POLITECNICO DI TORINO
Repository ISTITUZIONALE

Dependence calibration and portfolio fit with factor-based subordinators

Original
Dependence calibration and portfolio fit with factor-based subordinators / Luciano, E., Marena, M., Semeraro, P.. - In: QUANTITATIVE FINANCE. - ISSN 1469-7688. - 16:7(2016), pp. 1037-1052. [10.1080/14697688.2015.1114661]

Availability:
This version is available at: 11583/2642714 since: 2020-01-24T14:45:30Z

Publisher:
Routledge

Published
DOI:10.1080/14697688.2015.1114661

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

06 September 2026

Dependence Calibration and Portfolio Fit with Factor-Based Subordinators

Elisa Luciano[1] Marina Marena[2] and Patrizia Semeraro[3]

[1]Department of Economics and Statistics, University of Torino, Collegio Carlo Alberto and Netspar.

[2]Department of Economics and Statistics, University of Torino

[3]Politechnical School of Torino

Abstract

The paper explores the properties of a class of multivariate Lévy processes, used for asset returns, with a focus on describing in an economic sensible and empirically appropriate way both linear and nonlinear dependence. The processes are subordinated Brownian motions. The subordinator has a common and an idiosyncratic component, to reflect the properties of trade, which it represents. A calibration to a portfolio of ten US stock indices returns over the period 2009-2013 shows that the hyperbolic specification fits very well marginal distributions, the overall correlation matrix and the return distribution of both long-only and long-short random portfolios, which incorporate also nonlinear dependence. Their tail behavior is well captured also by the variance gamma specification. The main message is not only the goodness of fit, but also the flexibility in capturing dependence and the easiness of calibration on large sets of returns.
Journal of Economic Literature Classification: G12, G13
Keywords: Lévy processes, multivariate subordinators, dependence, correlation, multivariate asset modelling, multivariate subordinated processes, factor-based subordinators.

Introduction

A number of Lévy processes which extend the classical Gaussian benchmark have been adopted to represent stock returns. Among these, the variance gamma, the normal inverse Gaussian, as well as the generalized hyperbolic model have good analytical tractability as well as fit properties which explain their popularity, especially for derivative pricing. These models have been adopted to represent single-asset returns. All of them can be represented as subordinated Brownian motions.

They have been extended in various ways to represent returns on several assets - multivariate returns - whose marginal distributions belong to the same class. A multivariate model can be obtained which preserves variance gamma, normal inverse Gaussian, or hyperbolic marginal distributions, by applying the same subordinator to a multivariate Brownian motion. The technique is straightforward and the economic interpretation is appealing, because the time change represents the transformation from business time to calendar time (Clark (1973)). In business time returns can be Gaussian, thanks to the central-limit property. In calendar time though they are not Gaussian any more, since the subordinator represents information arrival and trading activity. The more intense the market activity, the faster economic time runs relative to calendar time and the greater, in a sense, the departure from Gaussianity. However, there is a long-standing empirical literature which shows that trade may be significantly different across assets (Harris (1986)), so the same subordinator may not be appropriate. Because recent empirical evidence shows that trade in different financial assets presents at least one common factor (Lo and Wang (2000)), a natural extension of subordinated processes from a univariate to a multivariate setting, which preserves the marginal distributions, is through a multivariate subordinator, with at least one common factor (see for instance Semeraro (2008)). The factor structure makes it possible to fully and analytically characterize the return process, and can reproduce a single subordinator if the common component is predominant. Alternative constructions, as such those which keep a distinct subordinator for each marginal distribution and correlate the Brownian motions without providing a full process construction and characterization, are an appealing, straightforward alternative. Eberlein and Madan (2010), for instance, correlate the unit-time random variables resulting from the marginal laws. On top of the greater or smaller economic rationale for the subordinated process construction, and the greater or smaller availability of analytic expressions and properties, calibration of the models which maintain variance gamma, normal inverse Gaussian or hyperbolic marginal distributions is not always satisfactory in terms of fit, especially with more than two processes.

In this paper we show that the factor-based processes can be constructed in a less cumbersome way than in previous studies and have full analytical characterization. **Furthermore, factor based processes can be specified to have marginal processes which are important for financial modelling for their distributional properties and their analitical tractability. We consider here the variance**

gamma (Madan and Seneta (1990)), normal inverse gaussian (Barndorff-Nielsen (1995)) and generalized hyperbolic specifications Barndorff-Nielsen (1978), which for their features are suitable for option pricing applications, see for example ?.

They are also simple to calibrate and provide a very good fit of single assets, large bundles of them and portfolios. The reason why we look at portfolios is that – unlike the correlation matrix – they reflect linear and nonlinear dependence. We focus on a class of factor-based processes, first introduced by Luciano and Semeraro (2010b) and Luciano and Semeraro (2010a), and on their ability to describe in an economic sensible and empirically appropriate way both linear and nonlinear dependence. We perform an extensive, not-pairwise calibration on ten MSCI stock indices over a turbolent period, from January 2009 to May 2013. The calibration shows that the hyperbolic specification fits marginal distributions and sample correlations with greater accuracy than either the Black-Scholes model or the Eberlein and Madan model. We choose the latter as a benchmark because it is as simple as ours to calibrate. When we compare the performance on portfolios, we still find that factor-based models outperform the others, both when evaluated on long-only and on long-short portfolios. In terms of overall distribution fit, as expressed by the Kolmogorov distance, both the variance gamma and hyperbolic distributions provide a good performance, the latter specification being better than the former for long-short portfolios. The main advantage of the factor-based process is not only the goodness of fit, but also the flexibility in capturing dependence and the easiness of calibration on large sets of returns.

The outline of the paper is as follows: Section 1 presents background and motivation. Section 2 introduces the factor-based multivariate subordinated processes, providing a new construction. Section 3 shows that, under suitable parameter restrictions, the model can be specified so as to have variance gamma, normal inverse Gaussian and generalized hyperbolic marginal distributions. Section 4 discusses linear and nonlinear dependence in the model. Section 5 illustrates the approach to correlate subordinated Brownian motions adopted by Eberlein and Madan. Section 6 is devoted to the two model calibrations. We first illustrate the calibration procedure, then the fit. Section 6.4 shows the portfolios fit. The final Section summarizes and concludes.

1 Background and motivation

The construction of multivariate Lévy processes by means of a subordinator acting on several assets dates back at least to Luciano and Schoutens (2006), Leoni and Schoutens (2008). Both use a single subordinator, i.e. a single non-negative process which represents the change of time; the former paper has independent Brownian motions, while the second has correlated Brownian motions, to incorporate the idea that returns may be correlated over and above the trade effect incorporated in the subordinator. The

shift from the first to the second model reflects the attempt to better capture the high correlation, frequently observed in the data, that independent Brownian motions do not reproduce.

The construction technique by means of a multivariate – instead of a single – subordinator is due to Barndorff-Nielsen *et al.* (2001). Semeraro (2008) introduces for the first time a factor-based subordinator, with a component common to all assets and an idiosyncratic component. The economic rationale is that trade can be explained by at least one factor, as in Lo and Wang (2000). She constructs a multivariate variance gamma process, named α-Variance Gamma (αVG). Recently, Buchmann *et al.* (2014) have generalized the αVG process by changing the distribution of the multivariate subordinator. Nevertheless, the independence of the Brownian motions limits the the possibility of capturing a wide range of dependence – what we call dependence flexibility – similarly to what happens in the αVG case. To improve the dependence flexibility, Buchmann *et al.* (2014) define the log price process as a linear transformation of the subordinated process. By so doing, however, it happens that each asset has not its own subordinator, i.e. change of time, any more. As a consequence, dependence flexibility increases at the cost of economic intuition.

Luciano and Semeraro (2010b) extend the αVG dependence structure, improving the correlation flexibility, by using correlated Brownian motions. This preserves the intuition that each asset has its own subordinator, but includes the possibility of comovements due to the Brownian component. Although the dependence structure of Luciano and Semeraro (2010b) holds in a general framework, they maintain the Lévy framework, by assuming that the change of time is a subordinator. They apply the technique not only to the variance gamma case but also to other marginal Lévy processes, including the compound Poisson, normal inverse Gaussian (**?**) and Carr Geman Madan Yor (Carr *et al.* (2002)) process. In Luciano and Semeraro (2010a) they use a similar construction with generalized hyperbolic marginal distributions (Eberlein and Prause (2002)). Also Bal (????) use a common and an idiosyncratic component both for the subordinator and for the returns.

The calibrated behavior of the previous models has not been tested massively, especially beyond the pairwase case. As a partial exception, Wallmeier and Diethelm (2012) study the calibration performance of the multivariate model of Leoni and Schoutens and the αVG model on a large dataset of reverse convertibles with three underlying stocks, traded on the Swiss market from January 2009 to May 2013. Both models provide a significant improvement over the multivariate Gaussian model, in terms of smile fit. The αVG model provides a better fit, since the kurtosis parameter can freely vary across stocks, while it cannot when the subordinator is unique, as in Leoni and Schoutens (2008). However, for a large portion of the reverse convertible dataset, both models fail to match the empirical correlation matrix of the underlying assets, and the Leoni and Schoutens model fits a higher percentage of the sample. Wallmeier and Diethelm conclude, as expected, that more flexible approaches are needed to capture the empirical

joint behavior of asset returns. The same conclusions on the αVG model are reached by Guillaume (2012) using the weekly quotes of four major stocks included in the S&P 500 index, from June 2, 2008 to October 20, 2009[1].

A partial improvement is obtained in Bal (????), who calibrate a multivariate Lévy model with fixed marginal distributions to option prices on three stocks belonging to the S&P100 index, considering the variance gamma, normal inverse Gaussian, Merton and Kou jump-diffusion specifications. Calibration is performed in two steps. First, marginal distributions are calibrated. In the second step, the parameters of the idiosyncratic and common component are determined, by fitting the sample correlations and enforcing the convolution conditions necessary to recover the desired marginal distributions. All specifications are able to match the market correlations up to a good level of accuracy. However, the second step of the calibration procedure becomes increasingly challenging from a numerical point of view, as the number of assets taken into consideration grows. Within this framework, Loregian (2013) propose a three-step estimation procedure which is feasible for large portfolios, by dropping the convolution conditions. As a consequence, marginal distributions do not belong anymore to the same Lévy class.

All in all, the few attempts to perform an extensive calibration of subordinated Brownian motions which preserve the variance gamma, the normal inverse Gaussian and the generalized hyperbolic nature of the marginal distributions have given a rather poor fit. This is because they are based on a model which is not flexible enough, or they cannot be easily performed, because of the model construction. Consequently, our main goal consists in selecting a dependence model which permits easy calibration with a large number of assets, provides a good fit of the marginal distributions, is flexible enough to capture (especially high) correlation, and, last but not least, captures well portfolio returns which - as opposed to the correlation matrix - reflect linear and nonlinear dependence.

2 Factor-based subordinated processes

In this section we recall the construction of factor-based subordinated Lévy processes and provide their characteristic function, which is used to derive the distribution of portfolio returns in closed form. We present different specifications of the model, which have variance gamma (VG), normal inverse Gaussian (NIG) and generalized hyperbolic marginal distributions (GH). The first two processes have been introduced in Luciano and Semeraro (2010b), the third one follows from a construction in Luciano and Semeraro (2010a).

