus on studying the prices of the at-the-money (ATM) European-style call spread options with up to 5 years to maturity. Figure 4 shows the term structure of

[3] In a real-world application the parameters of the model can be estimated using futures prices data for the corresponding commodities. Given the features of the model one can implement an estimation procedure based on the Kalman filter.

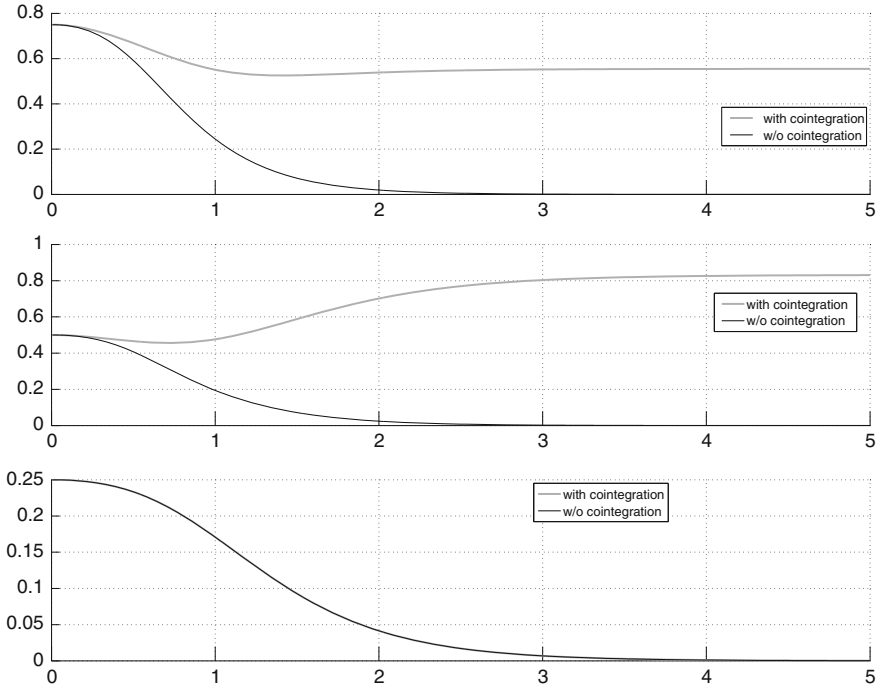

Fig. 3 Term structure of correlation, over a period of 5 years, between the futures log-returns of three commodities (from *top* to *bottom*: between 1 and 2, between 1 and 3, between 2 and 3)

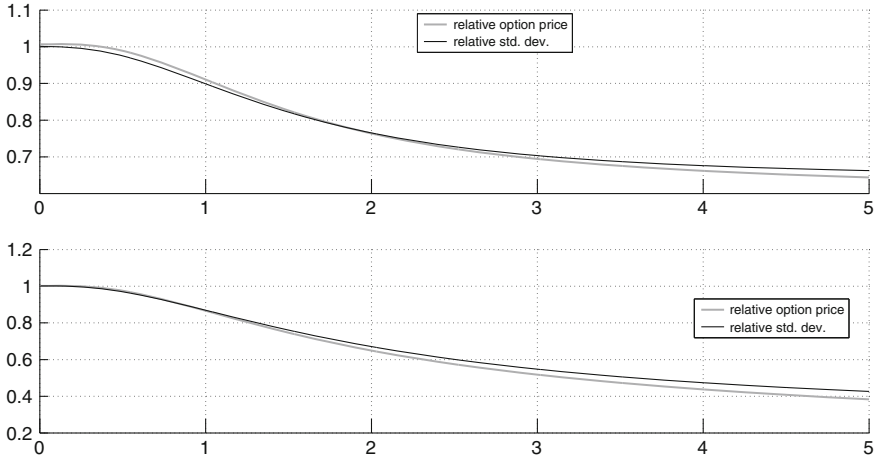

Fig. 4 Relative ATM call spread option prices with up to 5 years to maturity, and relative standard deviations of the spread distribution at maturities up to 5 years. *Top panel* for the spread $S_1(t) - S_2(t)$. *Bottom panel* for the spread $S_1(t) - 0.5(S_2(t) + S_3(t))$

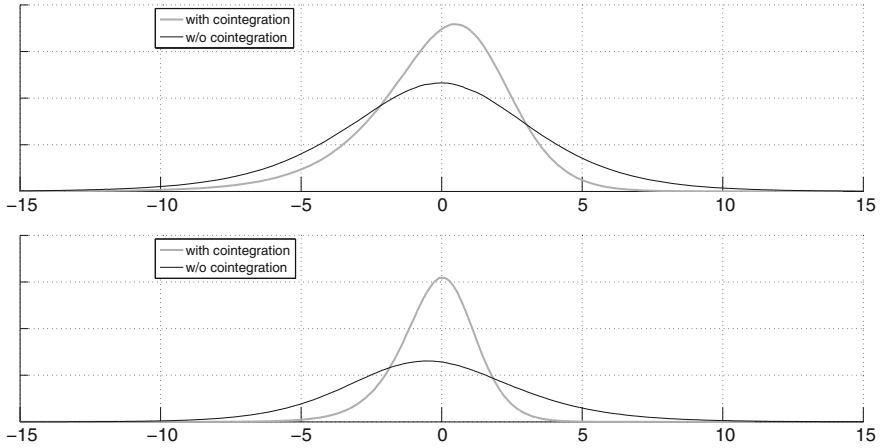

Fig. 5 The distribution of the spread at maturity (5 years). *Top panel* for the spread $S_1(t) - S_2(t)$. *Bottom panel* for the spread $S_1(t) - 0.5(S_2(t) + S_3(t))$

prices in the case with cointegration relative to the prices in the case the cointegration is not accounted for.[4]

Cointegration has a significant impact on spread option prices, with the price for the call with 5 years to maturity on the $S_1(t) - S_2(t)$ spread being almost 30 % lower in the case with cointegration and for the call on the $S_1(t) - 0.5(S_2(t) + S_3(t))$ spread being almost 60 % lower. This can be explained by the fact that cointegration induces additional correlation that acts to lower the standard deviation of the distribution of the spread at maturity. To give a better grasp of this fact, Fig. 5 depicts the distribution of the spread at maturity in the two cases. We omitted from the figures the other two spreads, because the results for the $S_1(t) - S_3(t)$ spread are similar to those for the $S_1(t) - S_2(t)$ spread, and for the $S_2(t) - S_3(t)$ spread there is, as expected given the "exogenous" nature of these two prices, no difference between the cases with and without cointegration.

