JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS Vol. 54, No. 5, Oct. 2019, pp. 2053–2083
COPYRIGHT 2018, MICHAEL G. FOSTER SCHOOL OF BUSINESS, UNIVERSITY OF WASHINGTON, SEATTLE, WA 98195
doi:10.1017/S0022109018001321

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

Estimation of Multivariate Asset Models with Jumps

Laura Ballotta, Gianluca Fusai, Angela Loregian, and M. Fabricio Perez*

Abstract

We propose a consistent and computationally efficient 2-step methodology for the estimation of multidimensional non-Gaussian asset models built using Lévy processes. The proposed framework allows for dependence between assets and different tail behaviors and jump structures for each asset. Our procedure can be applied to portfolios with a large number of assets because it is immune to estimation dimensionality problems. Simulations show good finite sample properties and significant efficiency gains. This method is especially relevant for risk management purposes such as, for example, the computation of portfolio Value at Risk and intra-horizon Value at Risk, as we show in detail in an empirical illustration.

I. Introduction

The importance of modeling financial assets under realistic distributional assumptions away from normality has been highlighted in particular after the subprime financial crisis. Realistic modeling is especially relevant for risk management, given the extreme price movements, event risk, and sudden and large trading losses observed in financial data. Nonnormality also directly affects the returns' tail distribution, which is crucial in the computation of the regulatory

*Ballotta, l.ballotta@city.ac.uk, City, University of London Cass Business School; Fusai, gianluca.fusai@uniupo.it, gianluca.fusai.1@city.ac.uk, Università del Piemonte Orientale Dipartimento SEI, City, University of London Cass Business School; Loregian, angela.loregian@arpm.co, Advanced Risk and Portfolio Management (ARPM); and Perez (corresponding author), mperez@wlu.ca, Wilfrid Laurier University Lazaridis School of Business and Economics. We thank Gurdip Bakshi (the referee), Cecilia Mancini, Roberto Renò, Lorenzo Trapani, and David Veredas for useful comments and suggestions. Perez acknowledges the financial support received by the Arthur Wesley Downe Professorship of Finance and the Social Sciences and Humanities Research Council of Canada. Fusai acknowledges the financial support received by Università del Piemonte Orientale. This article has been presented at the Arizona State University Economics Reunion Conference. A previous version of this article was circulated with the title "Multivariate Lévy Models by Linear Combination: Estimation" and has been presented at the 2014 International Conference of the Financial Engineering and Banking Society (FEBS). We thank all the participants for their helpful feedback. The usual caveat applies.

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

capital requirements of financial institutions. Lévy processes offer a natural and robust approach for incorporating different distributional assumptions by means of discontinuous movements (commonly described as jumps), which can accommodate the levels of skewness and excess kurtosis observed in financial data, in particular over short horizons (see, e.g., Aït-Sahalia (2004)).

Although Lévy processes offer agile distribution modeling for asset prices, they also present significant estimation challenges, especially in a multivariate setup; therefore, univariate models are generally used for portfolio analysis. For example, drawing on the flexible properties of the class of Lévy processes, Bakshi and Panayotov (2010) analyze portfolio tail-risk measures on the basis of a pure-jump Lévy model for portfolio returns. The Bakshi and Panayotov (2010) setup is univariate in the sense that it models portfolio returns, rather than each individual asset in the given portfolio, and it captures the main portfolio characteristics while retaining tractability and computational efficiency. However, their univariate setting allows for neither the analysis of the impact of dependence between the components of the portfolio nor the measurement of each asset's individual risk contribution to the portfolio, which is particularly relevant for scenario analysis. A multivariate model, which caters to different tail behaviors for each asset in the portfolio, is ideal not only for risk management but also for portfolio optimization and the analysis of multivariate structures such as basket options on equities and collateralized debt obligations.

In light of the aforementioned, in this article we adopt a general multivariate setting for Lévy processes, which accounts for the impact of dependence between the components of the portfolio, and propose a consistent and computationally efficient model estimation procedure, suitable for portfolio risk measurement and management. After showing the theoretical validity of our method and testing its finite-sample properties by simulations, we showcase its applicability to the computation of the risk measures Value at Risk (VaR) and intra-horizon Value at Risk (VaR-I).

Our estimation methodology is developed by combining the latest advances in the modeling of multivariate Lévy processes with the most recent developments in the estimation of latent-factor models. Specifically, we adopt the multivariate construction of Ballotta and Bonfiglioli (2016) that models a portfolio of asset returns as a linear combination of independent Lévy processes representing systematic and idiosyncratic components. Because neither the common nor the idiosyncratic components driving the margin processes are directly observable, maximum-likelihood (ML) estimation of the parameters is computationally burdensome and often unfeasible. Indeed, estimation of the parameters via a single maximization of the likelihood function presents significant issues in terms of implementation, particularly the curse of dimensionality due to the large number of parameters necessary to accommodate a multivariate model.

Thus, the first contribution of this article is to propose a consistent and computationally efficient 2-step estimation procedure for the multivariate model of Ballotta and Bonfiglioli (2016), which overcomes the aforementioned dimensionality problems by means of the principal-components method of Bai and Ng (2002) and Bai (2003). Specifically, principal-component estimation is employed in step 1 to consistently estimate the common risk process; then in step 2, we

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

focus on the estimation of the parameters of the idiosyncratic components. In both steps, the estimation of the Lévy process parameters is based on the maximization of univariate sample likelihood functions. Hence, our procedure not only simplifies estimation by improving computational efficiency but also solves the dimensionality problem while providing consistent estimation of the parameters. Our simulation study confirms the reliability and computational efficiency of our method in comparison with the (likely unfeasible) 1-step ML approach in which all parameters of the multivariate Lévy process are estimated in a single step.

As a second contribution of this article, we show how our methodology can be applied to the computation of portfolio risk measures. Similar to Bakshi and Panayotov (2010), specific attention is paid to VaR-I, which captures the exposure to losses throughout the investment horizon, contrasting with VaR, which is the industry standard for the estimation of regulatory capital requirements and measures the risk of possible losses at the end of a predetermined time horizon. Because VaR-I incorporates information about the dynamic path of possible losses, it offers an ideal tool for dealing with intra-horizon risk over a multiperiod investment horizon (see, e.g., Stulz (1996), Kritzman and Rich (2002), and Boudoukh, Richardson, Stanton, and Whitelaw (2004)). This is of paramount importance for monitored asset managers, leveraged investors, borrowers required to maintain a particular level of reserves as a condition of a loan agreement, or securities lenders required to deposit collateral. Our work, however, moves beyond that of Bakshi and Panayotov (2010) to a multivariate setting, which incorporates the impact of dependence between the components of the portfolio. The computation of the relevant risk measures in the multivariate setting is facilitated by the fact that the chosen factor construction gives immediate access to the portfolio's characteristic function; hence, we can use state-of-the-art numerical procedures required for the computation of VaR-I, such as the Fourier space time-stepping (FST) algorithm introduced by Jackson, Jaimungal, and Surkov (2008) for pricing path-dependent financial option contracts. Our methods avoid the implementation of numerical methods for partial integro-differential equations such as the ones used by Bakshi and Panayotov (2010), which might require approximations, especially for infinite activity jumps, that may lead to accuracy and stability problems (see Jackson et al. (2008) for further details). As an illustration, we provide a clear estimation procedure for portfolio VaR-I assuming a multivariate model following the normal inverse Gaussian process of Barndorff-Nielsen (1997) and the Merton jump-diffusion processes of Merton (1976).

The third contribution of this article is to show that the proposed framework also allows for the explicit computation of the risk contribution from each asset in the portfolio without the need for reestimating the multivariate model. This is of relevance, for example, in the active risk management of a portfolio because, by means of the proposed multivariate model, it is possible to assess changes in the portfolio risk profile resulting from one additional unit position of exposure to a given asset. This breakdown of the contribution to risk represents an invaluable "drill-down" exercise that enables managers to better control their risk profiles. We illustrate the application using the portfolio VaR-I previously mentioned.

The outline of the article is as follows: In Section II, we review the most relevant features of the multivariate Lévy model under consideration, and we discuss the estimation of the model, introducing the 2-step estimation procedure. In Section III, we assess the 2-step estimation procedure via simulations. Section IV illustrates how to compute the VaR-I for a portfolio of assets following the proposed model, with an application in real data. Section V concludes. Detailed simulation results are presented in Appendix A, and Appendix B offers a brief review of the literature related to the present article.