We first introduce the class of factor-based multivariate subordinators used to con-

[1]To improve the correlation fit, Guillaume proposed to relax some constraints to the marginal processes. In this paper, we restrict our attention to fixed marginals belonging to one of the classes specified above.

struct the $\mathbb{R}^n$-valued asset return process $\{\boldsymbol{Y}(t),\, t \geq 0\}$. A multidimensional factor-based subordinator $\{\boldsymbol{G}(t), t \geq 0\}$ is defined as follows

$$\boldsymbol{G}(t) = (X_1(t) + \alpha_1 Z(t), ..., X_n(t) + \alpha_n Z(t)), \quad \alpha_j > 0,\, j = 1, ..., n, \tag{2.1}$$

where $\boldsymbol{X}(t) = \{(X_1(t), ..., X_n(t)), t \geq 0\}$ and $\{Z(t), t \geq 0\}$ are independent subordinators with zero drift, and $\boldsymbol{X}(t)$ has independent components. They represent the idiosyncratic and the common factors of trading activity. The subordinated process $\boldsymbol{Y}(t)$ is constructed by subordinating n independent Brownian motions $B_j(t)$ with the independent subordinators $X_j(t)$ and by subordinating a multidimensional Brownian motion $\{\boldsymbol{B}^{\rho}(t),\, t \geq 0\}$, with correlations $\boldsymbol{\rho} = (\rho_{ij})_{i,j=1,...,n}$ **and Lévy triplet** $(\boldsymbol{\mu}^{\rho}, \boldsymbol{\Sigma}^{\rho}, \mathbf{0})$, with the unique subordinator $Z(t)$. Formally, let $\boldsymbol{B}(t) = \{(B_1(t), \ldots, B_n(t)),\, t \geq 0\}$ be a Brownian motion with independent components and Lévy triplet $(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \mathbf{0})$

$$\boldsymbol{\Sigma} = diag(\sigma_1^2, ..., \sigma_n^2) := \begin{pmatrix} \sigma_1^2 & 0... & 0 \\ 0 & \sigma_2^2.... & 0 \\ 0 & 0... & \sigma_n^2 \end{pmatrix}, \quad \boldsymbol{\mu} = (\mu_1, ..., \mu_n), \tag{2.2}$$

and let $\boldsymbol{B}(\boldsymbol{s})$ be the corresponding multi-parameter Brownian Motion, defined as follows. Consider a multiparameter $\boldsymbol{s} = (s_1, ..., s_n)^T \in \mathbb{R}_+^n$ and the partial order on $\mathbb{R}_+^n$

$$\boldsymbol{s}^1 \preceq \boldsymbol{s}^2 \;\; \Leftrightarrow \;\; s_j^1 \leq s_j^2,\, j = 1, ...n.$$

The multi-parameter Brownian motion $\{\boldsymbol{B}(\boldsymbol{s}),\, \boldsymbol{s} \in \mathbb{R}_+^n\}$ is defined by (Barndorff-Nielsen *et al.* (2001))

$$\boldsymbol{B}(\boldsymbol{s}) = \{(B_1(s_1), ..., B_n(s_n)), \boldsymbol{s} \in \mathbb{R}_+^n\} \tag{2.3}$$

With these two Brownian motions we can introduce the following definition

Definition 2.1. *Let $\boldsymbol{B}(\boldsymbol{s})$ be the multi-parameter Brownian Motion in* (2.3) *and let $\boldsymbol{B}^{\rho}(t) = (B_1^{\rho}(t), ..., B_n^{\rho}(t))$ be a multivariate Brownian motion, independent of $\boldsymbol{B}(t)$, with correlations $\boldsymbol{\rho} = (\rho_{ij})_{i,j=1,...,n}$. Let $\boldsymbol{B}^{\rho}(t)$ have the following Lévy triplet $(\boldsymbol{\mu}^{\rho}, \boldsymbol{\Sigma}^{\rho}, \mathbf{0})$, where*

$$\boldsymbol{\mu}^{\rho} = (\mu_1\alpha_1, ..., \mu_n\alpha_n),$$

and **the diffusion matrix is**

$$\boldsymbol{\Sigma}^{\rho} := \begin{pmatrix} \sigma_1^2\alpha_1 & \rho_{12}\sigma_1\sigma_2\sqrt{\alpha_1}\sqrt{\alpha_2} & \cdots & \rho_{1n}\sigma_1\sigma_n\sqrt{\alpha_1}\sqrt{\alpha_n} \\ \rho_{12}\sigma_1\sigma_2\sqrt{\alpha_1}\sqrt{\alpha_2} & \sigma_2^2\alpha_2 & \cdots & \rho_{2n}\sigma_2\sigma_n\sqrt{\alpha_2}\sqrt{\alpha_n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1n}\sigma_1\sigma_n\sqrt{\alpha_1}\sqrt{\alpha_n} & \rho_{2n}\sigma_2\sigma_n\sqrt{\alpha_2}\sqrt{\alpha_n} & \cdots & \sigma_n^2\alpha_n \end{pmatrix}, \tag{2.4}$$

with

$$\boldsymbol{\mu} \in \mathbb{R}^n, \quad \alpha_j > 0, \sigma_j > 0,\, j = 1, ..., n.$$

The $\mathbb{R}^n$-valued subordinated process $\{\boldsymbol{Y}(t), t > 0\}$ defined by

$$\boldsymbol{Y}(t) = \begin{pmatrix} Y_1^I(t) + Y_1^\rho(t) \\ ... \\ Y_n^I(t) + Y_n^\rho(t) \end{pmatrix} = \begin{pmatrix} B_1(X_1(t)) + B_1^\rho(Z(t)) \\ .... \\ B_n(X_n(t)) + B_n^\rho(Z(t)) \end{pmatrix}, \tag{2.5}$$

where $X_j(t)$ and $Z(t)$ are independent subordinators, independent of $\boldsymbol{B}(t)$ and $\boldsymbol{B}^\rho(t)$ is a factor-based subordinated Brownian motion.

The return on asset j is obtained as a sum of an idiosyncratic and a systematic component

$$Y_j(t) = Y_j^I(t) + Y_j^\rho(t).$$

While the former components are independent, the latter are correlated. The parameters $\boldsymbol{\mu}^\rho$ and $\boldsymbol{\Sigma}^\rho$ are choosen so that the each marginal return j is a Brownian motion with parameters μ_j and σ_j subordinated by $G_j(t)$ as defined in (2.1). Indeed, the following equality in law $\mathcal{L}$ holds (see Theorem 5.1 in Luciano and Semeraro (2010b)).

$$\mathcal{L}(Y_j(t)) = \mathcal{L}(\mu_j G_j(t) + \sigma_j W(G_j(t)), \tag{2.6}$$

The marginal laws of $\boldsymbol{Y}(t)$ are therefore one-dimensional subordinated Brownian motions, and the parameters of $\boldsymbol{Y}(t)$ may be specified so to have VG, NIG and GH marginal distributions. Obviously, whenever all the parameters ρ_{ij} collapse to 0 across different assets, i.e. $\rho_{ij} = 0$ for $i \neq j$, $\rho_{ij} = 1$, for $i = j$, we have a version of the model which preserves marginal distributions and in which the Brownian motions are independent.

Applying Theorem 3.3 in Barndorff-Nielsen *et al.* (2001) and its univariate version (Theorem 30.1 in Sato (1999)) to $\boldsymbol{Y}^I(t) = (B_1(X_1(t)), ..., B_1(X_1(t)))$ and $\boldsymbol{Y}^\rho(t) = (B_1^\rho(Z(t)), ..., B_n^\rho(Z(t)))$, we find the characteristic function $\psi_{\boldsymbol{Y}(t)}$ of $\boldsymbol{Y}(t)$

$$\begin{aligned} \psi_{\boldsymbol{Y}(t)}(\boldsymbol{u}) &= \psi_{\boldsymbol{Y}^I(t)}(\boldsymbol{u})\psi_{\boldsymbol{Y}^\rho(t)}(\boldsymbol{u}) = \\ &= \exp(t\sum_{j=1}^{n} l_{X_j}(\log(\psi_{B_j}(u_j)))) \exp(t l_Z(\log(\psi_{\boldsymbol{B}^\rho}(\boldsymbol{u})))). \end{aligned} \tag{2.7}$$

where l_{X_j} and l_Z are the Laplace exponents of the subordinators.

3 Specifications

Before discussing the dependence properties of the process $Y(t)$, we show that it can be restricted so that it has VG, NIG and GH marginal distributions, as in Luciano and Semeraro (2010b) and Luciano and Semeraro (2010a). This is not required in order to show some dependence features of the general model, but simply to provide examples. The different specifications are obtained using different subordinators. To do that, using

the correspondence between Lévy processes and their law at time 1, it is sufficient to specify the distributions of the subordinators at time 1. Let

$$X_j := X_j(1),\ Z := Z(1) \text{ and } \ G_j := G_j(1). \tag{3.1}$$

For each specification, we also provide the characteristic function and the linear correlation coefficients.

3.1 Variance gamma marginal distributions

Recall that the VG univariate process, introduced by Madan and Seneta (1990), is a real-valued Lévy process $\{L_{VG}(t), t \geq 0\}$ which can be obtained as a Brownian motion subordinated by a gamma process $\{G(t), t \geq 0\}$. Let $\sigma > 0$ and μ be real parameters, then the process $L_{VG}(t)$ is defined as

$$L_{VG}(t) = \mu G(t) + \sigma B(G(t)),$$

where $B(t)$ is a standard Brownian motion. Its characteristic function at time 1 is

$$\psi_{VG}(u) = \left(1 - iu\mu\alpha + \frac{1}{2}\sigma^2\alpha u^2\right)^{-\frac{1}{\alpha}}. \tag{3.2}$$

To build a multivariate, factor-based version of the VG univariate process, specify $\boldsymbol{G}(t)$ to have gamma marginal distributions. Let X_j and Z be distributed according to gamma laws

$$\mathcal{L}(X_j) = \Gamma\left(\frac{1}{\alpha_j} - a, \frac{1}{\alpha_j}\right),\ j = 1, ..., n, \quad \text{and} \quad \mathcal{L}(Z) = \Gamma(a, 1).$$

To the marginal distributions to have non-negative parameters, the parameters α_j and a must satisfy the constraints

$$0 < \alpha_j < \tfrac{1}{a} \quad j = 1, ..., n. \tag{3.3}$$

In this case the random variables G_j defined in (3.1) are gamma-distributed too

$$\mathcal{L}(G_j) = \Gamma\left(\frac{1}{\alpha_j}, \frac{1}{\alpha_j}\right), \quad j = 1, ..., n.$$

and the subordinator $\boldsymbol{G}(t)$ defined in equation 2.1 has gamma marginal distributions at all times t

$$\mathcal{L}(G_j(t)) = \Gamma\left(\frac{t}{\alpha_j}, \frac{1}{\alpha_j}\right), \quad j = 1, ..., n.$$

It follows that the factor-based subordinated process $\boldsymbol{Y}$ defined in (2.5) has VG marginal processes with parameters $\mu_j, \alpha_j, \sigma_j$ - denoted as $VG(\mu_j, \alpha_j, \sigma_j)$ - i.e.

$$\mathcal{L}(Y_j(t)) = \mathcal{L}(\mu_j G_j(t) + \sigma_j W(G_j(t))). \tag{3.4}$$

The process $\boldsymbol{Y}(t)$ is named $\rho\alpha$ variance gamma, shortly $\rho\alpha$VG. The model has a total of $1+3n+\frac{n(n-1)}{2}$ parameters. It has one common parameter a, three marginal parameters $\mu_j, \alpha_j, \sigma_j$ for each marginal distribution, $j = 1, ...n$, and as many additional parameters as the distinct Brownian motions correlations ρ_{ij}, $i, j = 1, ..., n$.

The characteristic function of $\boldsymbol{Y}(t)$ at time 1 is

$$\psi_{\boldsymbol{Y}}(\boldsymbol{u}) = \prod_{j=1}^{n} \left(1 - \alpha_j(i\mu_j u_j - \frac{1}{2}\sigma_j^2 u_j^2)\right)^{-\left(\frac{1}{\alpha_j} - a\right)} \left(1 - \left(i\boldsymbol{u}^T\boldsymbol{\mu}^\rho - \frac{1}{2}\boldsymbol{u}^T\boldsymbol{\Sigma}^\rho\boldsymbol{u}\right)\right)^{-a}. \quad (3.5)$$

By imposing $\rho_{ij} = 0$, for $i \neq j$, $\rho_{ij} = 1$, for $i = j$, we find as a subcase of the current model the αVG process introduced in Semeraro (2008).