If one were to add another cointegration relation to the system, linking the second and the third commodities in a long-run relationship, then the new cointegration relation would affect the prices of the options written on the $S_2(t) - S_3(t)$ spread. Moreover, the new cointegration relation might also affect the option prices written on the other three spreads, the magnitude of this influence depending on the structure of the K_y matrix that captures the strength of responses in various spot prices to deviations in the new long-run relationship.

To have a better grasp of the influence of cointegration, next we run a series of sensitivity analyses concerning the existence of a second cointegration relationship in the system. To account for the new cointegration relation, we assume a new structure

[4]Relative quantities in Fig. 4 are determined as the ratio between the quantity computed with the model accounting for cointegration and the corresponding quantity computed with the model without cointegration.

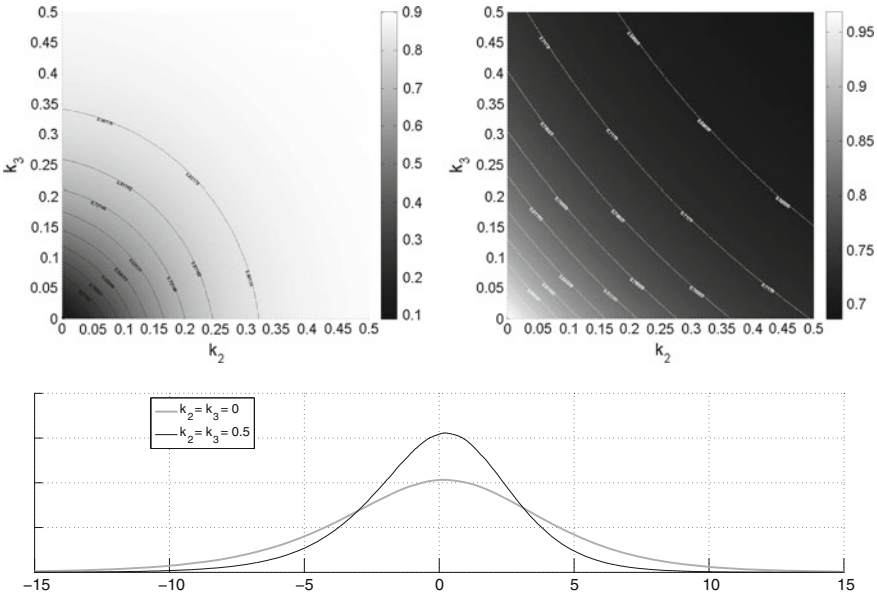

Fig. 6 The impact of k_2 and k_3 on the distribution of the spread $S_2(t) - S_3(t)$ at maturity (5 years). *Top left panel* correlation between the futures log-returns of the two commodities in the basket. *Top right panel* relative standard deviations of the spread distribution (the values are normalized by division with the standard deviation in the case $k_2 = k_3 = 0$). *Bottom panel* the distribution for the two extreme cases in the analysis

for $\Theta = \begin{bmatrix} 1 & -0.4 & -0.6 \\ -\theta & 1 & -0.8 \\ 0 & 0 & 0 \end{bmatrix}$ and $K_y = \begin{bmatrix} 1.5 & k_1 & 0 \\ 0 & k_2 & 0 \\ 0 & -k_3 & 0 \end{bmatrix}$, where $\theta, k_1, k_2, k_3 > 0$. The other parameters have the same values as before. We first focus on the impact of the parameters k_2 and k_3 on the $S_2(t) - S_3(t)$ spread. These two parameters quantify the strength that the second and, respectively, the third commodity reacts to deviations in the newly added cointegration relation. In the extreme case when both k_2 and k_3 are zero, we are in the same situation as before since the two commodities do not react to deviations. However, with the increase of these parameters the new cointegration relation will start to matter for the dynamics of the two commodities, and will have an impact on the distribution of the spread at maturity. Figure 6 presents the results of the sensitivity analysis when k_2 and k_3 are varied between 0 and 0.5, with the other parameters kept fixed at a level $\theta = 0.2$ and $k_1 = 0$.

A higher value for the two reaction parameters produces a higher extra correlation induced by the second cointegration relation, which, in turn, is reflected in a lower standard deviation of the distribution of the spread at maturity. Over a 5-years horizon, the standard deviation for the case $k_2 = k_3 = 0.5$ is 32 % lower than in the case the two parameters are equal to zero, and the ATM call price is 35 % lower.

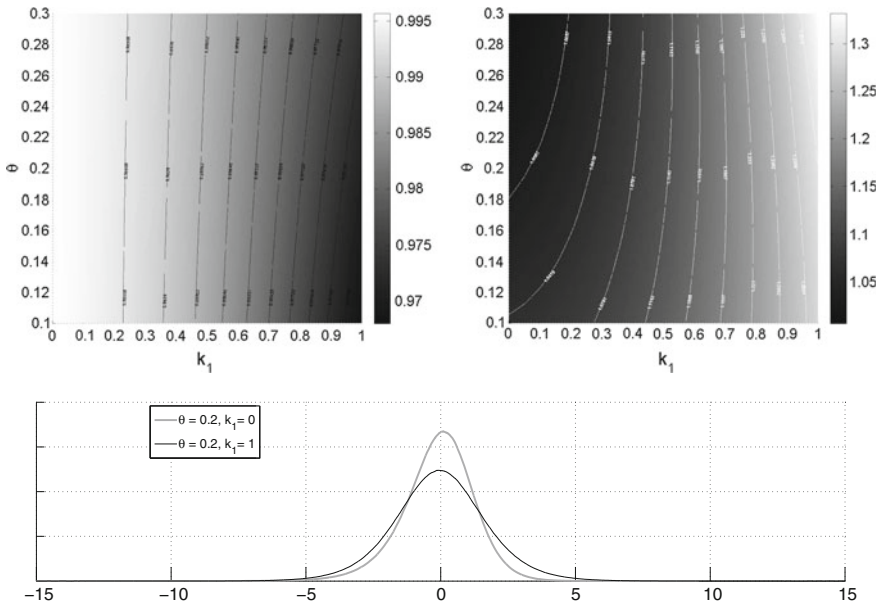

Fig. 7 The impact of k_1 and θ on the distribution of the spread $S_1(t) - 0.5(S_2(t) + S_3(t))$ at maturity (5 years). *Top left panel* correlation between the futures log-returns of the first commodity and the sum of the other two. *Top right panel* relative standard deviations of the spread distribution (the values are normalized by division with the standard deviation in the case $k_1 = 0, \theta = 0.2$). *Bottom panel* the distribution for two specific cases in the analysis

Next, we focus on the impact of θ and k_1 on the $S_1(t) - 0.5(S_2(t) + S_3(t))$ spread. The parameter θ is a free variable that determines the second cointegration relationship and the parameter k_1 measures the magnitude of the response of the first commodity to deviations from the second cointegration relation. Figure 7 presents the results of the sensitivity analysis when k_1 and θ are varied between 0 and 1 and, respectively, between 0.1 and 0.3, with the other parameters kept fixed at a level $k_2 = 0.25$ and $k_3 = 0.25$. An increase of k_1 generates a reduction in the correlation between the components of the spread, showing that the second cointegration relationship has the effect of pulling the components of the spread away from each other. This effect is marginally stronger for the smaller θ. The result of the reduction in correlation is a higher standard deviation of the distribution of the spread at maturity.