II. Multivariate Lévy Processes by Linear Combination: Model and Estimation

A. Model Specification

A Lévy process, L_t, on a filtered probability space is a continuous-time process with independent and stationary increments, whose distribution is infinitely divisible. Lévy processes have attracted attention in the financial literature due to the fact that they accommodate distributions with nonzero higher moments (skewness and excess kurtosis), therefore allowing a more realistic representation of the stylized features of market quantities such as asset returns. Further, they represent a class of processes with known characteristic functions in virtue of the Lévy–Khintchine representation, so that $\mathbb{E}(\exp(iuL_t)) = \exp(t\varphi(u)), u \in \mathbb{R}$, with $\varphi(\cdot)$ denoting the so-called characteristic exponent. This feature in particular allows for the development of efficient numerical schemes for the approximation of potentially unknown distribution functions and derivative prices using Fourier inversion techniques.

In this setting, let us denote by P_t the price of a financial instrument represented as

$$P_t \quad = \quad P_0 \exp(L_t);$$

assuming that we observe the price process on an equally spaced time grid $t = 1, 2, \ldots, T$, the log-return defined as

$$X_t \quad = \quad \log\left(\frac{P_t}{P_{t-1}}\right) \quad = \quad L_t - L_{t-1}$$

is a process with infinitely divisible distribution such that, for all t, X_t is distributed as L_1, and $X_1, X_2, \ldots, X_T$ are mutually independent. Hence, with a slight abuse of notation, we say that X_t is an independent and identically distributed (IID) process.

A convenient construction for an N-dimensional version of the Lévy process L_t is proposed by Ballotta and Bonfiglioli (2016) via a linear transformation of a vector of independent Lévy processes with components $\tilde{L}_t^{(n)}$, $n = 1, 2, \ldots, N$, each representing the idiosyncratic risk, and another independent Lévy process, $\tilde{L}_t^{(N+1)}$, modeling the common risk component, so that for $n = 1, 2, \ldots, N$, $L_t^{(n)} = \tilde{L}_t^{(n)} + a_n \tilde{L}_t^{(N+1)}$, $a_n \in \mathbb{R}$. Due to the property of independent and stationary increments of Lévy processes, the increments also respect the same linear transformation. In particular, let us denote by $Y_t^{(n)}, n = 1, \ldots, N$, Z_t the increments of $\tilde{L}_t^{(n)}$, $n = 1, 2, \ldots, N$, $\tilde{L}_t^{(N+1)}$, respectively. Then the following holds:

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

Proposition 1. Let $Z, Y^{(n)}, n=1,\dots,N$ be IID processes, with characteristic functions $\phi_Z(u;t)$ and $\phi_{Y^{(n)}}(u;t)$, for $n=1,\dots,N$, respectively. Then, for $a_n \in \mathbb{R}$, $n=1,\dots,N$

$$(1) \qquad \boldsymbol{X}_t = (X_t^{(1)},\dots,X_t^{(N)})' = (Y_t^{(1)}+a_1 Z_t,\dots,Y_t^{(N)}+a_N Z_t)'$$

is an IID process on $\mathbb{R}^N$ with the characteristic function

$$\phi_X(\boldsymbol{u};t) = \phi_Z\left(\sum_{n=1}^N a_n u_n;t\right)\prod_{n=1}^N \phi_{Y^{(n)}}(u_n;t), \quad \boldsymbol{u}\in\mathbb{R}^N.$$

It follows by conditioning on the systematic process, Z, that the joint probability density function of the multivariate IID process $\boldsymbol{X}_t$ is

$$(2) \qquad f_X(x_t^{(1)},\dots,x_t^{(N)}) = \int_{-\infty}^{\infty} f_{Y^{(1)}}(x_t^{(1)}-a_1 z) \times\cdots\times f_{Y^{(N)}}(x_t^{(N)}-a_N z) f_Z(z)\,dz.$$

We note that because the given multivariate model admits computable characteristic function, the joint distribution is always available (up to a Fourier inversion), even when the components' distributions, $f_{Y^{(1)}},\dots,f_{Y^{(N)}}, f_Z$, are not known analytically.

Proposition 1 implies that for each $X^{(n)}$, $n=1,\dots,N$, the process Z captures the systematic part of the risk originated by sudden changes affecting the whole market, whereas the process $Y^{(n)}$ represents the idiosyncratic shocks generated by company-specific issues. Consequently, the components of $\boldsymbol{X}_t$ are dependent and may jump together. In particular, for each $t\geq 0$, the components of $\boldsymbol{X}_t$ are positively associated if the loading factors a_n for $n=1,\dots,N$ are all either positive or negative; otherwise, the components of $\boldsymbol{X}_t$ are negative quadrant dependent. The resulting pairwise linear correlation coefficient is

$$(3) \qquad \rho^X_{n,m} = \operatorname{corr}(X_t^{(n)}, X_t^{(m)}) = \frac{a_n a_m \operatorname{var}(Z_1)}{\sqrt{\operatorname{var}(X_1^{(n)})}\sqrt{\operatorname{var}(X_1^{(m)})}},$$

which is well defined if all processes have finite moments of all orders (specifically the variance). We note, in fact, from expression (3) that for fixed $a_n, a_m \neq 0$, $\rho^X_{n,m}=0$ if and only if Z is degenerate and the components are independent, whereas $|\rho^X_{n,m}|=1$ if and only if $Y^{(n)}$ and $Y^{(m)}$ are degenerate (i.e., there is no idiosyncratic factor in the components $X^{(n)}$ and $X^{(m)}$). Further, $\operatorname{sign}(\rho^X_{n,m}) = \operatorname{sign}(a_n a_m)$; therefore, both positive and negative correlations can be accommodated. Finally, the resulting multivariate model shows nonzero indices of upper and lower tail dependence, which are controlled by the tail probabilities of the systematic risk process. For fuller details, refer to Ballotta and Bonfiglioli (2016) and Ballotta, Deelstra, and Rayée (2017).

Several features of the construction in Proposition 1 are worth noticing. In the first place, this construction is relatively parsimonious in terms of the number of parameters involved because this number grows linearly with the number of assets.

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

Further, the adopted modeling approach is quite flexible because it can be applied to any Lévy process; indeed, Proposition 1 allows for specifying any univariate Lévy process for the idiosyncratic and systematic risks. In this respect, we note that, differently from Ballotta and Bonfiglioli (2016), in this work we do not impose any convolution condition on the components aimed at recovering a known distribution for the margin processes, hence allowing for a more realistic portrayal of the asset log-return features and the dependence structure in place. Because factor models generally do not originate known marginal distributions (except in the Gaussian case), a large portion of the literature on multivariate Lévy processes focuses on finding suitable conditions for the model parameters under which this feature holds. However, as argued by several authors, such as Eberlein, Frey, and von Hammerstein (2008), recovering known distributions for the margin processes, although intuitive, leads to a biased view of the dependence structure in place. This is because it reduces the flexibility of the factor model and fails to recognize the different tail behaviors of the assets in the portfolio, which is an essential aspect in risk management, especially when it comes to the assessment of the marginal risk contribution of each individual asset in the portfolio. As observed by Ballotta et al. (2017) in a different context, the factor construction of Ballotta and Bonfiglioli (2016) retains its mathematical tractability and parsimonious parameter space regardless of the presence of these conditions.

Finally, the model is particularly tractable because the full description of the multivariate vector $\mathbf{X}_t$ only requires information on the univariate processes $Y_t^{(1)}, \ldots, Y_t^{(N)}$ and Z_t. However, from a practical point of view, these sources of risk are not directly observable.

Thus, for the purpose of the estimation of the given multivariate Lévy model, discussed in the following section, we distinguish between 1-step and 2-step approaches, which depend on the estimation methods used for the common factor. The 1-step approach involves joint estimation of the parameters of the common factor and the idiosyncratic components; however, the maximization of the resulting likelihood function is feasible only if we consider a limited number of assets. The 2-step approach that we propose instead involves, firstly, the estimation of latent factors and loadings by principal-component methods; conditioned on this information, the likelihood function admits a simple expression as a product of univariate densities. These facts simplify the estimation procedure to improve efficiency and solve the dimensionality problem while providing consistent estimation of the parameters. We observe that this approach is facilitated by the lack of convolution conditions imposed on the model parameters, as discussed previously. An alternative possibility would be to consider the unobservable common factor as a latent factor whose dynamic is assigned so that the estimation procedure can be reduced to a (in general) non-Gaussian Kalman filtering problem. However, the application of these techniques is in general not straightforward and, in any case, does not solve the dimensionality problem.