3.2 Normal inverse Gaussian marginal distributions

The univariate NIG process has been constructed by subordination by **?**. A NIG process with parameters $\gamma > 0$, $-\gamma < \beta < \gamma$, $\delta > 0$ is a Lévy process $\{L_{NIG}(t), t \geq 0\}$ with characteristic function at time 1

$$\psi_{NIG}(u) = \exp\left(-\delta\left(\sqrt{\gamma^2 - (\beta + iu)^2} - \sqrt{\gamma^2 - \beta^2}\right)\right).$$

It can be constructed by subordinating a Brownian motion with an inverse Gaussian distribution. An inverse Gaussian (IG) process with parameters (a, b) is a Lévy process with the following characteristic function

$$\psi_{IG}(u) = \exp\left(-a\left(\sqrt{-2iu + b^2} - b\right)\right).$$

To build a multivariate, factor-based version of the univariate process, consider that the subordinated process $\boldsymbol{Y}(t)$ has NIG marginal distributions, if we specify the subordinator $\boldsymbol{G}(t)$ in (2.1) to have IG marginal distributions, by defining

$$\begin{gathered} X_j \sim IG\left(1 - a\sqrt{\alpha_j}, \frac{1}{\sqrt{\alpha_j}}\right), \; j = 1, ..., n \\ Z \sim IG(a, 1). \end{gathered}$$

Using the closure properties of the IG distribution, we obtain that $\alpha_j Z \sim IG\left(a\sqrt{\alpha_j}, \frac{1}{\sqrt{\alpha_j}}\right)$ and that its sum with X_j is still IG

$$X_j + \alpha_j Z \sim IG\left(1, \frac{1}{\sqrt{\alpha_j}}\right).$$

The marginal distributions have non negative parameters if the following constraints are satisfied

$$0 < a < \frac{1}{\sqrt{\alpha_j}}, \quad j = 1, ..., n. \quad (3.6)$$

Let $\boldsymbol{G}(t)$ be as in (2.1), with X_j and Z as defined above. The marginal subordinator processes $G_j(t)$ are IG

$$\mathcal{L}(G_j(t)) = IG\left(t, \frac{1}{\sqrt{\alpha_j}}\right), \quad j = 1, ..., n.$$

Suitable constraints on the parameters make the subordinated process have NIG marginal distributions. Indeed, let $\gamma_j, \beta_j, \delta_j$ be such that

$$\gamma_j > 0, \quad -\gamma_j < \beta < \gamma_j, \quad \delta_j > 0.$$

Further, let

$$\frac{1}{\sqrt{\alpha_j}} = \delta_j \sqrt{\gamma_j^2 - \beta_j^2}. \tag{3.7}$$

If we set $\mu_j = \beta_j \delta_j^2$ and $\sigma_j = \delta_j$ in (2.4) the process $\boldsymbol{Y}$ defined in (2.5) has NIG marginal processes, i.e.

$$\mathcal{L}(Y_j(t)) = \mathcal{L}(\beta_j \delta_j^2 G_j(t) + \delta_j W(G_j(t))) \tag{3.8}$$

The process $\boldsymbol{Y}(t)$ is named $\rho\alpha$ normal inverse Gaussian, shortly $\rho\alpha$NIG. By imposing $\rho_{i,j} = 0$, for $i \neq j$, we have the αNIG model (Luciano and Semeraro (2010b)). Note that the process has a total of $1 + 3n + \frac{n(n-1)}{2}$ parameters. In particular a is a common parameter, $\gamma_j, \beta_j, \delta_j$, $j = 1, ..., n$ are marginal parameters and ρ_{ij} $(i > j)$ are the correlation coefficients between the Brownian components.

Its characteristic function at time one is

$$\psi_{\boldsymbol{Y}}(\boldsymbol{u}) = \exp\left\{ -\sum_{j=1}^{n} (1 - \frac{a}{\zeta_j}) \left(\sqrt{-2(i\beta_j \delta_j^2 u_j - \frac{1}{2}\delta_j^2 u_j^2) + \zeta_j^2} - \zeta_j \right) \right.$$
$$\left. -a \left(\sqrt{-2(i\boldsymbol{u}^T \boldsymbol{\mu}^\rho - \frac{1}{2}\boldsymbol{u}^T \boldsymbol{\Sigma}^\rho \boldsymbol{u}) + 1} - 1 \right) \right\}$$

where $\zeta_j = \delta_j \sqrt{\gamma_j^2 - \beta_j^2}$.

3.3 Generalized hyperbolic marginal distribution

Recall that the univariate generalized hyperbolic process is generated by the generalized hyperbolic distribution. The latter distribution, shortly GH$(\gamma, \beta, \delta, \lambda)$, has been introduced by Barndorff-Nielsen (1977) and has been first applied to finance by Eberlein and Keller (1995) and Eberlein and Prause (2002). Let $\lambda, \beta \in \mathbb{R}$, $\gamma, \delta \in \mathbb{R}_+$ satisfy the following constraints

$$\delta \geq 0, \ \ |\beta| < \gamma \ \text{ if } \ \lambda > 0$$
$$\delta > 0, \ \ |\beta| < \gamma \ \text{ if } \ \lambda = 0$$
$$\delta > 0, \ \ |\beta| \leq \gamma \ \text{ if } \ \lambda < 0$$

The characteristic function of a GH$(\gamma, \beta, \delta, \lambda)$ distribution is

$$\psi_{GH}(u) = \left(\frac{\gamma^2 - \beta^2}{\gamma^2 - (\beta + iu)^2}\right)^{\lambda/2} \frac{K_\lambda\left(\delta\sqrt{\gamma^2 - (\beta + iu)^2}\right)}{K_\lambda\left(\delta\sqrt{\gamma^2 - \beta^2}\right)}.$$

where $K_\lambda(x)$ denotes the modified Bessel function of the third kind with index λ. Generalized hyperbolic distributions have semi-heavy tails

$$f_{GH}(x; \gamma, \beta, \delta, \lambda) \sim |x|^{\lambda-1} \exp\left((\mp\gamma + \beta)\, x\right), \qquad \text{as } x \to \pm\infty \tag{3.9}$$

where $f_{GH}(x; \gamma, \beta, \delta, \lambda)$ is the density function of the GH$(\gamma, \beta, \delta, \lambda)$ distribution (see Prause (1999)). The univariate GH distribution in turn can be defined as a normal mean-variance mixture with a generalized inverse Gaussian (GIG) as mixing distribution; as a consequence the GH process can be constructed by time-changing a Brownian motion with a GIG subordinator. A GIG distribution with parameters $\lambda \in \mathbb{R}$, $a, b \in \mathbb{R}_+$, shortly $GIG(\lambda, a, b)$, is a three-parameter distribution defined on the positive half line. It is an infinitely divisible distribution and it generates a GIG subordinator. Its characteristic function is

$$\psi_{GIG}(u) = \frac{1}{K_\lambda(ab)} \left(1 - \frac{2iu}{b^2}\right)^{-\frac{\lambda}{2}} K_\lambda\left(ab\sqrt{1 - 2iub^{-2}}\right).$$

Last, if $G \sim GIG(\lambda, a, b)$ and W is standard normal, independent of G, then $\sqrt{G}W + \mu G$ has a GH distribution, with parameters $\gamma, \beta, \delta, \lambda$ where

$$\begin{aligned} a &= \delta \\ \mu &= \beta \\ b &= \sqrt{\gamma^2 - \beta^2}. \end{aligned}$$

To build a multivariate, factor-based version of the GH univariate process, i.e. to obtain GH marginal returns in the factor-based context, we specify a subordinator with marginal GIG laws. The main difficulty in the construction is that the GIG distribution is not closed under convolution. However, under a proper choice of the parameters, the convolution of a gamma and a GIG distribution is itself GIG distributed. We adopt the device of defining the subordinator by means of a gamma distributed common component. As we demonstrate below, this means that we do not recover as a limit case the multidimensional GH process analyzed by Eberlein and Prause (2002), and discussed in McNeil *et al.* (2010). If the idiosyncratic component degenerates instead we find the VG process with a common subordinator. The peculiarity of this model is then that it is a generalization of the multivariate VG model.

Let G_j be as in (2.1), where $X_j = R_j + V_j$, R_j and V_j are independent and

$$\begin{aligned} \mathcal{L}(R_j) &= GIG\left(-\lambda, \delta_j, \frac{1}{\sqrt{\alpha_j}}\right), \quad \mathcal{L}(V_j) = \Gamma\left(\lambda - a, \frac{1}{2\alpha_j}\right), \quad j = 1, ..., n, \\ \mathcal{L}(Z) &= \Gamma\left(a, \frac{1}{2}\right). \end{aligned}$$

If the parameters satisfy the following constraints

$$\lambda > 0, \alpha_j > 0, 0 < a < \lambda, \delta_j \geq 0 \tag{3.10}$$

the following equality in law holds (see Barndorff-Nielsen (1977))

$$\mathcal{L}(X_j + \alpha_j Z) = GIG\left(\lambda, \delta_j, \frac{1}{\sqrt{\alpha_j}}\right),$$

Notice that the law of X_j - the idiosyncratic component - is the convolution of a GIG and a gamma distribution.

If we set $\mu_j = \beta_j$ and $\sigma_j = 1$ in (2.4), the process $\boldsymbol{Y}$ defined in (2.5) has GH marginal distributions with parameters $\gamma_j, \beta_j, \delta_j, \lambda$, where

$$\frac{1}{\sqrt{\alpha_j}} = \sqrt{\gamma_j^2 - \beta_j^2}. \tag{3.11}$$

We name the above process $\rho\alpha$ generalized hyperbolic, shortly $\rho\alpha$GH. By imposing $\rho_{i,j} = 0$, for $i \neq j$, we have the αGH model (Luciano and Semeraro (2010a)). The process has a total of $2 + 3n + \frac{n(n-1)}{2}$ parameters, where λ and a are common parameters, $\gamma_j, \beta_j, \delta_j$, $j = 1, ..., n$, are marginal parameters and ρ_{ij} $(i > j)$ are the correlation coefficients between the Brownian components.

Because this construction holds for a subclass of generalized hyperbolic marginal distributions, namely, those with parameter $\lambda > 0$, the NIG marginal distribution which corresponds to $\lambda = -1/2$ is not a subcase of this construction. Note also that λ is the only common parameter of the marginal returns. In what follows, we consider the subcase of hyperbolic marginal distributions, which we call the $\rho\alpha$HYP process, introduced by Barndorff-Nielsen (1978), obtained by setting $\lambda = 1$. In this case the marginal returns do not have common parameters.

The $\rho\alpha$GH characteristic function at time 1 is

$$\begin{aligned}\psi_{\boldsymbol{Y}}(\boldsymbol{u}) &= \prod_{j=1}^{n}\left(1 - 2\left(\frac{\delta_j}{\zeta_j}\right)^2\left(iu_j\beta_j - \frac{1}{2}u_j^2\right)\right)^{-\frac{\lambda}{2}+a} \frac{K_\lambda\left(\zeta_j\sqrt{1 - 2\left(\frac{\delta_j}{\zeta_j}\right)^2\left(iu_j\beta_j - \frac{1}{2}u_j^2\right)}\right)}{K_\lambda(\zeta_j)} \\ &\quad \cdot \left(1 - 2\left(i\boldsymbol{u}^T\boldsymbol{\mu}^\rho - \frac{1}{2}\boldsymbol{u}^T\boldsymbol{\Sigma}^\rho\boldsymbol{u}\right)\right)^{-a}.\end{aligned}$$

where, again, $\zeta_j = \delta_j\sqrt{\gamma_j^2 - \beta_j^2}$.

4 Dependence

Factor-based processes introduced so far exhibit both linear and nonlinear dependence. The first one will be used in the calibration below, while the second points to the

dependence qualities that superimposing a Brownian motion on a subordinator produces. Since correlations between the components of $\boldsymbol{Y}(t)$ are independent of time (see Luciano and Semeraro (2010b)), we consider correlations at time 1. For each $\mathbb{R}^n$-valued process $\boldsymbol{Y}(t)$ or real valued process $Y(t)$, let $\boldsymbol{Y}$ and Y be the corresponding distributions at time 1.