For a maturity of 5 years, the standard deviation for the case $k_1 = 1$ is around 33 % higher than in the case the parameter equals zero, and the ATM call price is about 40 % higher. Therefore, the two cointegration relations influence the distribution of the $S_1(t) - 0.5(S_2(t) + S_3(t))$ spread in different directions, the first one generating a reduction, and the second one an increase in the standard deviation. The overall impact depends on the magnitude of the parameters quantifying the responses of the commodities to deviations in the two cointegration relations.

4 Concluding Remarks

In this work, we explored the implications of cointegration between commodity prices on the premiums of options written on various spreads between these commodities. We employed the continuous time model of cointegrated commodity prices developed in Farkas et al. [6] and conducted a simulation study for a cointegrated system of three commodities. We calculated the prices of several spread options and found that cointegration significantly influences these prices. Furthermore, we pointed out that cointegration leads to an upward sloping correlation term-structure which lowers the volatility of spreads and therefore it also lowers the value of options on spreads. Although we restricted in this chapter to a simulation study, it is worthwhile to mention that the model can also be estimated using futures prices on various commodities, as shown in Farkas et al. [6].

Acknowledgements The KPMG Center of Excellence in Risk Management is acknowledged for organizing the conference "Challenges in Derivatives Markets - Fixed Income Modeling, Valuation Adjustments, Risk Management, and Regulation".

References

1. Back, J., Prokopczuk, M.: Commodity price dynamics and derivative valuation: a review. Int. J. Theor. Appl. Financ. **16**(6), 1350032 (2013)
2. Caldana, R., Fusai, G.: A general closed-form spread option pricing formula. J. Bank. Financ. **37**, 48934906 (2013)
3. Cortazar, G., Naranjo, L.: An N-factor Gaussian model of oil futures prices. J. Futur. Mark. **26**(3), 243268 (2006)
4. Cortazar, G., Milla, C., Severino, F.: A multicommodity model of futures prices: using futures prices of one commodity to estimate the stochastic process of another. J. Futur. Mark. **28**(6), 537–560 (2008)
5. Engle, R.F., Granger, C.W.: Co-integration and error correction: representation, estimation, and testing. Econometrica **55**(2), 251–276 (1987)
6. Farkas, W., Gourie, E., Huitema, R., Necula, C.: A two-factor cointegrated commodity price model with an application to spread option pricing. Available at SSRN http://dx.doi.org/10.2139/ssrn.2679218 (2015)
7. Johansen, S.: Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. Econometrica **59**, 1551–1580 (1991)

8. Kirk, E.: Correlation in the energy markets. In: Managing Energy Price Risk, pp. 71–78. Risk Publications and Enron, London (1995)
9. Li, M., Zhou, J., Deng, S.-J.: Multi-asset spread option pricing and hedging. Quant. Financ. **10**(3), 305324 (2010)
10. Paschke, R., Prokopczuk, M.: Integrating multiple commodities in a model of stochastic price dynamics. J. Energy Mark. **2**(3), 4782 (2009)
11. Paschke, R., Prokopczuk, M.: Commodity derivatives valuation with autoregressive and moving average components in the price dynamics. J. Bank. Financ. **34**, 2742–2752 (2010)
12. Phillips, P.: Error correction and long-run equilibrium in continuous time. Econometrica **59**(4), 967–980 (1991)
13. Schwartz, E., Smith, J.E.: Short-term variations and long-term dynamics in commodity prices. Manag. Sci. **46**(7), 893–911 (2000)

The Dynamic Correlation Model and Its Application to the Heston Model

L. Teng, M. Ehrhardt and M. Günther

Abstract Correlation plays an essential role in many problems of finance and economics, such as pricing financial products and hedging strategies, since it models the degree of relationship between, e.g., financial products and financial institutions. However, usually for simplicity the correlation coefficient is assumed to be a constant in many models, although financial quantities are correlated in a strongly nonlinear way in the real market. This work provides a new time-dependent correlation function, which can be easily used to construct dynamically (time-dependent) correlated Brownian motions and flexibly incorporated in many financial models. The aim of using our time-dependent correlation function is to reasonably choose additional parameters to increase the fitting quality on the one hand, but also add an economic concept on the other hand. As examples, we illustrate the applications of dynamic correlation in the Heston model. From our numerical results we conclude that the Heston model extended by incorporating time-dependent correlations can provide a better volatility smile than the pure Heston model.

Keywords Time-dependent correlations · Heston model · Implied volatility · Nonlinear dependence

1 Introduction

Correlation is a well-established concept for quantifying interdependence. It plays an essential role in several problems of finance and economics, such as pricing financial

L. Teng (✉) · M. Ehrhardt · M. Günther
Lehrstuhl für Angewandte Mathematik und Numerische Analysis,
Fakultät Mathematik und Naturwissenschaften, Bergische Universität
Wuppertal, Gaußstr. 20, 42119 Wuppertal, Germany
e-mail: teng@math.uni-wuppertal.de

M. Ehrhardt
e-mail: ehrhardt@math.uni-wuppertal.de

M. Günther
e-mail: guenther@math.uni-wuppertal.de

K. Glau et al. (eds.), *Innovations in Derivatives Markets*, Springer Proceedings
in Mathematics & Statistics 165, DOI 10.1007/978-3-319-33446-2_21

products and hedging strategies. For example, in [3] the arbitrage pricing model is based on that correlation as a measure for the dependence among the assets, and in portfolio credit models the default correlation is one fundamental factor of risk evaluation, see [1, 2, 12].

In most of the financial models, the correlation has been considered as a constant. However, this is not a realistic assumption due to the well-known fact that the correlation is hardly a fixed constant, see e.g. [7, 13]. For example, in many situations the pure Heston model [9] cannot provide enough skews or smiles in the implied volatility surface as market requires, especially for a short maturity. A reason for this might be that deterministically correlated Brownian motions (BMs) of the price process and the variance process are used, as the correlation mainly affects the slope of implied volatility smile. If the correlation is modeled as a time-dependent dynamic function, better skews or smiles will be provided in the implied volatility surface by reasonably choosing additional parameters. Furthermore, compared with the way to extend a model by using time-dependent parameter, e.g., [6, 10] for the Heston model, a time-dependent correlation function adds an economic concept (nonlinear relationship) and its application will be considerably simpler.