We conclude by observing that in order to simplify the description of the model but without loss of generality, we assume that Z includes only 1 factor; however, all results in this article can be generalized to a multifactor model, as we explain in the following sections.

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

B. Model Estimation: A 2-Step Approach

From the joint density of the stock log-returns given by equation (2), it follows that the likelihood function of the sample $\boldsymbol{x} = \left\{(x_t^{(1)}, \ldots, x_t^{(N)})\right\}_{t=1,\ldots,T}$ is

$$
\begin{aligned}
(4) \qquad L(\boldsymbol{x}, \theta) \;=\; & \prod_{t=1}^{T} \int_{-\infty}^{\infty} f_{Y^{(1)}}(x_t^{(1)} - a_1 z_t; \theta_{Y^{(1)}}) \\
& \times \cdots \times f_{Y^{(N)}}(x_t^{(N)} - a_N z_t; \theta_{Y^{(N)}}) f_Z(z_t; \theta_Z) dz_t,
\end{aligned}
$$

where $\theta = [\theta_{Y^{(1)}}, \ldots, \theta_{Y^{(N)}}, \theta_Z, \boldsymbol{a}]$ is the parameter set to be estimated.

Thus, all parameters of the chosen multivariate Lévy model can be estimated via a single maximization of the likelihood function in equation (4). However, we note that this procedure presents significant issues in terms of implementation, in particular, the curse of dimensionality. This is caused by several elements: the dimension of the parameter space, due to a richer model parametrization; the number of assets N, which increases the complexity of the integrand function; the sample size T, which increases the number of integrals to be evaluated; and, in the case of extensions to multifactor models, the number of common factors, which increases the dimension of the integral in equation (4). In addition, in the case of non-Gaussian dynamics, the density functions might not be known in closed forms and therefore have to be computed numerically. All these issues only exacerbate the numerical optimization, leading to imprecise parameter estimates and cases of false convergence. Finally, we note that in the case in which more systematic factors are assumed, there would be an infinite set of possible orthogonal factors, and the ML equations would have an infinite number of solutions returning the same value of the likelihood. This fundamental indeterminacy, called a problem of rotation, is discussed by Anderson (1957).

A valid alternative to an estimation procedure based on equation (4), which improves on the implementation issues just mentioned, exploits the independence of the common factor and the idiosyncratic processes. Indeed, the factor model in equation (1), by virtue of standard results on the joint probability distribution of functions of random variables, gives access to a joint density for the pair $(\mathbf{X} = \mathbf{Y} + \mathbf{a}Z, Z)$ of the form

$$
f_Z(z_t) \prod_{n=1}^{N} f_{Y^{(n)}}\left(x_t^{(n)} - a_n z_t\right).
$$

This result, under the assumption of Z being observable (in a sense to be defined later), allows us to conveniently write the log-likelihood function of the sample $\boldsymbol{x}$ and $\boldsymbol{z} = \{z_t\}_{t=1,\ldots,T}$ as

$$
(5) \quad \ln L(\boldsymbol{x}, \boldsymbol{z}; \boldsymbol{\theta}_Z, \boldsymbol{\theta}_Y, \boldsymbol{a}) = \sum_{t=1}^{T} \ln f_Z(z_t; \boldsymbol{\theta}_Z) + \sum_{n=1}^{N} \sum_{t=1}^{T} \ln f_{Y^{(n)}}\left(x_t^{(n)} - a_n z_t; \boldsymbol{\theta}_{Y^{(n)}}\right),
$$

with the convenient feature of expression (5) being the separation of the log-likelihood of the systematic risk process from the log-likelihoods of the idiosyncratic components.

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

This leads to the following remark: If the systematic risk factor is observable, the additive structure of the log-likelihood function highlighted by expression (5) suggests that the optimization procedure for the model estimation can be performed in 2 steps, one for the systematic risk process and one for the idiosyncratic components. More specifically, the first step is represented by the following optimization with respect to the parameters of the observable systematic process Z:

$$\max_{\boldsymbol{\theta}_Z} \ln L\left(\boldsymbol{z};\boldsymbol{\theta}_Z\right) \quad = \quad \max_{\boldsymbol{\theta}_Z} \sum_{t=1}^{T} \ln f_Z\left(z_t;\boldsymbol{\theta}_Z\right). \tag{6}$$

Given θ_Z, the second step consists of N independent maximizations, one for each instrument, of the likelihood of the idiosyncratic components with respect to the loading coefficients and the parameters of the idiosyncratic processes, which can be formalized as

$$\max_{\boldsymbol{\theta}_{Y^{(n)}},a_n} \ln L\left(x^{(n)} - a_n \boldsymbol{z};\boldsymbol{\theta}_{Y^{(n)}}\right) \quad = \quad \max_{\boldsymbol{\theta}_{Y^{(n)}},a_n} \sum_{t=1}^{T} \ln f_{Y^{(n)}}\left(x_t^{(n)} - a_n z_t;\boldsymbol{\theta}_{Y^{(n)}}\right), \quad n=1,\ldots,N. \tag{7}$$

We notice that this estimation strategy, "observe, divide, and conquer," allows us to solve the curse of dimensionality because each maximization procedure involves only a subsection of the overall parameter space. In addition, generalizing the model for multiple factors has the minimal additional cost of solving more independent maximization problems. We also emphasize that once the common factor can be considered observable, our 2-step procedure is still based on the maximization of the likelihood function, and therefore, the estimator retains all theoretical limiting properties, such as consistency, asymptotic normality, and efficiency.

In practice, though, the common factor is not directly observable. One way to proceed is to use observable proxy variables such as a well-diversified index; however, these proxy variables are latent variables contaminated with an error that does not vanish, causing the estimation to lose all its theoretical limiting properties. To solve this problem, we propose an alternative approach based on the latest theoretical advances on factor models, as in Bai and Ng (2002) and Bai (2003).

Allowing for a multifactor structure, our model in equation (1) can be written as

$$X_t^{(n)} \quad = \quad a_n' Z_t + Y_t^{(n)}.$$

Hence, in terms of standard factor model notation, $X_t^{(n)}$ is the response variable, a_n is the $1 \times r$ vector of factor loadings specific to the cross-sectional unit n, Z_t is the $r \times 1$ vector of common factors, and $Y_t^{(n)}$ is the idiosyncratic error. In matrix notation, $\boldsymbol{X} = \boldsymbol{Z}\boldsymbol{a}' + \boldsymbol{Y}$, where the matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$ are $T \times N$, $\boldsymbol{Z}$ is $T \times r$, and a is $N \times r$. Equation (1) can be recovered from this more general setting in the case of $r = 1$.

Bai and Ng (2002), (2008) and Bai (2003) propose a principal-components method to consistently estimate factors, loadings, and the number of factors by

solving the following optimization problem:

$$\min_{a,Z}(NT)^{-1}\sum_{t=1}^{T}\sum_{n=1}^{N}\left(X_t^{(n)}-a_n'Z_t\right)^2,$$

subject to the normalization $a'a/N=\mathbb{I}_N$, where $\mathbb{I}_N$ is the N-dimensional identity matrix. The resulting (optimal) estimated loadings matrix, $\widetilde{a}$, is $\sqrt{N}$ times the eigenvectors associated with the r largest eigenvalues of the $N\times N$ matrix $\boldsymbol{X}'\boldsymbol{X}$. Given $\widetilde{a}$, the factors can be estimated by $\widetilde{\boldsymbol{Z}}=\boldsymbol{X}\widetilde{\boldsymbol{a}}/N$.

Five main assumptions are required for consistent estimation of the factors (see Bai and Ng (2002), (2008), Bai (2003) for further details):

Assumption 1. $\mathrm{E}\|Z_t\|^8\le M\le\infty$ and $(\sum_{t=1}^{T}Z_tZ_t')/T\xrightarrow{p}\Sigma_Z>0$, an $r\times r$ nonrandom matrix.

Assumption 2. The loading a_n is either deterministic or stochastic with $\mathrm{E}\|a_n\|^8\le M\le\infty$. In either case, $(\sum_{n=1}^{N}a_na_n')/N\xrightarrow{p}\Sigma_a>0$, an $r\times r$ nonrandom matrix as $N\to\infty$.