4.1 Linear dependence

The correlations of the subordinator $\boldsymbol{G}(t)$ at time 1 are

$$\rho_{\boldsymbol{G}}(l,j) = \frac{\alpha_l \alpha_j V(Z)}{\sqrt{[V(X_l) + \alpha_l^2 V(Z)][V(X_j) + \alpha_j^2 V(Z)]}}$$

where V is the variance of the corresponding random variable. It follows that the linear correlation coefficients of $\boldsymbol{Y}$ are

$$\begin{aligned} \rho_{\boldsymbol{Y}}(i,j) &= \frac{Cov(B_i^\rho, B_j^\rho)E(Z) + E(B_i^\rho)E(B_j^\rho)V(Z)}{\sqrt{V(Y_i)V(Y_j)}} \\ &= \frac{\rho_{ij}\sigma_i\sigma_j\sqrt{\alpha_i}\sqrt{\alpha_j}E(Z) + \mu_i\mu_j\alpha_i\alpha_j V(Z)}{\sqrt{V(Y_i)V(Y_j)}}, \end{aligned} \tag{4.1}$$

where E is the expectation of the correponding random variable.

As expected, the correlations of $\boldsymbol{Y}$ depend on the first two moments of the common component, $E(Z)$ and $V(Z)$, as well as on the variance of the marginal, subordinated processes Y_i and Y_j, and both the drift and variance of the Brownian motions proper of assets i and j.

In terms of correlation flexibility, there are good news and bad news. The good news is that the correlations are neither bounded below nor above by the Brownian motion correlations ρ_{ij}. They can be greater than the Brownian correlations, $\rho_{\boldsymbol{Y}}(i,j) > \rho_{ij}$. Consider for instance the subcase $\rho_{ij} = 0$, in which the process $\boldsymbol{Y}$ obtains from independent Brownian motions. It has positively correlated margins if

$$\frac{\mu_i\mu_j\alpha_i\alpha_j V(Z)}{\sqrt{V(Y_i)V(Y_j)}} > 0$$

for all i and j. Correlations can also be lower then the Brownian correlations. For instance, $\rho_{\boldsymbol{Y}}(i,j) < 1$ when $\rho_{ij} = 1$, provided that

$$\frac{Cov(B_i^\rho, B_j^\rho)E(Z) + E(B_i^\rho)E(B_j^\rho)V(Z)}{\sqrt{V(Y_i)V(Y_j)}} < 1. \tag{4.2}$$

Another good news is that the factor-based model constructed here overcomes an empirically restrictive feature of previous models, in which Brownian motions were independent, **as the αVG model (Semeraro (2008)), the αNIG (Luciano and**

Semeraro (2010b)) and the αGH (Luciano and Semeraro (2010a)). For them, negative return correlations were possible only if the drifts of the Brownian motions had different signs, $\mu_i\mu_j < 0$. In practice, calibrations are unlikely to give this feature, and those models did not capture negative correlation. In the construction with dependent Brownian motions proposed here, $\boldsymbol{Y}$ may have negative correlations also if both drifts have the same sign: $\mu_i\mu_j > 0$. Indeed, we have $\rho_{\boldsymbol{Y}}(i,j) < 0$ iff

$$\frac{Cov(B_i^\rho, B_j^\rho)}{E(B_i^\rho)E(B_j^\rho)} < -\frac{V(Z)}{E(Z)}.$$

i.e. iff

$$\frac{\rho_{ij}\sigma_i\sigma_j\sqrt{\alpha_i}\sqrt{\alpha_j}}{\mu_i\mu_j\alpha_i\alpha_j} < -\frac{V(Z)}{E(Z)}.$$

If $\mu_i\mu_j > 0$, the previous inequality may hold provided that $\rho_{ij} < 0$. The introduction of the Brownian motion correlations therefore allows to overcome the condition $\mu_i\mu_j < 0$.

The bad news is that the correlation of the subordinated process $\boldsymbol{Y}$ is bounded above by the correlation of the subordinators, $\rho_{\boldsymbol{Y}}(i,j) \leq \rho_{\boldsymbol{G}}(i,j)$. Equality may hold in degenerate cases, i.e. when the idiosyncratic components of the subordinator degenerate. The equality $\rho_{\boldsymbol{Y}}(i,j) = \rho_{\boldsymbol{G}}(i,j)$ may also hold in non degenerate cases; Luciano and Semeraro (2010b) provide one example in that sense, which falls into the VG specification. At least in applications, such as the calibration below, this drawback does not seem to be too binding.

4.2 Linear correlations of Lévy subclasses

The $\rho\alpha$VG linear correlations are

$$\rho_{\boldsymbol{Y}}(i,j) = \frac{\left(\mu_i\alpha_i\mu_j\alpha_j + \rho_{ij}\sigma_i\sqrt{\alpha_i}\sigma_j\sqrt{\alpha_j}\right)}{\sqrt{(\sigma_i^2 + \mu_i^2\alpha_i)(\sigma_j^2 + \mu_j^2\alpha_j)}}a. \tag{4.3}$$

Under (3.3) , they must satisfy the constraint

$$0 < a < \min_j \left(\frac{1}{\alpha_j}\right).$$

By exploiting (3.7) , the linear correlations of the $\rho\alpha$NIG case are

$$\rho_{\boldsymbol{Y}}(i,j) = \frac{\left(\beta_i\frac{\delta_i^2}{\zeta_i^2}\beta_j\frac{\delta_j^2}{\zeta_j^2} + \rho_{ij}\frac{\delta_i}{\zeta_i}\frac{\delta_j}{\zeta_j}\right)}{\sqrt{\left(\gamma_i^2\delta_i\left(\gamma_i^2 - \beta_i^2\right)^{-\frac{3}{2}}\right)\left(\gamma_j^2\delta_j\left(\gamma_j^2 - \beta_j^2\right)^{-\frac{3}{2}}\right)}}a. \tag{4.4}$$

Under (3.6) and (3.7) , they must satisfy the constraint

$$0 < a < \min_j \zeta_j.$$

Enforcing (3.11), the linear correlation coefficients of the $\rho\alpha$GH case are

$$\rho_{\boldsymbol{Y}}(i,j) = \frac{\left(4\beta_i \frac{\delta_i^2}{\zeta_i^2}\beta_j \frac{\delta_j^2}{\zeta_j^2} + 2\rho_{ij}\frac{\delta_i}{\zeta_i}\frac{\delta_j}{\zeta_j}\right)}{\sqrt{V(Y_i)V(Y_j)}}a, \tag{4.5}$$

where the marginal variance is given by

$$V(Y_i) = \frac{\delta_i^2}{\zeta_i}\frac{K_{\lambda+1}(\zeta_i)}{K_\lambda(\zeta_i)} + \beta_i^2\frac{\delta_i^4}{\zeta_i^2}\left(\frac{K_{\lambda+2}(\zeta_i)}{K_\lambda(\zeta_i))} - \frac{K_{\lambda+1}^2(\zeta_i)}{K_\lambda^2(\zeta_i)}\right). \tag{4.6}$$

Under (3.10) , the common parameter a must satisfy the constraint

$$0 < a < \lambda. \tag{4.7}$$

In all the Lévy subclasses, the linear correlation coefficients are increasing in a.

The $\rho\alpha$ models introduce an additional term in the expression of the return correlations with respect to the α versions which can be recovered by letting all the ρ_{ij} equate zero in (4.3), (4.4) and (4.5). By so doing, they widen the range of achievable linear dependence. The common parameter a acts as a scaling factor for the overall correlation level. Constraints on parameter a are necessary to preserve gamma, IG, and GIG marginal distributions of the subordinator G, under the convolution between the idiosyncratic and common subordinators.

In the $\rho\alpha$VG and $\rho\alpha$NIG cases, the constraints turn out to be closely related to the maximum kurtosis of the marginal processes. Actually, the kurtosis increases with the parameter α_j in the $\rho\alpha$VG case and decreases with ζ_j in the NIG case. While low empirical return correlations may always be accounted for by choosing a sufficiently low a, high correlations may be challenging to achieve if there is at least one asset with very high kurtosis. For example, in the $\rho\alpha$VG, if $\alpha_j \to \infty$ for one j, $a \to 0$ and $\rho_{\boldsymbol{Y}} \to 0$. On the other hand, if all assets show a moderate level of kurtosis and the correlation range is not too large in absolute value, then the constraint is not likely to limit the model ability to match sample correlations within the $\rho\alpha$VG and $\rho\alpha$NIG subclasses. The additional parameter $\lambda > 0$ of the $\rho\alpha$GH provides more flexibility in selecting the admissible range for the common parameter a. For a given $\lambda > 0$, the constraint on the common parameter a for the $\rho\alpha$GH process is independent of the marginal parameters. We perform the empirical analysis in the $\rho\alpha$HYP case ($\lambda = 1$). By setting $\lambda > 1$, the upper bound on the correlation level in the $\rho\alpha$GH case can be made less binding. Since the parameter λ is linked to the tail behaviour of the GH$(\gamma_j, \beta_j, \delta_j, \lambda)$ distributions (see 3.9), a trade-off between fit of marginal distributions and fit of linear dependence can arise.

In principle, if the range of correlations is very wide, it might be difficult to match both very low and very high correlation coefficients, since high correlations may require the parameter a to be close to its upper bound, while low correlations may require it to be close to zero. Even setting $\rho_{ij} = -1$ or $\rho_{ij} = 1$ for some Brownian correlations

could not be sufficient to match all sample correlations. **However, provided the importance of VG, NIG and GH processes to model single asset returs, due to their distributional properties we assume here these convolution condition and show that, in practice, realistic ranges are well captured by calibrated coefficients.**

4.3 Nonlinear dependence

As noticed in Semeraro (2008) for the subcase with independent Brownian motions, the process $\boldsymbol{Y}(t)$ has also nonlinear dependence which derives from subordination. The subordinator $\boldsymbol{G}(t)$ has a simple factor structure as the Gaussian copula model proposed in Linders and Schoutens (2014) to model asset returns. While in Linders and Schoutens returns have an idiosyncratic and a common factor, in our model the subordinator has an idiosyncratic and a common component. The superposition of a multivariate Brownian motion on the common component of the subordinator generates nonlinear dependence which clearly appears if the Brownian motions have zero drifts and zero correlations. In fact, $\rho_{\boldsymbol{Y}}\ (i,j)$ becomes

$$\rho_{\boldsymbol{Y}}(i,j) = \frac{\rho_{ij}\sigma_i\sigma_j\sqrt{\alpha_i}\sqrt{\alpha_j}E[Z]}{\sqrt{V(Y_i)V(Y_j)}}.$$

In this case, if the Brownian motions are not correlated, we have $\rho_{\boldsymbol{Y}} = 0$, and the subordination generates only nonlinear dependence (see also Luciano and Semeraro (2010b)). We may also have maximal dependence and non maximal correlation. Take the $\rho\alpha$VG specification for simplicity. Let $\alpha_i = \alpha_j = \frac{1}{a}$ and $\rho_{ij} = 1$ for all $i, j = 1, ..., n$. We have

$$\boldsymbol{Y}(t) = \begin{pmatrix} \mu_1\frac{1}{a}Z(t) + \sigma_1\sqrt{\frac{1}{a}}W(Z(t)) \\ .... \\ \mu_n\frac{1}{a}Z(t) + \sigma_n\sqrt{\frac{1}{a}}W(Z(t)) \end{pmatrix},$$

Each marginal process is a deterministic transformation of the pair $(Z, W(Z))$ and - as such - has maximal dependence. However, substituting in the $\rho\alpha$VG correlations, if $\sigma_i \neq \sigma_j$ and $\mu_i \neq -\mu_j$, we get

$$\rho^2_{\boldsymbol{Y}}(i,j) = \frac{\sigma_i^2\sigma_j^2 + \frac{\mu_i^2\mu_j^2}{a} + \frac{2\sigma_i\sigma_j\mu_i\mu_j}{a^2}}{\sigma_i^2\sigma_j^2 + \frac{\sigma_j^2\mu_i^2}{a} + \frac{\sigma_i^2\mu_j^2}{a} + \frac{\mu_i^2\mu_j^2}{a^2}} < 1,$$

since $\frac{2\sigma_i\sigma_j\mu_i\mu_j}{a^2} < \frac{\sigma_j^2\mu_i^2}{a} + \frac{\sigma_i^2\mu_j^2}{a}$, except for $\sigma_i = \sigma_j$ and $\mu_i = -\mu_j$. This shows that it is possible to have simultaneously maximal dependence and non-maximal correlation.