The key of modeling correlation as a time-dependent function is being able to ensure that the boundaries -1 and 1 of the correlation function are not attractive and unattainable for any time. In this work, we build up a appropriate time-dependent correlation function, so that one can reasonably choose additional parameters to increase the fitting quality on the one hand but also add an economic concept on the other hand.

The outline of the remaining part is as follows. Section 2 is devoted to a specific dynamic correlation function and its (analytical) computation. In Sect. 3, we present the concept of dynamically (time-dependent) correlated Brownian motions and the corresponding construction. The incorporation of our new dynamic correlation model in the Heston model is illustrated in Sect. 4. Finally, in Sect. 5 we conclude.

2 The Dynamic Correlation Function

In this section we introduce a dynamic correlation function. Actually, it is in high demand to find such a correlation function which must satisfy the correlation properties: it provides only the values in the interval $(-1, 1)$ for any time; it converges for increasing time. We find the following simple idea: we denote the dynamic correlation by $\bar{\rho}$ and propose simply using

$$\bar{\rho}_t := E\left[\tanh(X_t)\right], \quad t > 0 \tag{1}$$

for the *dynamic correlation function,* where X_t is any mean-reverting process with positive and negative values. For the known parameters of X_t, the correlation function $\bar{\rho}_t : [0, t] \to (-1, 1)$ depends only on t. We observe that the dynamic correlation model (1) satisfies the desired properties: first, it is obvious that $\bar{\rho}_t$ takes values only

in $(-1, 1)$ for all t. Besides, it converges for increasing time due to the mean reversion of the used process X_t.

Although we could intuitively observe that the function tanh is eminently suitable for transforming value to the interval $(-1, 1)$, one might still ask whether other functions can also be applied for this purpose, like trigonometric functions or $\frac{2}{\pi}\arctan(\frac{\pi}{2}x)$. In theory, such functions could be used for this purpose. However, the problem is whether one can obtain the expectation of the transformed mean-reverting process by such functions in a closed-form expression. Furthermore, our experiments show that the tendency of the function tanh is more suitable for modeling correlations, see [13].

X_t in (1) could be any mean-reverting process which allows positive and negative outcomes. As an example, let X_t be the *Ornstein–Uhlenbeck process* [14]

$$dX_t = \kappa(\mu - X_t)dt + \sigma dW_t, \quad t \geq 0. \tag{2}$$

We are interested in computing $E[\bar{\rho}_t]$ as a function of the given parameters in (2). We compute $\bar{\rho}_t = E[\tanh(X_t)]$ as

$$\bar{\rho}_t = E[\tanh(X_t)] = E\left[1 - \mathrm{e}^{-X_t} \cdot \frac{2}{\mathrm{e}^{-X_t} + \mathrm{e}^{X_t}}\right] = 1 - E\left[\mathrm{e}^{-X_t} \cdot \frac{1}{\cosh(X_t)}\right]. \tag{3}$$

We set $g(X_t) = 1/\cosh(X_t)$. Applying the results by Chen and Joslin [4], the expectation in (3) can be found in closed-form expression (up to an integral) as

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{g}(u) \cdot E[\mathrm{e}^{-X_t}\mathrm{e}^{iuX_t}]\, du, \tag{4}$$

where $i = \sqrt{-1}$ denotes the imaginary unit and $\hat{g}$ is the Fourier transform of g, in this case is known analytically by $\hat{g}(u) = \pi/\cosh(\frac{\pi u}{2})$. Denoting $CF(t, u|X_0, \kappa, \mu, \sigma)$ as the characteristic function of X_t, the expectation in (4) can be presented by $CF(t, i + u|X_0, \kappa, \mu, \sigma)$. Thus, we obtain the closed-form expression for $\bar{\rho}_t$:

$$\bar{\rho}_t = 1 - \frac{1}{2}\int_{-\infty}^{\infty} \frac{1}{\cosh(\frac{\pi u}{2})} \cdot CF(t, i + u|X_0, \kappa, \mu, \sigma)du. \tag{5}$$

The next step is to calculate $CF(t, i + u|X_0, \kappa, \mu, \sigma)$. The process X_t is an Ornstein–Uhlenbeck process and its characteristic function $CF(t, u|X_0, \kappa, \mu, \sigma)$ can be obtained analytically, e.g. using the framework of the affine process, see [5]. Then, we only need to substitute $u + i$ for u in the characteristic function of X_t to calculate $CF(t, i + u|X_0, \kappa, \mu, \sigma)$ which is given by

$$CF(t, i + u|X_0, \kappa, \mu, \sigma) = \mathrm{e}^{-A(t) - \frac{B(t)}{2} + iu(A(t)+B(t)) + u^2\frac{B(t)}{2}}, \tag{6}$$

with

$$A(t) = \mathrm{e}^{-\kappa t} X_0 + \mu(1 - \mathrm{e}^{-\kappa t}), \quad B(t) = -\frac{\sigma^2}{2\kappa}(1 - \mathrm{e}^{-2\kappa t}) \tag{7}$$

Finally, the dynamic correlation function $\bar{\rho}_t$ can be computed by

$$\bar{\rho}_t = 1 - \frac{\mathrm{e}^{-A(t)-\frac{B(t)}{2}}}{2} \int_{-\infty}^{\infty} \frac{1}{\cosh(\frac{\pi u}{2})} \cdot \mathrm{e}^{iu(A(t)+B(t))+u^2\frac{B(t)}{2}} du, \tag{8}$$

where $A(t)$ and $B(t)$ are defined in (7). In fact, X_0 in $A(t)$ is equal to $\operatorname{artanh}(\bar{\rho}_0)$.

To illustrate the role of each parameter in (8), we plot $\bar{\rho}_t$ for several values of the parameters. First in Fig. 1, we let $\kappa = 2$ and $\sigma = 0.5$ and display $\bar{\rho}_t$ with different values of μ, which is set to be 0.5, 0, and -0.5, respectively. Obviously, μ determines the long term mean of $\bar{\rho}_t$. However, μ is not the exact limiting value. Considering Fig. 1a where the initial value of the correlation function is 0, we see that $\bar{\rho}_t$ is increasing to a value around $\mu = 0.5$ and decreasing to a value around $\mu = -0.5$ as t become larger, when $\mu = 0.5$ and -0.5, respectively. Besides, for $\mu = \bar{\rho}_0 = 0$ we observe that the correlation function $\bar{\rho}_t$ yields always 0 which is the same as constant correlation $\rho = 0$. Now, we set $\bar{\rho}_0 = 0.3$ and keep the value of all other parameters unchanged, then display the curves of $\bar{\rho}_t$ in Fig. 1b.