Assumption 3. $Y_t^{(n)}$ is weakly correlated, both over time and cross-sectionally.

Assumption 4. $\{a_n\}$, $\{Z_t\}$, and $\{Y_t^{(n)}\}$ are mutually independent. Dependence within each group is allowed.

Assumption 5. $\mathrm{E}\|N^{-\frac{1}{2}}\sum_{n=1}^{N}a_nY_t^{(n)}\|^8\le M\le\infty$ for all t.

In particular we note the following: According to Assumption 1, the factors Z_t are allowed to be non-IID, meaning some level of autocorrelation is permitted; by definition, X_t is an IID process, and thus Z_t is also an IID process, and the assumptions holds. Assumption 2 rules out nonsignificant factors, or factors with a trivial contribution to the variance of the response variables; factor loadings are assumed to be nonrandom variables, which is the case in our model. Assumption 3 allows for limited time-series and cross-sectional dependence in the idiosyncratic components $Y_t^{(n)}$ and also heteroscedasticity in both the time and cross-sectional dimensions (see Bai and Ng (2002) for formal details). This assumption is irrelevant in our model because $Y_t^{(n)}$ are also IID processes. However, it is important to note that in general, some level of cross-sectional correlation is allowed because under standard approximate-factor models, common factors may not fully capture the total systematic variation on the response variables. Assuming that $Y_t^{(n)}$ are IID processes implies that all systematic variations are captured by the common factors (exact-factor model); however, our factor estimation allows for some level of correlation left in the residuals, which may be relevant when dealing with real data. Assumption 4 is a standard assumption in factor-analysis models and holds in our case by definition in Proposition 1. Assumption 5 allows some weak correlation between idiosyncratic errors and loadings and guarantees that the estimated factors uniformly converge to the space spanned by Z_t.

Under these assumptions, Bai and Ng (2002) show that estimated factors and loadings are consistent when $N,T\to\infty$. By means of Monte Carlo simulations,

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

Bai (2003) shows also that the estimator has good finite-sample properties for values of N as small[1] as 25.

Moreover, Bai and Ng (2008) show that the estimated factors $\widetilde{\boldsymbol{Z}}$ can be treated as observed variables in extremum estimation as ML estimation under suitable assumptions.[2] Thus, equations (6) and (7) can be rewritten as functions of estimated factors $\widetilde{\boldsymbol{Z}}$ and estimated loadings $\widetilde{\boldsymbol{a}}$, and the parameters $\boldsymbol{\theta}_Z$ and $\boldsymbol{\theta}_{Y^{(n)}}$ can be consistently estimated by ML estimation.

Two additional comments are worth mentioning for the factor estimation. First, in the case of multifactor models, the estimation of factors and loadings also requires estimation of the number of factors. Bai and Ng (2002) propose an estimation method based on classical model-selection methods. More recently, Ahn and Horenstein (2013) propose an estimation method based on the ratio of the eigenvalues of the matrix $\boldsymbol{X}'\boldsymbol{X}/NT$, according to which, for eigv_k denoting the kth-largest eigenvalue of the matrix $X'X/NT$, the number of factors is the value of k that maximizes the ratio criterion function

$$ER(k) = \frac{\text{eigv}_k}{\text{eigv}_{k+1}}.$$

The method is very simple and has very good small-sample properties; hence, we adopt it for our empirical application in Section IV.[3]

Second, as in any factor model with latent factors and loadings, the true factors and loadings can only be identified up to a scale. Specifically, the estimated factors $\widetilde{\boldsymbol{Z}}$ are consistent estimators of $\boldsymbol{ZH}$, where $\boldsymbol{Z}$ represents the true factors and $\boldsymbol{H}$ is an invertible rotation matrix. This is known as the factor-rotation problem, and it affects the interpretation of factors and loadings. However, our model in equation (1) does not require factors and loadings to be separately identified but only requires the identification of the components $\boldsymbol{Za}'$, which are, in fact, identifiable.

As mentioned previously, the second step consists of N separate likelihood-maximization problems, as in equation (7). Because the loadings are estimated jointly with the factors, the objective function will only require estimation of the parameters $\theta_{Y^{(n)}}$.

Finally, a practitioner's note: A proxy factor could be used instead of the estimated principal-component factor. Although, as explained earlier, this method does not guarantee estimation consistency, it could nevertheless represent a practical alternative. Indeed, the optimization problem in equation (7) can in this case be solved iteratively by maximizing first with respect to the idiosyncratic parameters and then with respect to the loading parameters until no further significant improvement in the objective functions is achieved. In particular, the loadings can be constrained to fit the covariance matrix to correctly recover the dependence

[1] In the simulations performed by Bai (2003), the average correlation between the true factor and the estimated factor for $N = 25, T = 50$ over 2,000 repetitions is 0.98, improving as T increases.

[2] Additional technical assumptions required to prove the consistency of estimated parameters are in most part also required for ML estimation with observed variables; for technical details, refer to Bai and Ng (2008).

[3] For cases with small N or T, the number of factors can be estimated by the methods of Ahn and Perez (2010).

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

structure, as described in the following discussion, and then the maximization of the likelihood in equation (7) is performed only with respect to the idiosyncratic parameters.

Specifically, the vector $\boldsymbol{a}$ can be initialized by fitting the nondiagonal entries of the sample covariance matrix to their theoretical counterparts predicted by the multivariate model in equation (1). This is achieved by solving

$$\min_{\boldsymbol{a}} \|\mathrm{cov}(\boldsymbol{X}) - \boldsymbol{\sigma}^2\|_F, \tag{8}$$

where $\|\cdot\|_F$ denotes the Frobenius norm,

$$\mathrm{cov}(\boldsymbol{X}) = \boldsymbol{a}\boldsymbol{a}'\mathrm{var}(Z_1) + \mathrm{diag}([\mathrm{var}(Y^{(1)}), \ldots, \mathrm{var}(Y^{(N)})]) \tag{9}$$

is the model covariance matrix (see equation (3) as well), and $\boldsymbol{\sigma}^2$ denotes the sample covariance matrix (we set the diagonal entries equal to 0 in both). In expression (9), we can use either the sample variance of the increments of the proxy variable for Z or the parametric expression for the variance computed with the parameters estimated in step 1; in the former case, this step turns out to be independent of the specification of the Lévy processes involved in the multivariate model construction.

III. Estimation Assessment

In this section, we evaluate the performance of the 2-step estimation approach presented in Section II.B by simulations. Our objectives are to assess the efficiency gains of the 2-step approach in comparison with the 1-step approach and also to analyze the finite-sample properties of the 2-step estimator. We focus on two particular specifications of the multivariate model in equation (1): the case in which all the involved processes come from the normal inverse Gaussian (NIG) process of Barndorff-Nielsen (1997) with drift (all-NIG) and the case in which all the involved processes are generated by Merton jump-diffusion processes (MJD) of Merton (1976) (all-MJD). These jump structures were also considered by Ornthanalai (2014). The features of these processes are reviewed in the following discussion of this section. At this stage we note that all required densities are generated via Fourier numerical inversion of the corresponding characteristic functions. The numerical inversion has been performed adopting the COS method introduced by Fang and Oosterlee (2008) by virtue of its high numerical accuracy.

The NIG process, introduced by Barndorff-Nielsen (1997), is a normal-tempered stable process obtained by subordinating an arithmetic Brownian motion with an (unbiased) independent inverse Gaussian process. Its characteristic function reads

$$\phi(u;t) = \exp\left(i\mu t + \frac{t}{k}\left(1 - \sqrt{1 - 2iu\theta k + u^2\sigma^2 k}\right)\right), \quad u \in \mathbb{R}, \tag{10}$$

for $\mu, \theta \in \mathbb{R}$ and $\sigma, k > 0$.

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

It follows by differentiation of the (log of the) characteristic function that the first four cumulants of the NIG process are

$$\begin{aligned} c_1 &= (\mu+\theta)t, & c_2 &= \left(\sigma^2+\theta^2 k\right)t, \\ c_3 &= 3\theta k\left(\sigma^2+\theta^2 k\right)t, & c_4 &= 3k\left(\sigma^4+6\sigma^2\theta^2 k+5\theta^4 k^2\right)t. \end{aligned}$$

From this, we observe that θ primarily controls the sign of the skewness of the process distribution, σ affects the overall variability, and k primarily controls the kurtosis of the distribution. The drift parameter μ affects the mean of the distribution, which otherwise would be concordant with the skewness, allowing us to model return distributions with positive means and negative skewness as well (and vice versa). Finally, the tails of the distribution are characterized by a power-modified exponential decay, or semi-heavy tail (see, e.g., Cont and Tankov (2004)).