5 An alternative approach: correlating marginal processes

Before calibrating the $\rho\alpha-$ models introduced above, we recall here the approach followed by Eberlein and Madan (2010) to correlate the unit-time random variables resulting from univariate subordinated Brownian motions which, together with Gaussian returns, we will use as a benchmark in the calibrations. Although they explored the VG specification, their model of dependence applies also to marginal processes of the NIG and GH types.

Let us model individual returns as one dimensional subordinated Brownian motions

$$Y_j(t) = \mu_j G_j(t) + \sigma_j W_j(G_j(t)), \;\; j = 1, ...n,$$

and assume that the subordinators are independent. Eberlein and Madan consider introducing dependence between returns at unit time by merely correlating the Brownian motions and keeping the subordinators independent. Therefore

$$\mathcal{L}(Y_j) = \mathcal{L}(\mu_j G_j + \sigma_j \sqrt{G_j} W_j), \;\; j = 1, ...n,$$

where W_j are standard normal variates with correlations ρ_{ij}^W. Under this model the correlation $\rho_{\boldsymbol{Y}}^{EM}(i,j)$ between returns becomes

$$\rho_{\boldsymbol{Y}}^{EM}(i,j) = \frac{cov(Y_i, Y_j)}{\sqrt{V(Y_i)V(Y_j)}} = \frac{\sigma_i \sigma_j E(\sqrt{G_i}) E(\sqrt{G_j}) \rho_{ij}^W}{\sqrt{V(Y_i)V(Y_j)}},$$

which implies the following correlations between the Brownian motions

$$\rho_{ij}^W = \frac{cov(Y_i, Y_j)}{\sigma_i \sigma_j E(\sqrt{G_i}) E(\sqrt{G_j})},$$

For the VG specification (VG-EM) $\sqrt{V(Y_i)V(Y_j)}$ is provided in equation (4.3) and (see Eberlein and Madan (2010))

$$E(\sqrt{G_i}) = \frac{\sqrt{\alpha_i}\Gamma((\frac{1}{\alpha_i}) + \frac{1}{2})}{\Gamma(\frac{1}{\alpha_i})}.$$

For the NIG specification (NIG-EM) $\sqrt{V(Y_i)V(Y_j)}$ is provided in equation (4.4) and

$$E(\sqrt{G_i}) = \int_{\mathbb{R}^+} \sqrt{x} \frac{1}{\sqrt{2\pi}} exp(\frac{1}{\sqrt{\alpha_i}}) x^{-3/2} exp(-\frac{1}{2}(x^{-1} + \frac{x}{\alpha_i})) dx.$$

For the HYP specification (HYP-EM) the marginal variance $V(Y_i)$ is given by (4.6) and

$$E(\sqrt{G_i}) = \int_{\mathbb{R}^+} \sqrt{x} \frac{\frac{1}{\sqrt{\alpha_i}\delta_i}}{2K_1(\frac{\delta_i}{\sqrt{\alpha_i}})} exp(-\frac{1}{2}(\delta_i^2 x^{-1} + \frac{x}{\alpha_i})) dx.$$

6 Calibration and fit

Define an $n-$dimensional price process, $\boldsymbol{S} = \{\boldsymbol{S}(t),\, t \geq 0\}$, by

$$\boldsymbol{S}(t) = \boldsymbol{S}(0)\exp(\boldsymbol{c}t + \boldsymbol{Y}(t)),\, \boldsymbol{c} \in \mathbb{R}^n, \tag{6.1}$$

where $\boldsymbol{c}$ is the drift term (equivalently, $S_j(t) = S_j(0)\exp(c_j t + Y_j(t))$, $t \geq 0, j = 1, 2, ..n$). The different specifications of the factor-based subordinated process $\boldsymbol{Y}(t)$ listed above can then be fitted to asset return data. We perform the calibration under the historical measure, since we want to assess how the previous models describe historically observed returns. The calibration is performed in two steps. The first consists in fitting the marginal parameters from marginal return data; the second in selecting the common parameters by matching the historical return correlation matrix. We calibrate the marginal return parameters by maximum likelihood (MLE), and then we calibrate the common parameters, i.e. the common subordinator parameter a and the correlation coefficients of the $\boldsymbol{B}^{\rho}$ component, by minimizing the distance between model and empirical correlations. We use as a distance the Frobenius norm. We include the Gaussian case (shortly, G), and the Eberlein and Madan models as benchmarks.

We are interested first in the marginal fit, as measured by the Kolmogorov-Smirnov and Anderson-Darling statistics, which is addressed in Section 6.2. We then explore the linear-correlation fit of the different $\rho\alpha-$ Lévy models in Section 6.3. We examine first the calibration fit when all assets are calibrated at the same time, then the results of pairwise calibration. In Section 6.4, with a construction to be explained below, we test the ability of the model to describe the sample distribution of portfolio returns as a way to investigate nonlinear dependence.

6.1 Data and summary statistics

We consider daily logreturns on MSCI US Investable Market Indices from January 2, 2009 to May 31st, 2013. Specifically, we look at the following sector indices: consumer discretionary (CD), consumer staples (CS), energy (EN), financials (FN), health care (HC), industrials (IN), information technology (IT), materials (MT), telecommunications (TC), and utilities (UT). We have 10 indices, with a total of 1109 observations.

The analysis of heterogeneous indices over the closer time window will allow us to provide a description of different market behaviours during a very turbulent period. Sample moments and correlations are reported in Table 1.

[Insert Table 1]

We observe that all means, referring to daily returns, are very low, skewnesses are all negative, except for the Financials Index, and kurtosis levels are significantly high. Because of the model properties explained in Section 4.2 above, matching market correlations appears to be particularly challenging with our dataset, since it includes assets

with high kurtosis together with high sample correlations. For instance, the Financials Index has a sample kurtosis equal to 11.946 and several correlation coefficients are above 0.8, with the correlation coefficient between the Consumer Discretionary Index and the Industrials Index as high as 0.933. Sample correlations range between 0.6503 and 0.933.

6.2 Calibration of the marginal distributions

We use MLE to estimate the marginal return distribution on each stock individually. The density function is recovered applying the Fractional Fast Fourier Transform (FRFT) (Chourdakis (2004)). Initial conditions are chosen according to the method of moments (see Seneta *et al.* (2004) and Prause (1999) for details.). The calibrated marginal parameters are reported in Table 2 for the different model specifications, including the G model. In all cases, the parameters in $\boldsymbol{c}$ are the drift terms in equation (4.2). Obviously, these parameters will apply both to our joint model and to the Eberlein and Madan one, since in both cases the two-step procedure is permitted.

[Insert Table 2]

The goodness of fit is evaluated by the Kolmogorov-Smirnov (KS) test, whose results are reported in Table 3. For all Lévy models and all indices, we cannot reject the null hypothesis that the sample comes from the model distribution, at the 5% level of significance. On the contrary, the G model provides a poor fit for all assets. Obviously, there are nuances across assets and models, in particular for assets with high kurtosis, such as the Financials Index.

[Insert Table 3]

Table 4 shows the Anderson and Darling (AD) distance between the model and the sample cumulative distribution functions. This statistics puts more weight on the tails than the Kolmogorov-Smirnov test. Again, the poor fit of the Gaussian distribution is reflected by the high values of the AD statistic, while all Lévy models provide a good fit. Overall, none of them seems to be superior to the others, considering the whole basket of assets. Differences in fit across models exist, with the NIG outperforming alternative specifications for assets with high kurtosis.

[Insert Table 4]

6.3 Calibration of the dependence structure

Given the marginal parameters, we jointly calibrate the common parameter a, and the correlation between the Brownian components, by fitting sample return correlations.

Specifically, we minimize the root-mean-squared error between the empirical and the $\rho\alpha$-model return correlations

$$RMSE\,(a,\rho) = \sqrt{\frac{2}{n\,(n-1)}\sum_{i=1}^{n}\sum_{j>i}\left(\rho_{\boldsymbol{Y}}^{emp}(i,j) - \rho_{\boldsymbol{Y}}(i,j,a,\rho_{ij})\right)^2}$$

where $\boldsymbol{\rho} = \{\rho_{ij},\, i = 1,\ldots n,\, j = 2,\ldots n\}$ are the correlation coefficients between the Brownian components collected in the $\boldsymbol{B}^{\rho}$, $\rho_{\boldsymbol{Y}}^{emp}(i,j)$ and $\rho_{\boldsymbol{Y}}(i,j)$ are the sample and model return correlations, respectively. Since all the marginal parameters are fixed from step 1, the correlation coefficients $\rho_{\boldsymbol{Y}}(i,j)$ depend on a and ρ_{ij} only, $\rho_{\boldsymbol{Y}}(i,j) = \rho_{\boldsymbol{Y}}(i,j,a,\rho_{ij})$.

In case cross-return correlations result in a matrix that is not symmetric positive semidefinite, a regularization procedure is applied. We follow the one proposed by Higham (2002), which allows to find the symmetric positive-semidefinite matrix which is closest to an arbitrary real matrix (in our case the one formed by the $\rho_{\boldsymbol{Y}}^{emp}(i,j)$) in the Frobenius norm. Note that this calibration step can be easily implemented. Moreover, the convergence of the algorithm is fast and robust with respect to the initial conditions.

We provide both joint and pairwise calibration results for the different Lévy $\rho\alpha-$ subclasses. As concerns the joint calibration, the upper bound of the common parameter a are $a = 0.77$ in the $\rho\alpha$VG case, $a = 0.25$ in the $\rho\alpha$NIG case and $a = 1$ in the $\rho\alpha$HYP case. Because the parameter a rescales the overall correlation level in the sample, the upper bound for the $\rho\alpha$NIG case turns out to be particularly restrictive, compared to the $\rho\alpha$VG and $\rho\alpha$HYP cases. We notice that the $\rho\alpha$NIG specification provides the best fit of marginal distributions for assets with high kurtosis, and it is high kurtosis that most affects the upper bound on a in the $\rho\alpha$VG and $\rho\alpha$NIG settings. On the contrary, the upper bound in the $\rho\alpha$HYP case is independent of the marginal parameters, and therefore is not affected by the marginal moments. The optimal value of the common parameter a is equal to its maximal value in all submodels for the 10 asset portfolio, which is consistent with high correlations and their range. We emphasize that this is not the case when subportfolios are considered. The intuition for this is again the behavior of the correlation submatrix and the range of its components. The calibrated correlation coefficients between the Brownian components for the $\rho\alpha-$ models are reported in Table 5. In the $\rho\alpha$NIG case, all correlations are equal to one. In all cases, we applied the Higham procedure (Higham (2002)) in order to enforce a positive-semidefinite matrix.

[Insert Table 5]

Table 6 provide the RMSE and the maximum absolute errors between the sample and model return correlations for both α and $\rho\alpha-$ models. They clearly show that moving from the $\alpha-$ models to the $\rho\alpha-$ models allows to mitigate the correlation problem in the VG and NIG cases, although in both models the constraint on the common parameter

a turns out to be binding and several correlation coefficients between the Brownian components are close to 1. In the HYP case, while the $\alpha-$ model cannot recover market correlations, the $\rho\alpha-$ model achieves an almost perfect matching across all pairs.

[Insert Table 6]

The calibration errors in matching correlations for the $\rho\alpha-$ models are shown in Table 7.