Next, we fix $\kappa = 2$ and $\mu = 0.5$ and then display $\bar{\rho}_t$ for the varying $\sigma = 0.5,\ 1$ and 2 in Fig. 2. Obviously, σ shows the magnitude of variation from the transformed mean value of X_t ($\mu = 0.5$). In Fig. 2a we see, the larger the value of σ is, the stronger the deviations of $\bar{\rho}_t$ is from the transformed mean value of X_t. More interesting is that $\bar{\rho}_t$ first decreases until $t \approx 0.25$, then increases and converges to a value, see Fig. 2b where $\bar{\rho}_0 = 0.3$ and $\sigma = 2$.

Again, in order to illustrate the role of κ, we set $\mu = 0.5,\ \sigma = 2$ and vary the value of κ, see Fig. 3. From Fig. 3a it is easy to observe that κ represents the speed of $\bar{\rho}_t$ tending to its limit. Especially, as we have seen in Fig. 2b, the curve is more

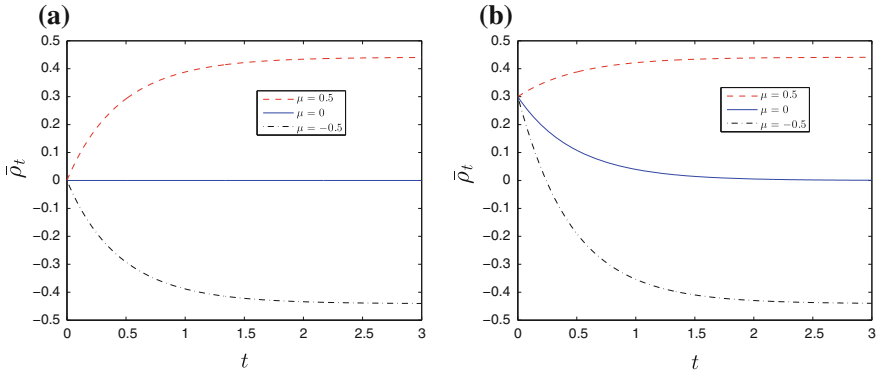

Fig. 1 Dynamic correlation $\bar{\rho}_t$ for varying μ ($\kappa = 2$ and $\sigma = 0.5$). **a** $\bar{\rho}_0 = 0$. **b** $\bar{\rho}_0 = 0.3$

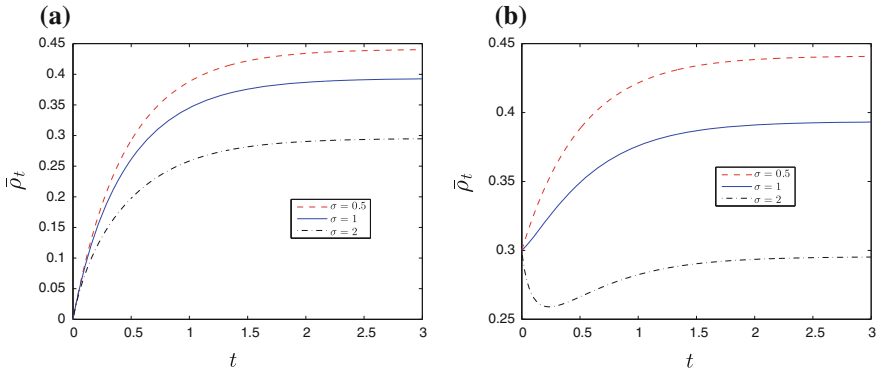

Fig. 2 Dynamic correlation $\bar{\rho}_t$ for varying σ ($\kappa = 2$ and $\mu = 0.5$). **a** $\bar{\rho}_0 = 0$. **b** $\bar{\rho}_0 = 0.3$

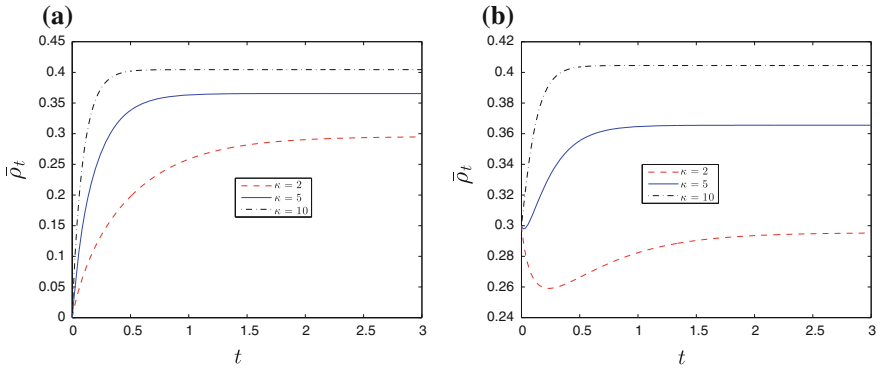

Fig. 3 Dynamic correlation $\bar{\rho}_t$ for varying κ ($\mu = 0.5$ and $\sigma = 2$). **a** $\bar{\rho}_0 = 0$. **b** $\bar{\rho}_0 = 0.3$

unstable for $\kappa = 2$ and $\sigma = 2$ in Fig. 3b. However, if σ remains constant while the value of κ is increased, we can see that curves of $\bar{\rho}_t$ become more stable and tend straightly to its limit. If one incorporates the dynamic correlation function (8) to a financial model, the parameter $\bar{\rho}_0$, κ, μ, and σ could be estimated by fitting the model to market data.

3 Dynamically Correlated BMs and Their Construction

We fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and an information filtration $(\mathscr{F}_t)_{t\in\mathbb{R}^+}$, satisfying the usual conditions, see e.g. [11]. At a time $t > 0$, the correlation coefficient of two Brownian motions (BMs) W_t^1 and W_t^2 is defined as

$$\rho_t^{1,2} = \frac{E\left[W_t^1 W_t^2\right]}{t}. \tag{9}$$

If we assume that $\rho_t^{1,2}$ is constant, $\rho_t^{1,2} = \rho^{1,2}$ for all $t > 0$, say W_t^1 and W_t^2 are correlated with the constant $\rho^{1,2}$.

Therefore, we give the definition of dynamically correlated BMs.

Definition 1 Two Brownian motions W_t^1 and W_t^2 are called *dynamically correlated* with correlation function ρ_t, if they satisfy

$$E\left[W_t^1 W_t^2\right] = \int_0^t \rho_s ds, \tag{10}$$

where $\rho_t : [0, t] \to [-1, 1]$. The *average correlation* of W_t^1 and W_t^2, ρ_{Av}, is given by $\rho_{Av} := \frac{1}{t}\int_0^t \rho_s ds$.