Because the density function is known in (semi-)closed form (as it is expressed in terms of the modified Bessel function of the second kind, see, e.g., Cont and Tankov (2004)), the parameters of the NIG model can be estimated directly using ML estimation, initialized via the method of moments based on the first four theoretical cumulants derived previously.

A Lévy jump-diffusion process has the form

$$\mu t+\sigma W_t+\sum_{i=1}^{N_t} J_i, \tag{11}$$

where W_t is a standard Brownian motion, N_t is a Poisson process with rate $\lambda>0$ counting the jumps of the overall process, and $\{J_i\}_{i\in\mathbb{N}}$ are IID random variables describing the jump sizes. All the random quantities involved, W_t, N_t, and J_i (for all i), are assumed to be mutually independent. In the MJD model (Merton (1976)), jump sizes are all normally distributed (i.e., $J_i\sim N(\nu,\tau^2)$, $\nu\in\mathbb{R}$, $\tau>0$, for all i). It follows that the characteristic function is

$$\phi(u;t) = \exp\left(iu\mu t-\frac{u^2\sigma^2}{2}t+\lambda t\left(e^{iu\nu-(\tau^2u^2)/2}-1\right)\right), \quad u\in\mathbb{R}. \tag{12}$$

The first four cumulants are

$$\begin{aligned} c_1 &= (\mu+\lambda\nu)t, & c_2 &= (\sigma^2+\lambda(\nu^2+\tau^2))t, \\ c_3 &= \lambda\nu(3\tau^2+\nu^2)t, & c_4 &= \lambda(3\tau^4+6\tau^2\nu^2+\nu^4)t. \end{aligned}$$

We can observe that the parameters λ, ν, and τ control the non-Gaussian part of the process; in particular, ν primarily controls the sign of skewness (the density function is symmetric when $\nu=0$), whereas λ governs the jump frequency and, therefore, the level of excess kurtosis. Further, the MJD process has an infinite Gaussian-mixture distribution with mixing coefficients given by a Poisson distribution with parameter λ; hence, the probability density function can be expressed as a fast-converging series. Finally, the tails are heavier than in the pure Gaussian case (see, e.g., Cont and Tankov (2004)).

We note that the estimation of the MJD model is far from trivial because the ML method requires a careful numerical optimization, as discussed by

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

Honoré (1998). Consequently, in the numerical study, we implement the expectation maximization (EM) algorithm in the formulation proposed by Duncan, Randal, and Thomson (2009), which has simple closed-form solutions for the M-step.

We conclude this review by highlighting the main difference between the NIG and the MJD processes. Although they both cater to small movements occurring with a high frequency (i.e., they are infinite-variation processes), in the MJD process, these movements are generated by the Brownian motion and, therefore, are Gaussian (skewness and kurtosis are generated by the "big" jumps controlled by the compound Poisson process). In the NIG process, instead, these small movements are purely discontinuous, and therefore, their distribution is already skewed and leptokurtic.

Using the all-NIG and all-MJD processes described previously, we generate data for our simulations. The first objective of our simulation is to evaluate the computation-efficiency gains of the 2-step approach versus the 1-step approach. Our second objective is to assess the error, bias, and efficiency of the parameter estimates obtained by adopting the proposed 2-step approach. We perform simulation experiments for different data-generating processes, sample sizes, and numbers of assets in the portfolio. The simulations show that the 2-step approach has very large computation-efficiency gains relative to the 1-step approach. Our most conservative test shows that after controlling for errors, parameter estimation using the 2-step approach is more than 3,000 times faster than using the 1-step method. Moreover, the 2-step estimation has good finite-sample properties with low bias and root-mean-squared errors (RMSEs) that decrease with increased sample size. The number of assets included in the portfolio has a minimal impact on the estimation errors, and the 2-step approach provides reliable estimates even for samples with more than 15 assets in the portfolio, a case for which the 1-step approach becomes imprecise and often noncomputable. Detailed simulation procedures and results, including tables, are presented in Appendix A.

IV. Application to Portfolio Risk Measures: VaR and VaR-I

In this section, we illustrate our estimation method for the computation of portfolio risk measures such as VaR and VaR-I.

The trading portfolios of financial institutions can be adversely exposed to a multitude of risk factors that may lead to extreme losses. Therefore, asset managers are required to maintain a particular level of reserves as protection against these trading losses. VaR, defined as the lower-tail percentile for the distribution of returns, is the market risk measure recommended by U.S. and international banking regulators to estimate the minimum capital requirements. For example, the Basel Capital Accord amended in 1996 established that the minimum capital requirement on a given day is equal to the sum of a charge to cover credit risk and a charge to cover general market risk, where the market-risk charge is equal to a multiple of the average reported 2-week VaRs in the last 60 trading days. A drawback of VaR estimates is that they do not take into consideration the magnitude of possible losses incurred before the end of the specified trading horizon. An improved alternative risk measure is the VaR-I because it takes into account the exposure to losses throughout the investment's life of the portfolio.

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

Estimation of tail-risk measures such as VaR and VaR-I crucially depends on modeling financial assets under realistic distributional assumptions. Market participants and regulators are likely to be concerned about the effects of event risk and sudden large trading losses or jumps. Therefore, the use of the traditional Brownian-motion framework based on normality will likely understate such market risk. Lévy processes offer a natural and robust approach to incorporate jumps that can accommodate the levels of skewness and excess kurtosis observed in financial data, in particular over short horizons.

Moreover, asset managers will be concerned not only about the risk measures of the whole portfolio but also about the risk contribution of each asset in the portfolio. For example, an active portfolio manager will be interested in evaluating the effect on the portfolio risk profile of a change in the portfolio weight of a given asset. This breakdown of the contribution to risk represents an invaluable "drill-down" exercise that enables managers to better control their risk profiles. However, it requires a multivariate model that is, on the one hand, capable of incorporating the impact of dependence between the assets in the portfolio and, on the other hand, is sufficiently flexible to cater to different returns' distributions for each asset in the portfolio, a feat that lies out of the range of a univariate setting.

Therefore, the methods developed in this article are particularly suitable for the estimation of portfolio VaR and VaR-I because they accommodate realistic distributional assumptions, including jumps, in a multivariate setting that allows for the evaluation of the risk contribution of each individual asset in the portfolio.

In this section, we first provide a step-by-step general procedure to estimate VaR and VaR-I for portfolios following a multivariate Lévy model (equation (1)). Then we apply the proposed estimation method to a portfolio of the 20 most capitalized stocks in the Standard & Poor's (S&P) 500 index, using the two Lévy models' specifications introduced in Section III (i.e., the all-NIG model and the all-MJD model). For comparison, we also consider the case in which all assets follow a normal distribution ("all-Gaussian" model). After computing the model parameters and relevant risk measures, we assess the quality of our estimation and conclude the section with the identification of the risk contribution of each asset in the portfolio.

The results presented in the following discussion of this section show that the multivariate Lévy models correctly capture the observed distribution of portfolio returns, with important improvements over the Gaussian model. Estimation under multivariate Lévy models in general provide the most conservative risk estimates, with a VaR and VaR-I between 3.5% and 10% larger than using a Gaussian model. Finally, we identify that in our sample, the assets with larger risk contribution to the portfolio risk are stocks of financial institutions.

A. Estimation of VaR and VaR-I under Multivariate Lévy Models

The VaR-I is defined on the distribution of the minimum return. Thus, let X_t, for $t \in [0, T]$, be the real-valued random process describing possible paths of an instrument or portfolio log-return over the interval $[0, t]$; without loss of generality, we set $X_0 = 0$. For practical implementation, let us assume that the process is

observed on an equally spaced time grid $0, \Delta, \ldots, K\Delta = T$. In standard financial applications Δ is set at 1 day, and $K\Delta = T$ is 10 days. Define the process of the minimum, M_k, up to the kth monitoring date as $M_k := \min_{i=0,\ldots,k} X_{i\Delta}$. The VaR-I at confidence level $(1-\alpha)$ is defined as the absolute value of the α-quantile of the distribution of the random variable M_k, that is,

$$P(M_k \leq -\text{VaR-I} | X_0 = x) \quad = \quad \alpha. \tag{13}$$

The idea is that during the investment life, the path of returns can reach high negative values, which investors may care about. In such cases, the left tail of the minimum return distribution better represents risk than the left tail of the return distribution itself.