[Insert Table 7]

We remark that, in the $\rho\alpha$VG and $\rho\alpha$NIG cases, the constraint on a does not allow to accurately reproduce the sample correlations of our dataset. The restriction on the common parameter a limits the ability of these models to match high market correlations. On the contrary, the $\rho\alpha$HYP model proves to be able to reproduce even very high correlation coefficients. This happens, in the sample, because the idiosyncratic component has mean and variance close to zero, while the common component has not. This shows that the factor-based model captures high correlation, simply by playing on the relative importance of the two factors, without relying ex ante on a single subordinator.

Let us consider now the pairwise calibration which is quite common in the Lévy literature. We expect it to give a good performance also for those models that are relatively unable to fit joint distributions with all assets, both for analytical reasons and given our previous numerical explorations. Indeed, the $\rho\alpha$VG and $\rho\alpha$NIG models, though not able to generate the sample correlation in the joint calibration of this portfolio, since it is large and has critical features, still provide a good matching across most pairs, when the calibration is performed pairwise. Furthermore, a pairwise calibration confirms that, in the $\rho\alpha$VG and $\rho\alpha$NIG models, the only correlations that cannot be satisfactorily matched involve either assets with a very high kurtosis or very high correlation levels. Calibration errors under the pairwise optimization procedure are given in Table 8.

[Insert Table 8]

6.4 Portfolio returns and their fit

We compare the fit on portfolio returns of the factor-based multivariate model, the Gaussian and the Eberlein and Madan benchmark (extended to include also the NIG and HYP margins), since the return on a portfolio, being a weighted average of single-asset returns, incorporates all of their dependence, linear and nonlinear. The behavior of the factor-based models on portfolios therefore adds to the ability to capture linear dependence studied so far, providing information on their ability to capture other types of dependence, without resorting to any specific statistical measure (see Eberlein and Madan (2010) and Wang (2009)). To compare the performance of different models

in explaining portfolio returns we perform the Kolmogorov-Smirnov test, deriving the density function of model portfolio returns by FRFT inversion of their characteristic function. For each vector w of randomly selected portfolio weights, we compute the sample portfolio return from the whole time series data. The model portfolio daily linear return can be approximated[2] by the log-return

$$\boldsymbol{R}_P = \boldsymbol{w} \cdot (\boldsymbol{c} + \boldsymbol{Y}(1)),$$

and its characteristic function is

$$\psi_{\boldsymbol{R}_P}(u) = \exp(\boldsymbol{c} \cdot u\boldsymbol{w})\psi_{\boldsymbol{Y}}(u\boldsymbol{w}).$$

We consider long-short and long-only portfolios. Long-short portfolio weights are generated from independent normal variates, rescaled by the sum of their squares. For the long-only case, we take the absolute value of the standard normal variates and rescale them by their sum.

Since the characteristic function of the portfolio returns is not available in the Eberlein and Madan model, we rely on Monte Carlo simulation of 100,000 portfolio returns to derive their empirical cumulative distribution functions. For testing purposes, we also derive the cumulative distribution functions of portfolio returns for the $\rho\alpha-$ models by Monte Carlo simulation. The simulation procedure of our multivariate Lévy subclasses is straightforward, thanks to the time-change representation provided in (2.5). The extension to other Lévy margins, such as CGMY, may prove to be not particularly efficient (see, e.g., Ballotta and Kyriakou (2014)).

For each model and portfolio type, we compute the empirical complementary distribution $F(p)$ of the p-values across our sample of 1000 randomly generated portfolios, and we provide the proportion of portfolios with p-value greater than p. The higher is the complementary function, the better is the model ability to capture the laws of the randomly generated portfolios. Figure 1 shows the results obtained for all the specifications of $\rho\alpha-$ models and for the VG-EM specification. We do not plot all other EM specifications, since they are rejected at 1% level of significance. On the left we present long-only portfolios, while on the right we have long-short portfolios.

[Insert Figure 1]

We find that all Lévy multivariate models outperform the multivariate Gaussian model, since the $F(p)$ of the Gaussian is close to zero at almost all p-values, while the $F(p)$ of the Lévy models stays high. In terms of overall distribution fit, as expressed by the Kolmogorov distance, the $\rho\alpha$VG and $\rho\alpha$HYP subclasses provide a good performance on both long-only and long-short portfolios, the $\rho\alpha$HYP specification being superior to

[2]For longer horizon, this approximation is not accurate and linear portfolio returns should be considered.

$\rho\alpha$VG one for long-short portfolios. In this case, the VG specification of the Eberlein and Madan model shows a performance comparable to the $\rho\alpha$ VG specification. As noticed in Eberlein and Madan (2010), we find a departure of long-only portfolios from the model. All other specifications which indeed were not examined by Eberlein and Madan have p-values very close to zero, for both long-only and long-short portfolios.

Figure 2 (a) provides the Gaussian kernel density estimator of the Kolmogorov distance and Figure 2 (b) shows the Gaussian kernel density estimator of the Anderson and Darling statistics for the Lévy specifications. Figure 2 confirms the relatively good fit of all the Lévy models, in the KS sense. The $\rho\alpha$VG specification provides a slightly better description of sample portfolio returns for long-only portfolios, despite the fact that it is not able to match all sample correlations. An exception is the $\rho\alpha$NIG, consistently with the results of Figure 1. In terms of fit of the tails, as reflected by the Anderson and Darling statistics in Figure 2, the $\rho\alpha$HYP specification provides the best description of sample portfolio returns for both long-only and long-short portfolios.

[Insert Figure 2]

The three Figures above and the underlying tests show that, once evaluated on portfolios which reflect not only the linear dependence, but also the nonlinear one, the factor-based models are more flexible than a simple superposition of linear dependence on marginal distributions. Here the marginal distributions of the Eberlein and Madan specifications and the factor-model ones are indeed the same, with the same parameter values. The former construction, although appealing for its conceptual simplicity and similar to the factor-based construction in calibration, seems to be less prone to reproduce nonlinear dependence. Nonlinear dependence depends on the common component of the subordinator. Not surprisingly, the specifications of the factor-model in which the common component is stronger ($\rho\alpha$VG and $\rho\alpha$HYP) provide a better portfolio fit, both in comparison to the other factor-based construction ($\rho\alpha$NIG) and to the EM dependence structure. This happens on long-only portfolios. With long-short portfolios, the simple contruction of VG-EM gains in competitiveness.

Last, we explore the presence of a size effect in the ability of different models to explain portfolio returns by comparing the subportfolios of two (EN and FN) and five indices (EN, FN, HC, IN and IT) with the overall portfolio of 10 indices. Results are reported in Figure 3. Among the factor-based models, in our dataset we find a significant size effect only for the $\rho\alpha$NIG specification, which is most evident for long-only portfolios. We remark that the $\rho\alpha$NIG specification has the best marginal fit on the FN index, included in the two asset porfolio. Furthermore, it shows an improvement of the correlation fit with respect to the 10 assets case, although it has the highest absolute error if compared with the other submodels. Nevertheless, it provides the best fit for long-only portfolios.

[Insert Figure 3]

7 Summary and further research

This paper explores the dependence and fit properties of a class of multivariate processes named multivariate factor-based subordinated Brownian motions. The dependence structure applies to time-changed Brownian motions and permits us to stay in the Lévy class by choosing a subordinator as a change of time. The Lévy class can be specified so that it includes either variance gamma or normal inverse Gaussian and generalized hyperbolic marginal distributions, which are the specifications used here to investigate the dependence structure flexibility.

A joint - as opposed to pairwise - calibration to the level and linear correlation of a portfolio of ten US daily stock indices returns over the period 2009-2013 shows that several specifications of the model provide a strong improvement with respect to the traditional, Gaussian returns hypothesis. The fit of the Hyperbolic specification is very good, both in terms of marginal distributions and dependence. We then evaluate the performance of the joint calibrated factor-based models on a sample of randomly generated portfolios. We find that the calibrated factor-based models outperform the multivariate Gaussian model and the model by Eberlein and Madan which shows an adequate performance only for the $\rho\alpha$VG specification and long-short portfolios. In terms of overall distribution fit, as expressed by the Kolmogorov distance, our $\rho\alpha$VG and $\rho\alpha$HYP subclasses provide a good performance on both long-only and long-short portfolios, the $\rho\alpha$HYP specification being better than the $\rho\alpha$VG's one for long-short portfolios. In terms of fit of the tails, as reflected by the Anderson and Darling statistics, the $\rho\alpha$VG subclass provides a good description of sample portfolio returns, despite the fact that it doesn't match perfectly sample correlations.

Overall, the model - which clarifies and extends similar attempts in the same direction - seems to be flexible and general enough to incorporate empirical behavior of asset returns without computational efforts and keeping run times very limited. The model is able to capture also nonlinear dependence. Further research on the dependence structure, based on the computation of the implied copula and its behaviour over time, or on the higher moments of the joint distribution, is needed, even though it is out of the scope of the present paper. Equally out of the scope of this paper is the extension to the case of different marginal distributions.

Acknowledgments. *The authors thank Stefan Stöckl and two anonymous referees for providing valuable comments.*

References

, ????

Ballotta, L. and Kyriakou, I., Monte Carlo simulation of the CGMY process and option pricing. *Journal of Futures Markets*, 2014, **34**, 1095–1121.

Barndorff-Nielsen, O.E., Exponentially decreasing distributions for the logarithm of particle size. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 1977, **353**, 401–419.

Barndorff-Nielsen, O.E., Hyperbolic distributions and distributions on hyperbolae. *Scandinavian Journal of statistics*, 1978, pp. 151–157.

Barndorff-Nielsen, O.E., Normal inverse Gaussian distributions and the modeling of stock returns. Technical report, Research report, 1995.

Barndorff-Nielsen, O.E., Pedersen, J. and Sato, K., Multivariate subordination, self-decomposability and stability. *Advances in Applied Probability*, 2001, pp. 160–187.

Buchmann, B., Kaehler, B. and Maller, R., Multivariate subordination using generalised gamma convolutions with applications to V.G. processes and option pricing. Technical report, Presented at 8th World Congress of the Bachelier Finance Society, 2014.

Carr, P., Geman, H., Madan, D.B. and Yor, M., The Fine Structure of Asset Returns: An Empirical Investigation. *The Journal of Business*, 2002, **75**, 305–333.

Chourdakis, K., Option pricing using the fractional FFT. *Journal of Computational Finance*, 2004, **8**, 1–18.

Clark, P.K., A subordinated stochastic process model with finite variance for speculative prices. *Econometrica: Journal of the Econometric Society*, 1973, **41**, 135–155.

Eberlein, E. and Keller, U., Hyperbolic distributions in finance. *Bernoulli*, 1995, **1**, 281–299.

Eberlein, E. and Madan, D., On correlating Lévy processes. *The Journal of Risk*, 2010, **13**, 3–16.

Eberlein, E. and Prause, K., The generalized hyperbolic model: financial derivatives and risk measures. In *Proceedings of the Mathematical FinanceBachelier Congress 2000*, pp. 245–267, 2002.

Guillaume, F., Sato two-factor models for multivariate option pricing. *Journal of Computational Finance*, 2012, **15**, 159.

Harris, L., Cross-security tests of the mixture of distributions hypothesis. *Journal of financial and Quantitative Analysis*, 1986, **21**, 39–46.

Higham, N.J., Computing the nearest correlation matrixa problem from finance. *IMA Journal of Numerical Analysis*, 2002, **22**, 329–343.

Leoni, P. and Schoutens, W., Multivariate smiling. *Wilmott Magazine*, 2008, **8**, 82–91.

Linders, D. and Schoutens, W., Basket option pricing and implied correlation in a Lévy copula model. *Available at SSRN 2468778*, 2014.

Lo, A.W. and Wang, J., Trading volume: definitions, data analysis, and implications of portfolio theory. *Review of Financial Studies*, 2000, **13**, 257–300.

Loregian, A., Multivariate Lévy models: estimation and asset allocation. PhD thesis, Università degli Studi di Milano-Bicocca, 2013.