We consider first the two-dimensional case and let ρ_t be a correlation function. For two independent BMs W_t^1 and W_t^3 we define

$$W_t^2 = \int_0^t \rho_s dW_s^1 + \int_0^t \sqrt{1-\rho_s^2}\, dW_s^3, \tag{11}$$

with the symbolic expression

$$dW_t^2 = \rho_t dW_t^1 + \sqrt{1-\rho_t^2}\, dW_t^3. \tag{12}$$

It can be easily verified that W_t^2 is a BM and correlated with W_t^1 dynamically by ρ_t. Besides, the covariance matrix and the average correlation matrix of $\mathbb{W}_t = (W_t^1, W_t^2)$ can be determined, given by

$$\begin{pmatrix} t & \int_0^t \rho_s ds \\ \int_0^t \rho_s ds & t \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & \frac{1}{t}\int_0^t \rho_s ds \\ \frac{1}{t}\int_0^t \rho_s ds & 1 \end{pmatrix},$$

respectively.

The construction above could be also generalized to n-dimensions. We denote a standard n-dimensional BM by $\mathbb{Z}_t = (Z_{1,t}, \ldots, Z_{n,t})$ and the matrix of dynamic correlations $\mathscr{R}_t = (\rho_t^{i,j})_{1<i,j<n}$ which has the Cholesky decomposition for each time t, $\mathscr{R}_t = \mathbb{A}_t \mathbb{A}_t^\top$ with $\mathbb{A}_t = (a_t^{i,j})_{1<i,j<n}$. We define a new n-dimensional process $\mathbb{W}_t = (W_{1,t}, \ldots, W_{n,t})$ by

$$W_{i,t} = \sum_{j=1}^{n} a_t^{ij}\, dZ_{j,t}, \quad i = 1, \ldots, n. \tag{13}$$

We can easily verify that $\mathbb{W}_t$ satisfies the following properties:

- $\mathbb{W}_0 = \mathbf{0}$ and the paths are continuous with probability 1.
- The increments $\mathbb{W}_{t_1} - \mathbb{W}_{t_0}$ and $\mathbb{W}_{t_2} - \mathbb{W}_{t_1}$ are independent for $0 \le t_0 < t_1 < t_2 < t$.
- For $0 \le s < t$, the increment $\mathbb{W}_t - \mathbb{W}_s$ is multivariate normally distributed with mean zero and covariance matrix $\Sigma : \mathbb{W}_t - \mathbb{W}_s \sim N(0, \Sigma)$ with

$$\Sigma = \begin{pmatrix} t-s & \int_s^t \rho_u^{1,2} du & \dots & \int_s^t \rho_u^{1,n} du \\ \int_s^t \rho_u^{2,1} du & t-s & \dots & \int_s^t \rho_u^{2,n} du \\ \vdots & \vdots & \ddots & \vdots \\ \int_s^t \rho_u^{n,1} du & \int_s^t \rho_u^{n,2} du & \dots & t-s \end{pmatrix}.$$

We call the process $(\mathbb{W}_t)_{t \ge 0}$ *an n-dimensional dynamically correlated Brownian motion,* with the correlation matrix $\mathscr{R}_t$.

4 Dynamic Correlation in the Heston Model

As mentioned before, in many situations the pure Heston model has a limitation on reproducing properly a volatility smile. For this problem, several time-dependent Heston models have been proposed for good fitting to implied volatilities, e.g. [6] and [10]. In this section, we show how to incorporate our time-dependent correlation function into the Heston model.

4.1 Incorporating Dynamic Correlations

Heston's stochastic volatility model is specified as

$$dS_t = \mu_S S_t dt + \sqrt{v_t}\, S_t\, dW_t^S, \tag{14}$$
$$dv_t = \kappa_v(\mu_v - v_t)dt + \sigma_v \sqrt{v_t}\, dW_t^v, \tag{15}$$

where (14) is the price of the spot asset, (15) is the volatility (variance) and W_t^S and W_t^v are correlated with a constant correlation ρ_{Sv}. To incorporate the time-dependent correlations, we assume that dS_t and dv_t are correlated by a time-dependent correlation function $\bar{\rho}_t$ instead of the constant correlation ρ_{Sv}. The extended Heston model with dynamic correlation $\bar{\rho}$ is specified as

$$dS_t = \mu_S S_t dt + \sqrt{v_t}\, S_t\, dW_t^1, \tag{16}$$
$$dv_t = \kappa_v(\mu_v - v_t)dt + \sigma_v \sqrt{v_t} \left(\bar{\rho}_t\, dW_t^1 + \sqrt{1 - \bar{\rho}_t^2}\, dW_t^2 \right), \tag{17}$$

where W_t^1 and W_t^2 are independent. Applying Itô's lemma and no-arbitrage arguments yields [9]

$$\frac{1}{2}\nu S^2\frac{\partial^2 U}{\partial S^2}+\bar{\rho}_t\sigma_\nu\nu S\frac{\partial^2 U}{\partial S\partial\nu}+\frac{1}{2}\sigma_\nu^2\nu\frac{\partial^2 U}{\partial\nu^2}+rS\frac{\partial U}{\partial S}+[\kappa_\nu(\mu_\nu-\nu)-\tilde{\lambda}(S,\nu,\bar{\rho},t)\nu]\frac{\partial U}{\partial\nu}-rU+\frac{\partial U}{\partial t}=0, \tag{18}$$

where $\bar{\rho}_t$ is defined in (8) but with the parameter $\bar{\rho}_0$, κ_ρ, μ_ρ, and ν_ρ. It is worth mentioning that the market price of volatility risk depends also on the dynamic correlation, which could be written as $\tilde{\lambda}(S,\nu,\bar{\rho}_t,t)$. This means, the price of correlation risk embedding in the price of volatility risk has been considered.