Although the distribution of the minimum return is analytically known under the assumption of an arithmetic Brownian motion (see Kritzman and Rich (2002) for the case with continuous monitoring, and see Fusai, Abrahams, and Sgarra (2006) for the discrete-monitoring one), under more general assumptions for the driving process, it must be recovered numerically (see Fusai, Germano, and Marazzina (2016) for a review and comparison of different approaches in computing the distribution of the minimum under Lévy processes). For this purpose, we resort to the FST algorithm introduced by Jackson et al. (2008) for option-pricing purposes. Our problem is indeed equivalent to finding the value of a down-and-out binary option, that is, an option paying 1 if the underlying does not hit a certain lower barrier within a given time period, and 0 otherwise. However, due to the nature of the application under consideration, our computations are performed under the physical probability measure.

With a fixed arbitrary threshold y, the FST algorithm allows us to recover the value function

$$v(0,x) \quad := \quad \mathrm{E}[1_{\{M_K > y\}} | X_0 = x] \quad = \quad P(M_K > y | X_0 = x)$$

via backward recursion so that

$$\begin{aligned} v^K(x) \quad &:= \quad v(T,x) \quad = \quad 1_{\{x>y\}}, \\ v^{k-1}(x) \quad &= \quad \mathrm{FFT}^{-1}[\mathrm{FFT}[v^k(x)]e^{\varphi\Delta}]1_{\{x>y\}}, \quad k = 1, \ldots, K, \end{aligned} \tag{14}$$

where φ is the characteristic exponent of X_t, $\mathrm{FFT}(X)$ computes the discrete Fourier transform of the vector X using the fast Fourier transform (FFT) algorithm, and $\mathrm{FFT}^{-1}(\cdot)$ denotes the inverse discrete Fourier transform. For further details on the FST algorithm, refer to Jackson et al. (2008). Further, in virtue of the translation-invariance property of Lévy processes, it follows that $v(0,x) = P(M_K > y - x | X_0 = 0)$. Hence, the computation of the $(1-\alpha)$-VaR-I can be summarized in the following steps:

Step 1. Choose an arbitrary threshold y.

Step 2. Compute the function $v(0,x) = v^0(x)$ by means of the FST algorithm.

Step 3. Find the value x such that $v(0,x) = 1-\alpha$.

Step 4. Compute the VaR-I as $\text{VaR-I} = -(y-x)$.

https://doi.org/10.1017/S0021090180001321 Published online by Cambridge University Press

The implementation of the FST iteration in equation (14) requires the expression of the characteristic function of the process X_t of the log-return of a portfolio of assets with weights w_n. In the given multivariate setting of equation (1), for short time horizons, this expression can be easily derived in virtue of the approximation of the portfolio returns as linear combinations of the asset log-returns. Exploiting the independence of the idiosyncratic process, $Y^{(n)}$, $n=1,\ldots,N$ and the systematic process, Z, we can in fact obtain

$$
\begin{aligned}
(15) \qquad \mathrm{E}\left[\exp(iuX)\right] &= \mathrm{E}\left[\exp\left(iu\left(\sum_{n=1}^{N} w_n Y^{(n)} + Z\sum_{n=1}^{N} w_n a_n\right)\right)\right] \\
&= \left(\prod_{n=1}^{N} \phi_{Y^{(n)}}(uw_n)\right)\phi_Z\left(u\sum_{n=1}^{N} w_n a_n\right), \quad u\in\mathbb{R},
\end{aligned}
$$

where we omit time subscripts to simplify the notation.[4] The characteristic functions in equation (15) are then chosen according to the specified model.

B. Estimation Results

We estimate the 10-day 99% VaR and VaR-I for an equally weighted portfolio under the all-NIG, all-MJD, and all-Gaussian (henceforth Gaussian) models. We include in the portfolio 20 of the most capitalized stocks[5] in the S&P 500 index. Our sample includes daily log-returns from May 24, 2011, to May 20, 2013.

We start by estimating the parameters of the characteristic functions of X_t required for implementation of the FST iteration as described previously. We apply the 2-step procedure presented in Section II.B. We estimate the number of factors using the Ahn and Horenstein (2013) method and find one factor, which we estimate via the principal component method. We do not report full estimation results, due to the large number of parameters (complete parameter estimation results are available from the authors). Instead, in Table 1 we report the estimated mean, standard deviation, Pearson's moment coefficient of skewness, and index of excess kurtosis of the return distribution using both the all-NIG and all-MJD estimated model parameters for a selection of assets and the equally weighted portfolio of all 20 assets. Our objective is to assess the assumption of nonnormality of returns by testing if skewness and excess kurtosis are significantly different from 0. For the test, we use bootstrap (Efron (1979)) to generate 5,000 resampled data sets from our observed one, and we estimate the model parameters on each of the resampled data sets. Then, we compute the moments of the distribution of the 20 stocks across the 5,000 data sets according to the model; the α-confidence levels for a given moment are built using the $(1-\alpha)/2$ and $(1+\alpha)/2$ quantiles of that moment over the 5,000 data sets. Based on these confidence intervals, in

[4] If returns are very volatile or the horizon is longer, it becomes essential to work with linear returns. In this case, equation (15) no longer holds. For more details, see Meucci (2005).

[5] Apple, Exxon Mobil Corporation, Wal-Mart Stores, Microsoft Corporation, Google, General Electric, IBM, Chevron Corporation, Berkshire Hathaway, AT&T, Procter & Gamble, Pfizer, Johnson & Johnson, Wells Fargo & Co., Coca-Cola, JPMorgan Chase & Co., Oracle, Merck & Co., Verizon Communications, and Amazon.

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

TABLE 1
Estimated Distribution Characteristics

In Table 1, we report the estimated daily mean (expressed in basis points (bps)), standard deviation, Pearson's moment coefficient of skewness, and index of excess kurtosis for the returns of a selection of assets and the equally weighted portfolio under the all-NIG and all-MJD estimated models. *, **, and *** represent statistical significance at the 90%, 95%, and 99% levels, respectively (confidence levels obtained by bootstrap resampling techniques with 5,000 iterations). We also report the loadings **a**.

	Apple	Google	AT&T	Coca-Cola	Amazon	Portfolio
Panel A. All-NIG						
Mean (bps)	0.0174	9.1666	6.9314	1.8355	7.9301	8.3977
Standard deviation	0.01887***	0.01537***	0.01037***	0.00977***	0.02147***	0.01387***
Skewness	0.0789	−0.2719	−0.4215**	−0.1067	0.3198	−0.0433
Excess kurtosis	1.9837***	3.1409***	2.5857***	1.7734***	8.3284***	4.8165***
Panel B. All-MJD						
Mean (bps)	0.0175	9.1666	6.9313	1.8355	7.9301	8.3978
Standard deviation	0.01897***	0.01597***	0.01067***	0.00967***	0.02307***	0.01407***
Skewness	−0.1205	−0.5464	−0.9966**	−0.0147	0.2459	−0.1640
Excess kurtosis	3.1272***	5.8531***	5.1519***	0.9730***	10.7512***	7.7416***
a	0.0087***	0.0092***	0.0072***	0.0067***	0.0097***	0.0092***

Table 1, the *, **, and *** represent statistical significance at the 90%, 95%, and 99% levels, respectively.

The results in Table 1 show that in general the skewness is not significantly different from 0, indicating generally symmetric asset distributions; however, the excess kurtosis is always statistically significant, indicating heavier tails than in the case of normal distributed assets' returns. The features of the returns of the equally weighted portfolio indicate that the nonnormality inherited from each asset is persistent regardless of the level of diversification in place. These results confirm the importance of modeling asset returns under realistic distributional assumptions away from normality. The differences in the level of skewness and excess kurtosis between the all-NIG and all-MJD models reflect the different flexibility offered by the processes in portraying the distribution tails.