Luciano, E. and Schoutens, W., A multivariate jump-driven financial asset model. *Quantitative finance*, 2006, **6**, 385–402.

Luciano, E. and Semeraro, P., A generalized normal mean-variance mixture for return processes in finance. *International Journal of Theoretical and Applied Finance*, 2010a, **13**, 415–440.

Luciano, E. and Semeraro, P., Multivariate time changes for Lévy asset models: Characterization and calibration. *Journal of Computational and Applied Mathematics*, 2010b, **233**, 1937–1953.

Madan, D.B. and Seneta, E., The variance gamma (VG) model for share market returns. *Journal of business*, 1990, **63**, 511–524.

McNeil, A.J., Frey, R. and Embrechts, P., *Quantitative risk management: concepts, techniques, and tools*, 2010, Princeton University Press.

Prause, K., The generalized hyperbolic model: Estimation, financial derivatives, and risk measures. PhD thesis, University of Freiburg, 1999.

Sato, K.i., *Lévy processes and infinitely divisible distributions*, 1999, Cambridge University Press.

Semeraro, P., A multivariate variance gamma model for financial applications. *International Journal of Theoretical and Applied Finance*, 2008, **11**, 1–18.

Seneta, E. *et al.*, Fitting the variance-gamma model to financial data. *Journal of Applied Probability*, 2004, **41**, 177–187.

Wallmeier, M. and Diethelm, M., Multivariate downside risk: Normal versus variance gamma. *Journal of Futures Markets*, 2012, **32**, 431–458.

Wang, J., The multivariate variance gamma process and its applications in multi-asset option pricing. PhD thesis, University of Maryland, 2009.

MOMENTS										
	CD	CS	EN	FN	HC	IN	IT	MT	TC	UT
Mean	0.0009	0.0005	0.0004	0.0004	0.0006	0.0005	0.0007	0.0006	0.0003	0.0002
Variance	0.0002	0.0001	0.0003	0.0005	0.0001	0.0003	0.0002	0.0003	0.0001	0.0001
Skewness	-0.1663	-0.3873	-0.3566	0.1573	-0.4573	-0.2487	-0.0972	-0.3907	-0.1627	-0.3457
Kurtosis	6.0003	5.4783	5.4490	11.9466	6.5265	6.1286	5.4574	5.1771	6.2099	6.0961
CORRELATION MATRIX										
CD	1.0000									
CS	0.7985	1.0000								
EN	0.8214	0.7404	1.0000							
FN	0.8267	0.6781	0.7707	1.0000						
HC	0.7975	0.8338	0.7566	0.6985	1.0000					
IN	0.9337	0.7946	0.8690	0.8403	0.8089	1.0000				
IT	0.9052	0.7429	0.8244	0.7886	0.7728	0.8926	1.0000			
MT	0.8822	0.7462	0.8952	0.7989	0.7661	0.9165	0.8698	1.0000		
TC	0.7393	0.7391	0.6939	0.6921	0.6832	0.7352	0.7226	0.7012	1.0000	
UT	0.7247	0.7705	0.7457	0.6503	0.7343	0.7552	0.6863	0.7089	0.7022	1.0000

Table 1: Sample moments and sample correlation matrix of daily logreturns on MSCI US Investable Market Indices from January 2, 2009 to May 31st, 2013. 1109 observations. Sector indices are consumer discretionary (CD), consumer staples (CS), energy (EN), financials (FN), health care (HC), industrials (IN), information technology (IT), materials (MT), telecommunications (TC), and utilities (UT).

	G				$\rho\alpha$VG			
Index	c	σ			c	γ	β	δ
CD	0.0010	0.0146			0.0021	-0.0013	0.0144	0.9026
CS	0.0005	0.0084			0.0013	-0.0009	0.0083	0.7602
EN	0.0005	0.0168			0.0018	-0.0014	0.0166	0.7348
FN	0.0006	0.0224			0.0005	-0.0001	0.0206	1.2848
HC	0.0006	0.0106			0.0019	-0.0013	0.0105	0.8420
IN	0.0007	0.0158			0.0018	-0.0013	0.0156	0.9973
IT	0.0008	0.0139			0.0016	-0.0009	0.0139	0.9103
MT	0.0007	0.0172			0.0025	-0.0019	0.0170	0.7812
TC	0.0004	0.0113			0.0019	-0.0017	0.0111	0.7293
UT	0.0003	0.0100			0.0015	-0.0013	0.0099	0.6646
	$\rho\alpha$NIG				$\rho\alpha$HYP			
Index	c	γ	β	δ	c	γ	β	δ
CD	0.0020	51.7708	-5.0441	0.0112	0.0019	100.6824	-5.0872	0.0022
CS	0.0014	108.3392	-12.8277	0.0076	0.0014	180.4814	-13.5324	0.0024
EN	0.0021	54.9486	-6.0927	0.0155	0.0020	90.2830	-5.6782	0.0053
FN	0.0012	22.7119	-1.7045	0.0113	0.0014	71.5114	-2.6797	0.0000
HC	0.0021	82.5935	-13.7078	0.0090	0.0020	142.6616	-13.1937	0.0025
IN	0.0019	45.0711	-5.2494	0.0115	0.0020	92.0004	-5.8528	0.0012
IT	0.0015	57.3094	-4.3395	0.0114	0.0016	103.4022	-4.8038	0.0018
MT	0.0028	54.3748	-7.4708	0.0159	0.0026	86.9305	-6.8379	0.0045
TC	0.0018	81.6045	-12.0085	0.0101	0.0020	138.4737	-13.9333	0.0039
UT	0.0010	97.9514	-7.5590	0.0098	0.0012	157.8498	-9.4632	0.0043

Table 2: Maximum likelihood estimates of marginal return distributions for different model specifications: Gaussian (top-right table), $\rho\alpha$VG (top-left table), $\rho\alpha$NIG (bottom-right table), and $\rho\alpha$HYP (bottom-left table). MSCI US **Investable** Market Index from January 2, 2009 to May 31st, 2013.

Index	p-value				KS distance			
	G	$\rho\alpha$VG	$\rho\alpha$NIG	$\rho\alpha$HYP	G	$\rho\alpha$VG	$\rho\alpha$NIG	$\rho\alpha$HYP
CD	0.0000	0.8143	0.9509	0.8191	0.0815	0.0189	0.0154	0.0188
CS	0.0003	0.9961	0.9996	0.9999	0.0628	0.0122	0.0105	0.0093
EN	0.0001	0.9048	0.9378	0.9479	0.0682	0.0169	0.0159	0.0155
FN	0.0000	0.2074	0.8720	0.1055	0.1181	0.0314	0.0168	0.0354
HC	0.0000	0.5645	0.8765	0.6599	0.0842	0.0235	0.0176	0.0218
IN	0.0000	0.9179	0.9121	0.8729	0.0843	0.0165	0.0167	0.0177
IT	0.0000	0.9963	0.9844	0.9984	0.0731	0.0121	0.0136	0.0114
MT	0.0001	0.9635	0.9325	0.9866	0.0662	0.0149	0.0161	0.0134
TC	0.0006	0.9692	0.9969	0.9830	0.0602	0.0146	0.0119	0.0128
UT	0.0004	0.6943	0.6615	0.7422	0.0619	0.0203	0.0218	0.0203

Table 3: Kolmogorov-Smirnov test at the 5% level of significance for different model specifications: G, $\rho\alpha$VG, $\rho\alpha$NIG, and $\rho\alpha$HYP. The table on the right side shows the p-value. The critical value is 0.0406. The table on the left side exhibits the Kolmogorov-Smirnov (KS) distance between the model and the sample cumulative distribution function.

Index	G	$\rho\alpha$VG	$\rho\alpha$NIG	$\rho\alpha$HYP
CD	1.6075	0.0780	0.0428	0.0721
CS	0.6727	0.0597	0.0482	0.0495
EN	0.5621	0.0364	0.0355	0.0336
FN	132.4952	0.2194	0.0568	0.3901
HC	1.0028	0.0629	0.0507	0.0678
IN	1.9102	0.0693	0.0439	0.0733
IT	1.7122	0.0444	0.0447	0.0413
MT	0.6557	0.0332	0.0411	0.0299
TC	4.4743	0.1769	0.0933	0.1642
UT	1.5534	0.0498	0.0439	0.0409

Table 4: Anderson and Darling distance between the model and the sample cumulative distribution function for different model specifications: G, $\rho\alpha$VG, $\rho\alpha$NIG, and $\rho\alpha$HYP.

$\rho\alpha$VG										
	CD	CS	EN	FN	HC	IN	IT	MT	TC	UT
CD	1.0000									
CS	0.9866	1.0000								
EN	0.9999	0.9847	1.0000							
FN	0.9627	0.9055	0.9656	1.0000						
HC	0.9830	0.9998	0.9809	0.8967	1.0000					
IN	0.9989	0.9930	0.9984	0.9492	0.9904	1.0000				
IT	0.9969	0.9963	0.9960	0.9385	0.9944	0.9995	1.0000			
MT	0.9999	0.9847	1.0000	0.9656	0.9809	0.9984	0.9960	1.0000		
TC	0.9948	0.9981	0.9936	0.9301	0.9966	0.9984	0.9997	0.9936	1.0000	
UT	0.9915	0.9994	0.9901	0.9194	0.9985	0.9965	0.9987	0.9901	0.9996	1.0000
$\rho\alpha$HYP										
CD	1.0000									
CS	0.8611	1.0000								
EN	0.8942	0.8298	1.0000							
FN	0.8441	0.7156	0.8185	1.0000						
HC	0.8489	0.9165	0.8364	0.7279	1.0000					
IN	0.9615	0.8440	0.9320	0.8462	0.8477	1.0000				
IT	0.9392	0.7962	0.8898	0.8003	0.8179	0.9131	1.0000			
MT	0.9414	0.8235	0.9852	0.8379	0.8343	0.9646	0.9260	1.0000		
TC	0.8115	0.8365	0.7908	0.7445	0.7601	0.7940	0.7888	0.7860	1.0000	
UT	0.8127	0.8944	0.8684	0.7145	0.8406	0.8350	0.7657	0.8188	0.8290	1.0000

Table 5: Correlation matrices between the Brownian components for different model specifications: $\rho\alpha$VG, $\rho\alpha$NIG, and $\rho\alpha$HYP. Given the marginal parameters, the common parameter a, and the correlation between the Brownian components, are jointly calibrated, by fitting sample return correlations. MSCI US Investable Market Index from January 2, 2009 to May 31st, 2013.

Model	RMSE	MAE
αVG	1.0970	0.9287
αNIG	1.0941	0.9270
αHYP	1.0919	0.9275
$\rho\alpha$VG	0.2094	0.3056
$\rho\alpha$NIG	0.5645	0.5957
$\rho\alpha$HYP	0.0026	0.0107
$\rho\alpha$VG (pairwise)	-	0.0315
$\rho\alpha$NIG (pairwise)	-	0.2593
$\rho\alpha$HYP (pairwise)	-	0.0000

Table 6: Root mean square error (RMSE) and Maximum absolute error (MAE) between empirical and model asset correlations for each specifications: $\rho\alpha$VG, $\rho\alpha$NIG, $\rho\alpha$HYP. The MAE for pairwise calibration is also provided.