We consider, e.g. a European call option with strike price K and maturity T in the Heston model

$$C(S,\nu,t,\bar{\rho}_t)=SP_1-KP(t,T)P_2,\quad \tau=T-t, \tag{19}$$

where $P(t,T)$ is the discount factor and both in-the-money probabilities P_1,P_2 must satisfy the PDE (18) as well as their characteristic functions, $f_1(S_t,\nu_t,\bar{\rho}_t,\phi,t)$ and $f_2(S_t,\nu_t,\bar{\rho}_t,\phi,t)$

$$f_j(S_t,\nu_t,\bar{\rho}_t,\phi,t)=E[\mathrm{e}^{i\phi\ln S_T}|S_t,\nu_t,\bar{\rho}_t]=\mathrm{e}^{C_j(\tau,\phi)+D_j(\tau,\phi)\nu+i\phi\ln S_t},\quad j=1,2, \tag{20}$$

where $C_j(0,\phi)=0$ and $D_j(0,\phi)=0$. By substituting this functional form (20) into the PDE (18) we can obtain the following ordinary differential equations (ODEs) for the unknown functions C and D:

$$-\frac{1}{2}\phi^2+\bar{\rho}_t\sigma_\nu\phi iD_j+\frac{1}{2}\sigma_\nu^2D_j^2+u_j\phi i-b_jD_j+\frac{\partial D_j}{\partial t}=0, \tag{21}$$

$$r\phi i+\kappa_\nu\mu_\nu D_j+\frac{\partial C_j}{\partial t}=0, \tag{22}$$

with the initial conditions $C_j(0,\phi)=D_j(0,\phi)=0$

$$u_1=0.5,\quad u_2=-0.5,\quad b_1=\kappa_\nu+\lambda-\bar{\rho}_t\sigma_\nu\quad\text{and}\quad b_2=\kappa_\nu+\lambda, \tag{23}$$

where

$$\bar{\rho}_t=1-\frac{\mathrm{e}^{-A(t)-\frac{B(t)}{2}}}{2}\int_{-\infty}^{\infty}\underbrace{\frac{1}{\cosh(\frac{\pi u}{2})}\cdot\mathrm{e}^{iu(A(t)+B(t))+u^2\frac{B(t)}{2}}}_{:=g(u)}du, \tag{24}$$

with $A(t)=\mathrm{e}^{-\kappa_\rho t}\mathrm{artanh}(\bar{\rho}_0)+\mu_\rho(1-\mathrm{e}^{-\kappa_\rho t})$, $B(t)=-\frac{\sigma_\rho^2}{2\kappa_\rho}(1-\mathrm{e}^{-2\kappa_\rho t})$.

Obviously, (21) and (22) cannot be solved analytically. Therefore, we need to find an efficient way to compute the option price numerically. We firstly generate the

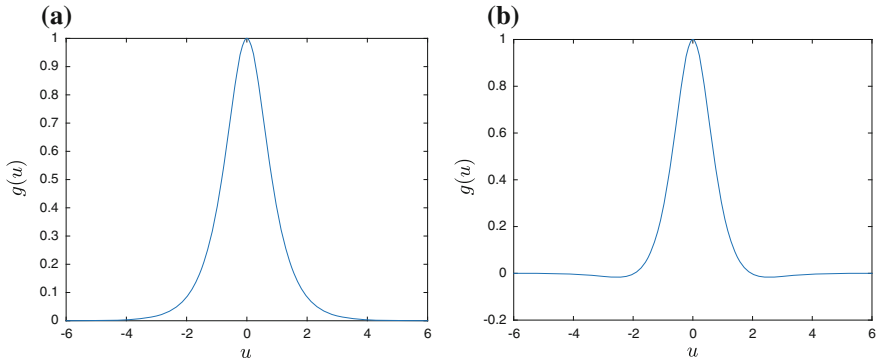

Fig. 4 $g(u)$ under $\rho_0 = 0.3, \kappa_\rho = 2, \mu_\rho = -0.8, \sigma_\rho = 0.1$. **a** $t = 0.1$. **b** $t = 10$

dynamic correlations using (24). We observe that $g(u)$ is a symmetric function about $u = 0$ and vanishes (approaches zero) for a sufficiently large absolute value of u, see Fig. 4. For these two reasons, the numerical integration in (24) is computationally fast. Next we use an explicit Runge–Kutta method, the matlab routine `ode45`, to obtain C and D in (21) and (22) and thus also the characteristic functions (20). Finally, we employ the COS method [8] to obtain the option price $C(S, \nu, t, \bar{\rho})$ in (19). Thanks to the COS method, although we solved that ODE system numerically, the time for obtaining European option prices is less than 0.1 s so that a calibration can be performed. Obviously, the error consists of the error using `ode45` for (21) and (22) and the error using COS method. The detailed analysis of error using COS method has been provided in [8].

4.2 Calibration of the Heston Model Under Dynamic Correlation

In this section we calibrate the Heston model extended by our time-dependent correlation function to the real market data (Nikk300 index call options on July 16, 2012) and compare these to the pure Heston model [9] and the time-dependent Heston model [10].

We consider a set of N maturities T_i, $i = 1, \ldots, N$ and a set of M strikes K_j, $j = 1, \ldots, M$. Then for each combination of maturity and strike we have a market price $V^M(T_i, K_j) = V_{ij}^M$ and a corresponding model price $V(T_i, K_j; \Theta) = V_{ij}^\Theta$ generated by using (19). We choose the relative mean error sum of squares (RMSE) for the loss function $\frac{1}{M \times N} \sum_{i,j} \frac{(V_{ij}^M - V_{ij}^\Theta)^2}{V_{ij}^M}$, which can be minimized to obtain the parameter estimates

$$\hat{\Theta} = \arg\min \frac{1}{M \times N} \sum_{i,j} \frac{(V_{ij}^M - V_{ij}^\Theta)^2}{V_{ij}^M}. \tag{25}$$

For the optimization we restrict $\bar{\rho}_0$ to the interval $(-1, 1)$ but not the value of μ_ρ. Since it is not the direct limit of the correlation function but the mean reversion of the Ornstein–Uhlenbeck process, thus, it could take any value in $\mathbb{R}$. Our experiments showed, that it is sufficient and appropriate to restrict μ_ρ to the interval $[-4, 4]$.

We state our estimated parameters and the estimation error for the pure Heston model (abbr. PH), the Heston model under our time-dependent correlations (CH), the time-dependent Heston model by Mikhailov and Ngel [10] (MN) in Tables 1, 2 and 3, respectively. We see that the estimation error using the CH model is significantly less than the error using the PH model and almost the same to the error (sum of errors for each maturity) under the MN model. To illustrate more clearly, for each maturity we compare the implied volatilities for all the models to the market volatilities in Fig. 5. We can observe that the implied volatilities for the CH model are much closer to the market volatilities than the implied volatilities for the PH model, especially the CH model has better volatility smile for the short maturity $T = 1/12$. Compared to the MN model, the implied volatilities for our model are almost the same. However, our CH model has an economic interpretation, namely the correlation is nonlinear

Table 1 The estimated parameters for the pure Heston model using call options on the Nikk300 index on July 16, 2012 for the maturities 1/12, 1/4, 1/2, 1

The pure heston model					
$\hat{v}_0$	$\hat{\kappa}_v$	$\hat{\mu}_v$	$\hat{\sigma}_v$	$\hat{\rho}$	Estimation error
0.029	4.746	0.053	1.108	−0.355	1.10×10^{-3}

Table 2 The estimated parameters for the Heston model under time-dependent correlations using call options on the Nikk300 index on July 16, 2012 for the maturities 1/12, 1/4, 1/2, 1