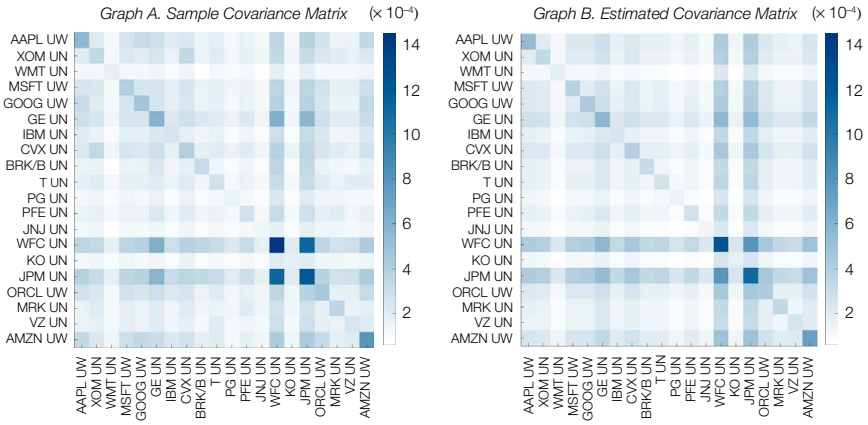
Next, we evaluate whether our fitted multivariate Lévy model is able to capture the dependence observed in real data. To do so, we compare the sample covariance of the assets in the portfolio with the covariance matrix estimated assuming the model in equation (1) for the all-NIG case. Figure 1 shows the two covariances, using two color-coded matrices in which each entry is colored according to its value, and the conversion color-code is provided in the lateral color bar. We notice that the all-NIG model accurately reproduces the sample covariance among the assets in our data set. Similar results are obtained for the all-MJD model (available from the authors).

To further explore whether our fitted multivariate Lévy models correctly capture the observed distribution of portfolio returns, we perform a simulation exercise. We randomly generate 1,000 long-only portfolios and 1,000 long–short portfolios using our estimated parameters for the all-NIG and all-MJD models. For comparison, we also generate such portfolios under a Gaussian model. We compare the simulated distributions with the observed sample distribution (as in, e.g., Eberlein and Madan (2009), Luciano, Marena, and Semeraro (2016)). Long-only weights are generated by drawing an IID sample from a standard normal distribution, taking the absolute value, and rescaling by the sum. Long–short

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

FIGURE 1
Sample and Estimated Covariance Matrices

Figure 1 shows the sample and estimated covariance matrices for the all-NIG model. Estimation of the common factor Z is performed via the principal-components method. The color-values conversion is provided in the side color bar. Constituents: Apple, Exxon Mobil Corporation, Wal-Mart Stores, Microsoft Corporation, Google, General Electric, IBM, Chevron Corporation, Berkshire Hathaway, AT&T, Procter & Gamble, Pfizer, Johnson & Johnson, Wells Fargo & Co., Coca-Cola, JPMorgan Chase & Co., Oracle, Merck & Co., Verizon Communications, and Amazon.



portfolio weights are generated similarly, drawing an IID sample from a standard normal distribution and rescaling it by the sum of the squares. We perform the Kolmogorov–Smirnov test,[6] with the null hypothesis that the simulated distribution is drawn from the sample distribution. The results, presented in Table 2, show the proportion of portfolios for which the null hypothesis is rejected using 1%, 5%, and 10% significance levels for all models. We note that both Lévy-based models significantly outperform the Gaussian one; in fact, for the case of long-only portfolios, the Gaussian model is rejected 100% of the time. Further, the

TABLE 2
Goodness-of-Fit Test

In Table 2, we report the proportion of simulated portfolios for which the null hypothesis of the Kolmogorov–Smirnov test is rejected at the 1%, 5%, and 10% significance levels. The null hypothesis is that the simulated distribution is drawn from the sample distribution. Portfolios are generated under the all-NIG, all-MJD, and Gaussian models. Constituents: Apple, Exxon Mobil Corporation, Wal-Mart Stores, Microsoft Corporation, Google, General Electric, IBM, Chevron Corporation, Berkshire Hathaway, AT&T, Procter & Gamble, Pfizer, Johnson & Johnson, Wells Fargo & Co., Coca-Cola, JPMorgan Chase & Co., Oracle, Merck & Co., Verizon Communications, and Amazon.

	Long-Only			Long–Short		
Significance Level	All-NIG	All-MJD	Gaussian	All-NIG	All-MJD	Gaussian
0.01	0.00%	0.00%	100%	0.00%	0.00%	65.10%
0.05	0.00%	0.00%	100%	0.80%	0.70%	74.20%
0.10	0.00%	0.00%	100%	2.10%	2.00%	79.00%

[6]We derive the probability density function by inverting the portfolio characteristic function in equation (15) using the COS method of Fang and Oosterlee (2008).

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

all-NIG and all-MJD model specifications fit the sample distribution of returns for both long-only and long–short portfolios equally well.

Once we have confirmed that our estimated model is able to successfully replicate the observed distribution of portfolio returns, we proceed to estimate the portfolio VaR and VaR-I following the 4 steps described in Section IV.A. In Table 3, we report our estimates of 10-day 99% VaR and VaR-I and the corresponding confidence intervals at the 95% level computed using bootstrap resampling methods. The confidence intervals of the portfolio VaR/VaR-I are calculated using the quantiles of the VaR/VaR-I across the 5,000 instances stemming from the bootstrapped samples. We also report the ratio between the Lévy model estimate and the Gaussian model estimate ("Multiples"). We observe that VaR-I consistently exceeds the traditional VaR and that jump risk tends to amplify intra-horizon risk.

These results indicate that the pure-jump all-NIG model has the (marginally) thickest tails for both the return and the minimum return distributions and thus provides the most conservative risk estimates, with a VaR 1.09 times higher than the VaR under the Gaussian model and a VaR-I approximately 1.11 times higher with respect to the Gaussian one. The VaR and the VaR-I under the jump-diffusion all-MJD specification are, respectively, 1.03 and 1.04 times higher with respect to the corresponding measures under the Gaussian model. These results reflect the slower decay in the distribution tails of the NIG and the MJD processes compared with the Brownian motion, as discussed in Section II.A.

As a final consideration, we observe the following: Although in a few cases Table 1 highlights fairly different estimates for skewness and excess kurtosis in the distribution of each asset between the all-NIG and the all-MJD models, the relatively similar figures reported in Table 3 for the risk measures of interest show the effect of the allocation in place in the portfolio, which diversifies away the tail risk of the idiosyncratic components but maintains the exposure to the tail risk of the systematic risk factor. This also suggests that a change in the portfolio's weights can potentially generate a significant change in the overall risk measure depending on which asset becomes "predominant," so to speak. This latter effect can only be captured by means of a multivariate model for the returns of each

TABLE 3
VaR and VaR-I Estimates

In Table 3, we report the 10-day horizon 99% Variance at Risk (VaR) and intra-horizon Variance at Risk (VaR-I) of an equally weighted portfolio under the all-NIG, all-MJD, and Gaussian models, with confidence intervals at the 95% level. The common factor Z is estimated via the principal-component method. Confidence intervals are computed using bootstrap resampling methods (5,000 iterations). Multiples: Lévy model estimate/Gaussian model estimate. Constituents: Apple, Exxon Mobil Corporation, Wal-Mart Stores, Microsoft Corporation, Google, General Electric, IBM, Chevron Corporation, Berkshire Hathaway, AT&T, Procter & Gamble, Pfizer, Johnson & Johnson, Wells Fargo & Co., Coca-Cola, JPMorgan Chase & Co., Oracle, Merck & Co., Verizon Communications, and Amazon.

	VaR				VaR-I			
	Estimate	CI(lb)	CI(ub)	Multiples	Estimate	CI(lb)	CI(ub)	Multiples
Gaussian	0.0699	0.0572	0.0831	1.0000	0.0738	0.0614	0.0869	1.0000
All-MJD	0.0723	0.0579	0.0888	1.0341	0.0769	0.0630	0.0929	1.0413
All-NIG	0.0764	0.0599	0.0961	1.0939	0.0818	0.0652	0.1016	1.1085

portfolio's components, as opposed to a univariate model for the overall portfolio returns.

Motivated by the previous analysis, we conclude this section with the identification of the risk contribution of each asset to the whole-portfolio risk. This is especially relevant for active risk portfolio managers, who require identification of the effects of changes in portfolios' weights on the overall portfolio risk in order to modify the overall risk profile most effectively.