$\rho\alpha$VG										
	CD	CS	EN	FN	HC	IN	IT	MT	TC	UT
CD	0.0000									
CS	-0.1624	0.0000								
EN	-0.1876	-0.1676	0.0000							
FN	-0.0223	0.0159	-0.0421	0.0000						
HC	-0.1307	-0.2114	-0.1564	0.0232	0.0000					
IN	-0.1961	-0.1216	-0.2038	-0.0065	-0.1027	0.0000				
IT	-0.2020	-0.0980	-0.1904	0.0000	-0.0960	-0.1515	0.0000			
MT	-0.2288	-0.1555	-0.3056	-0.0490	-0.1468	-0.2307	-0.2163	0.0000		
TC	-0.1117	-0.1610	-0.1286	0.0034	-0.0754	-0.0731	-0.0899	-0.1177	0.0000	
UT	-0.1270	-0.2177	-0.2075	0.0076	-0.1529	-0.1239	-0.0822	-0.1536	-0.1606	0.0000
$\rho\alpha$NIG										
CD	0.0000									
CS	-0.4290	0.0000								
EN	-0.4588	-0.4334	0.0000							
FN	-0.1425	-0.0734	-0.1763	0.0000						
HC	-0.4089	-0.5026	-0.4312	-0.1174	0.0000					
IN	-0.4646	-0.3924	-0.4738	-0.1413	-0.3857	0.0000				
IT	-0.4902	-0.3913	-0.4790	-0.1657	-0.4039	-0.4537	0.0000			
MT	-0.5226	-0.4414	-0.5957	-0.2593	-0.4463	-0.5362	-0.5304	0.0000		
TC	-0.3714	-0.4268	-0.3871	-0.1407	-0.3554	-0.3462	-0.3754	-0.3958	0.0000	
UT	-0.3823	-0.4817	-0.4620	-0.1351	-0.4315	-0.3931	-0.3632	-0.4273	-0.4145	0.0000
$\rho\alpha$HYP										
CD	0.0000									
CS	0.0001	0.0000								
EN	0.0025	0.0002	0.0000							
FN	-0.0001	0.0000	-0.0003	0.0000						
HC	-0.0001	0.0000	-0.0002	0.0000	0.0000					
IN	0.0006	0.0000	0.0020	-0.0001	0.0000	0.0000				
IT	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000			
MT	-0.0031	-0.0002	-0.0107	0.0004	0.0002	-0.0025	-0.0001	0.0000		
TC	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	-0.0001	0.0000	
UT	-0.0006	0.0000	-0.0021	0.0001	0.0000	-0.0005	0.0000	0.0027	0.0000	0.0000

Table 7: Differences between model and sample correlations for different model specifications: $\rho\alpha$VG, $\rho\alpha$NIG, and $\rho\alpha$HYP.

$\rho\alpha$VG										
	CD	CS	EN	FN	HC	IN	IT	MT	TC	UT
CD	0.0000									
CS	0.0000	0.0000								
EN	0.0000	0.0000	0.0000							
FN	0.0000	0.0000	0.0162	0.0000						
HC	0.0000	0.0000	0.0000	0.0000	0.0000					
IN	0.0000	0.0000	0.0106	0.0000	0.0000	0.0000				
IT	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000			
MT	0.0000	0.0000	0.0000	0.0223	0.0000	0.0315	0.0000	0.0000		
TC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
UT	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$\rho\alpha$NIG										
CD	0.0000									
CS	0.0000	0.0000								
EN	0.0000	0.0000	0.0000							
FN	0.1425	0.0734	0.1763	0.0000						
HC	0.0000	0.0000	0.0000	0.1174	0.0000					
IN	0.0000	0.0000	0.0721	0.1413	0.0000	0.0000				
IT	0.0000	0.0000	0.0000	0.1657	0.0000	0.0075	0.0000			
MT	0.0674	0.0000	0.0000	0.2593	0.0000	0.1496	0.0033	0.0000		
TC	0.0000	0.0000	0.0000	0.1407	0.0000	0.0000	0.0000	0.0000	0.0000	
UT	0.0000	0.0000	0.0000	0.1351	0.0000	0.0250	0.0000	0.0000	0.0000	0.0000

Table 8: Pairwise calibration. Differences between model and sample correlations for the $\rho\alpha$VG and the $\rho\alpha$HYP models.

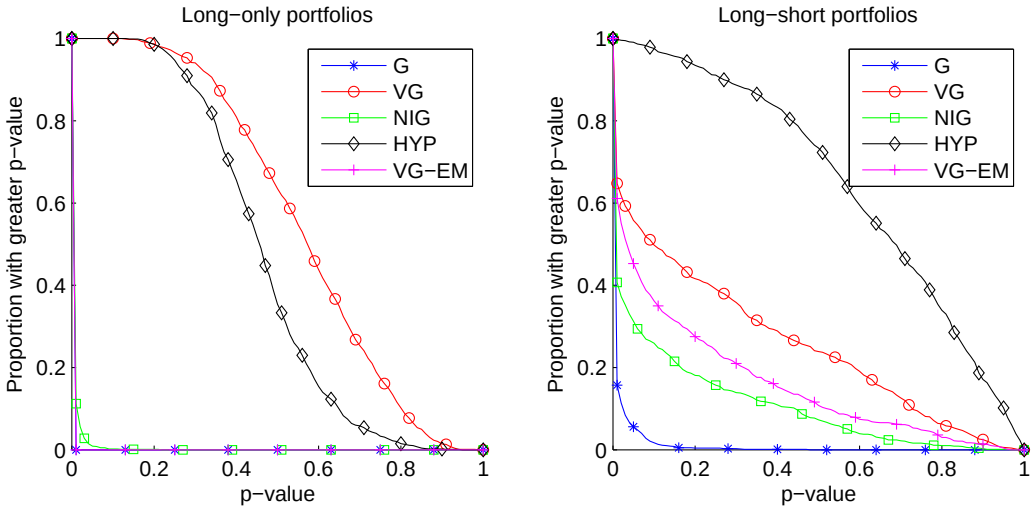
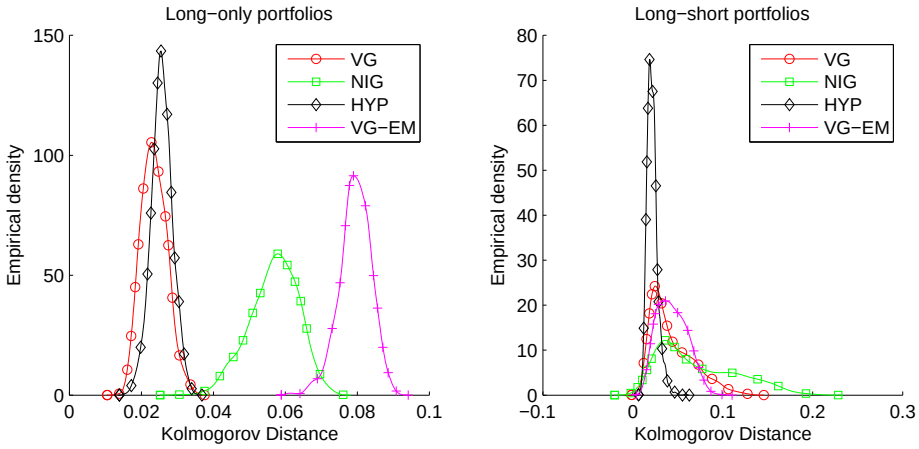
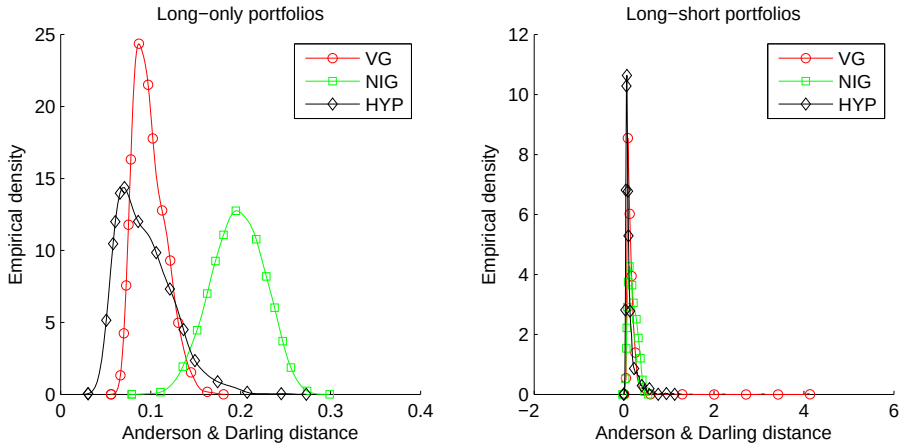


Figure 1: P-values complementary distribution function of the KS test for the following models: $\rho\alpha$VG, $\rho\alpha$NIG, $\rho\alpha$HYP, and VG-EM. For each model and portfolio type, the empirical complementary distribution $F(p)$ of the p-values across a sample of 1000 randomly generated portfolios is considered, and the proportion of portfolios with p-value greater than p is computed. The higher is the complementary function, the better is the model ability to capture the laws of the randomly generated portfolios.

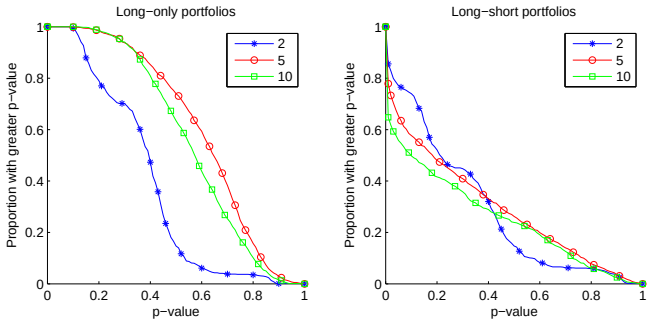


(a) Kolmogorov distance

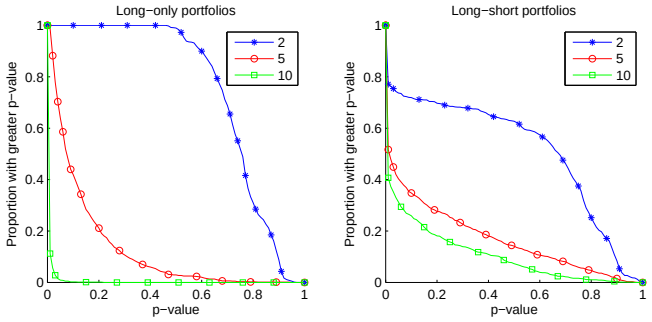


(b) Anderson and Darling distance

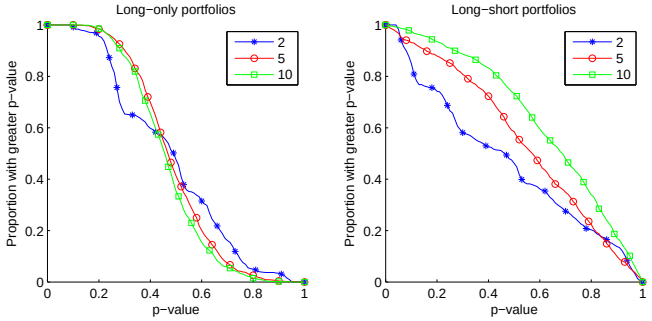
Figure 2: Gaussian kernel density estimator of the Kolmogorov distance (a) and the Anderson and Darling distance (b) for the following model specifications: $\rho\alpha$VG, $\rho\alpha$NIG, $\rho\alpha$HYP, and VG-EM.



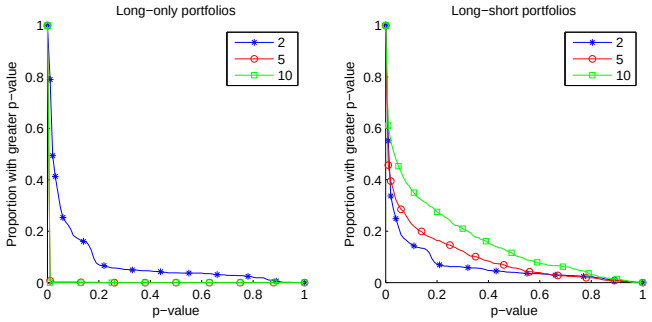
(a) $\rho\alpha$VG model



(b) $\rho\alpha$NIG model



(c) $\rho\alpha$HYP model



(d) VG-EM model

Figure 3: P-values complementary distribution function of the KS test for the following models: $\rho\alpha$VG, $\rho\alpha$NIG, $\rho\alpha$HYP, and VG-EM. Results refer to subportfolios of two (EN and FN), five (EN, FN, HC, IN and IT) and 10 indices (the overall portfolio). The Financial Index (FN) which has the highest kurtosis is always selected.