The extended Heston model by using our time-dependent correlation function								
$\hat{v}_0$	$\hat{\kappa}_v$	$\hat{\mu}_v$	$\hat{\sigma}_v$	$\hat{\bar{\rho}}_0$	$\hat{\kappa}_\rho$	$\hat{\mu}_\rho$	$\hat{\sigma}_\rho$	Estimation error
0.027	5.542	0.055	1.224	−0.165	5.333	−0.752	0.434	2.38×10^{-4}

Table 3 The estimated parameters for the time-dependent Heston model by Mikhailov and Ngel using call options on the Nikk300 index on July 16, 2012

The time-dependent Heston model by Mikhailov and Ngel						
Maturity	$\hat{v}_0$	$\hat{\kappa}_v$	$\hat{\mu}_v$	$\hat{\sigma}_v$	$\hat{\rho}$	Estimation error
1/12	0.025	2.749	0.095	1.172	−0.201	1.78×10^{-4}
1/4	0.012	2.936	0.076	0.524	−0.411	2.45×10^{-5}
1/2	0.011	2.890	0.058	0.592	−0.430	1.14×10^{-5}
1	0.001	2.911	0.051	0.558	−0.389	4.28×10^{-6}

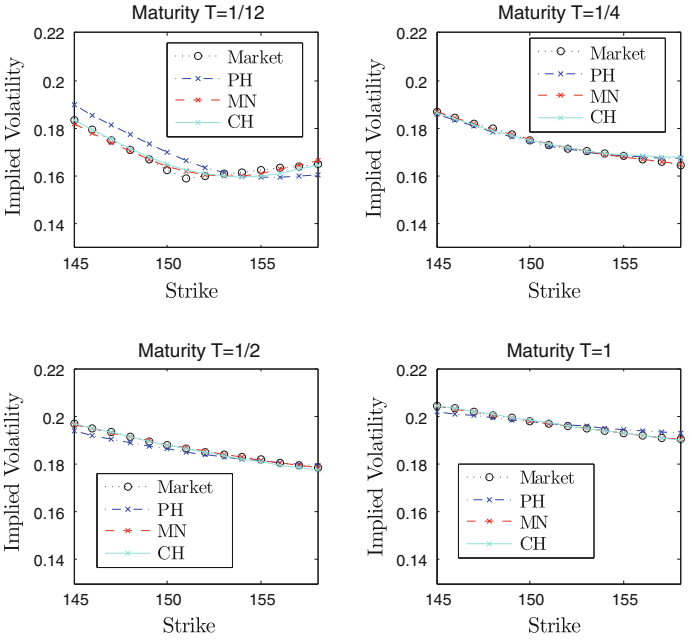

Fig. 5 The comparison of implied volatilities for all the models to the market volatilities of the call options on the Nikk300 index on July 16, 2012, where the spot price is 150.9

and time-dependent as market requires. We conclude that the Heston model extended by incorporating our time-dependent correlations can provide better volatility smiles compared to the pure Heston model. The time-dependent correlation function can be easily and directly introduced into the financial models.

5 Conclusion

In this work, we first investigated the dynamically (time-dependent) correlated Brownian motions and their construction. Furthermore, we proposed a new dynamic correlation function which can be easily incorporated into another financial model. The aim of using our dynamic correlation function is to reasonably choose additional parameters to increase the fitting quality on the one-hand side, but also add an economically meaningful perspective.

As an application, we incorporated our time-dependent correlation function into the Heston model. An experiment on estimation of the models using real market data has been provided. The numerical calibration results show that the Heston model extended by using our time-dependent correlation function provides better volatility

smiles compared to the pure Heston model. Besides, this time-dependent correlation function could be easily and directly imposed to the financial models and thus it is preferred to use instead of a constant correlation.

Acknowledgements The authors acknowledge the much appreciated inspiration and in-depth discussions with Dr. Jörg Kienitz from Deloitte Düsseldorf, Germany.
The work was partially supported by the European Union in the FP7-PEOPLE-2012-ITN Program under Grant Agreement Number 304617 (FP7 Marie Curie Action, Project Multi-ITN STRIKE—Novel Methods in Computational Finance).
The KPMG Center of Excellence in Risk Management is acknowledged for organizing the conference "Challenges in Derivatives Markets—Fixed Income Modeling, Valuation Adjustments, Risk Management, and Regulation".

References

1. Brigo, D., Capponi, A.: Bilateral counterparty risk with application to CDSs. Risk Mag. **23**(3), 85–90 (2010)
2. Brigo, D., Chourdakis, K.: Counterparty risk for credit default swaps: impact of spread volatility and default correlation. Int. J. Theoret. Appl. Financ. **12**, 1007–1026 (2009)
3. Campbell, J., Lo, A., MacKinlay, A.: The Econometrics of Financial Markets. Princeton University Press, Princeton (1997)
4. Chen, H., Joslin, S.: Generalized transform analysis of affine processes and applications in finance. Rev. Fin. Stud. **25**(7), 2225–2256 (2012)
5. Duffie, D., Filipović, D., Schachermayer, W.: Affine processes and applications in finance. Ann. Appl. Probab. **13**(3), 984–1053 (2003)
6. Elices, A.: Affine concatenation. Wilmott J. **1**(3), 155–162 (2009)
7. Escobar, M., Götz, B., Neykova, D., Zagst, R.: Stochastic correlation and volatility mean-reversion - empirical motivation and derivatives pricing via perturbation theory. Appl. Math. Fin. **21**(6), 555–594 (2014)
8. Fang, F., Oosterlee, C.W.: A novel pricing method for European options based on Fourier-cosine series expansions. SIAM J. Sci. Comput. **31**, 826–848 (2008)
9. Heston, S.L.: A closed-form solution for options with stochastic volatility with applications to bond and currency options. Rev. Fin. Stud. **6**(2), 327–343 (1993)
10. Mikhailov, S., Nögel, U.: Heston's stochastic volatility model: implementation, calibration and some extensions. Wilmott Mag. 74–79 (2003)
11. Øksendal, B.: Stochastic Differential Equations. Springer, Berlin (2002)
12. Teng, L., Ehrhardt, M., Günther, M.: Bilateral counterparty risk valuation of CDS contracts with simultaneous defaults. Int. J. Theoret. Appl. Financ. **16**(7), 1350040 (2013)

13. Teng, L., van Emmerich, C., Ehrhardt, M., Günther, M.: A versatile approach for stochastic correlation using hyperbolic functions. Int. J. Comput. Math. **93**(3), 524–539 (2016)
14. Uhlenbeck, G.E., Ornstein, L.S.: On the theory of Brownian motion. Phys. Rev. **36**, 823–841 (1930)