The study is based on a sensitivity analysis of the VaR-I with respect to the portfolio weights, which is performed by finite difference. For illustration purposes, we consider the case of an equally weighted portfolio; for each asset in the portfolio, we perturb its weight by 1/100 and then recompute the VaR-I. The change in the VaR-I (marginal VaR-I) is the discrete analogue of a derivative. If we multiply marginal VaR-I by the asset weight in the portfolio and divide this product by the original VaR-I, we obtain a percentage measure of the risk contribution of this asset, the so-called Component VaR-I. This measure can be interpreted as the percentage increase in VaR-I for a 1% change in the weight of a given asset in the portfolio. Notice that, given any prespecified portfolio, risk measures and risk attribution are obtained without reestimating the underlying multivariate model parameters.

The results on the Component VaR-I of the 20 assets in the portfolio considered in this section are reported in Table 4. The decomposition highlights the positions that the portfolio is most sensitive to. Interestingly, JPMorgan Chase & Co. and Wells Fargo & Co. are the two assets that contribute the most to the portfolio risk measured by VaR-I. A 1% increase in the portfolio weight of JPMorgan Chase & Co. increases VaR-I by almost 9%. The asset with the smallest Component VaR-I is Wal-Mart Stores, with an increase of approximately 2.5% when its weight in the portfolio is increased by 1%. There are no

TABLE 4

Asset's Individual Risk Contribution to the Equally Weighted Portfolio VaR-I

In Table 4, we report the component intra-horizon Variance at Risk (VaR-I), which can be interpreted as the percentage increase in VaR-I for a 1% increase in the weight of the asset in the portfolio. Finite difference is calculated with perturbation 1/100.

Asset	All-NIG	All-MJD
Apple	4.88%	4.93%
Exxon Mobil Corporation	5.37%	5.38%
Wal-Mart Stores	2.44%	2.24%
Microsoft Corporation	4.88%	4.93%
Google	4.88%	4.93%
General Electric	6.34%	6.73%
IBM	4.39%	4.48%
Chevron Corporation	5.85%	5.83%
Berkshire Hathaway	5.85%	5.83%
AT&T	3.41%	3.59%
Procter & Gamble	2.93%	2.69%
Pfizer	4.39%	4.48%
Johnson & Johnson	2.93%	3.15%
Wells Fargo & Co.	8.29%	8.07%
Coca-Cola	3.41%	3.59%
JPMorgan Chase & Co.	8.78%	8.97%
Oracle	7.32%	7.17%
Merck & Co.	4.39%	4.04%
Verizon Communications	2.93%	3.14%
Amazon	6.34%	5.83%

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

significant differences between the all-NIG and the all-MJD models, which suggests that the proposed estimation method is robust with respect to the model choice. Similar results can be obtained for the case of VaR and are available from the authors.

In unreported experiments, we repeat the estimation using the S&P 500 index as a proxy of the common factor. The estimated VaR and VaR-I are close to those obtained by estimating the common factor via principal components, displayed in Table 3. However, the ability of the model to fit the portfolio distribution, measured by the Kolmogorov–Smirnov test on the distributions of 1,000 randomly generated long-only portfolios and 1,000 randomly generated long–short portfolios, turns out to be substantially lower than in the case in which the principal-component method is adopted for the estimation of the common factor. The results are available from the authors.

V. Conclusions

We propose an estimation procedure for multivariate asset models based on the linear transformation of Lévy processes as in Ballotta and Bonfiglioli (2016), allowing for an extension of the use of multivariate Lévy models to risk and portfolio management applications. We note that factor constructions are in line with recommendations from the Basel Committee on Banking Supervision (Basel (2013)) for the development of internal models.

For the case of an N-asset portfolio, the 2-step estimation procedure proposed in this article reduces to the estimation of the common factor Z and the loadings $\boldsymbol{a}$ via principal components and N univariate estimations, one per each idiosyncratic component; therefore, it is fast to implement, and its complexity does not increase with the number of components of the multivariate model. Our simulation study reveals that this approach is almost as accurate as a more traditional direct ML estimation of the whole set of parameters, as long as proper univariate estimation methods are used; however, the 2-step procedure proves to be significantly more efficient from the computational point of view. The proposed approach is flexible with respect to the number of assets included in the portfolio and does not impose any convolution condition on the factors, as it is assumed in other multivariate constructions proposed in the literature. Although in the numerical studies presented in this article we conveniently assume that all factors are modeled using the same type of process, this assumption can be relaxed so as to allow any Lévy process for the idiosyncratic part across all the names included in the portfolio in order to accommodate different tail behaviors.

As an application, we employ the proposed estimation procedure for the calculation of the VaR-I of a portfolio of assets following the model under consideration by means of the FST algorithm. The numerical study reveals the importance of properly capturing realistic features of asset log-returns, such as skewness and excess kurtosis, by incorporating jumps in the risk dynamic. The results from the empirical study, in fact, highlight the more conservative risk estimates offered by the VaR-I, especially for the case of the NIG, reflecting the different decay behavior of the distribution tails.

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

Due to the short horizons typical of risk management operations, in this article we have considered the case of a non-Gaussian multivariate model built on Lévy processes (i.e., processes characterized by independent and stationary increments). For applications aimed at longer horizons, stochastic volatility features are required. Lévy processes can be conveniently equipped with such features by means of suitably constructed time changes, as proposed by Carr and Wu (2004), (2007), for example, and more recently extended by Ballotta and Rayée (2017). Although multivariate extensions of the time-changed Lévy processes framework are currently investigated, in the case in which a similar factor structure is adopted, we envisage the potential of the 2-step methodology proposed in this article for the estimation under the physical probability measure. The analysis of the validity of conditions ensuring the consistency of the estimated factors and loadings in this context is left to future research. Finally, we observe that the methodology proposed in this article can also be directly applied to other areas, such as portfolio optimization problems based on multivariate Lévy processes (as in, e.g., Loregian (2013)).

Appendix A. Detailed Simulation Procedures and Results

In Appendix A, we provide detailed procedures and complete results of our simulation experiments designed to evaluate the performance of the proposed methodology. We present our results in comparison with the 1-step estimation approach (i.e., via single maximization of the likelihood function in equation (4)). Because the detailed simulation study of Bai (2003) shows that the factor-estimation method has strong small-sample properties, without loss of generality, our simulation study assumes that the systematic risk factor is proxied by a well-diversified index.[7]

Our data-generation process is based on daily log-returns of the S&P 500 index and a selection of its constituent stocks; further, we assume that the S&P 500 index is the true driver of the commonality in stocks returns. The observation period ranges from Sept. 10, 2007, to May 20, 2013, for a total of 1,434 observations per series. These data are extracted from Bloomberg and adjusted for dividends. We first estimate the chosen multivariate model using the index log-returns as a proxy for the systematic process Z and use the estimated parameters to simulate series of the returns of the assets under consideration. Then the 1-step and 2-step estimation procedures are applied to the generated data to recover the distribution of each parameter.

1. Computation Efficiency

The first objective of our simulation is to evaluate the computation-efficiency gains of the 2-step approach versus the 1-step approach. We use estimation errors and estimation time to calculate an efficiency-gain index commonly used in Monte Carlo simulation analysis and defined, for example, by Glasserman (2004). Given a specification of the model in equation (1), characterized by $\bar{k}$ parameters, we compute the efficiency gain, E_{21}, of the

[7] We also perform a small simulation study to check the reliability of the factor-estimation method; we find that the average correlation between estimated factors and loadings with true factors and loadings is larger than 0.98. Also, simulation results using estimated factors and loadings are similar in magnitude to the ones obtained assuming that the factor is observable, as reported in this section.

2-step procedure to the 1-step ML approach as

$$E_{21} = \frac{\overline{\text{MSE}}_1 \tau_1}{\overline{\text{MSE}}_2 \tau_2}, \tag{A-1}$$

where $\overline{\text{MSE}}$ denotes the average mean-square error

$$\overline{\text{MSE}} = \frac{\sum_{k=1}^{\bar{k}} \text{MSE}(\hat{\theta}_k)}{\bar{k}},$$

of the parameters estimated by the 1-step (1) and the 2-step (2) approach. $\text{MSE}(\hat{\theta}_k)$ is the mean-square error for S simulation iterations

$$\frac{1}{S}\sum_{s=1}^{S}\left(\hat{\theta}_s - \theta\right)^2,$$

and $\tau_{1\ (2)}$ is the average time needed to estimate the model parameters using the 1 (2) approach. In particular, we compute the efficiency-gain index corresponding to the all-NIG and all-MJD models with $N=5$ and $N=15$ components. In the case with 5 assets, for each of the two approaches, we consider the mean-square errors based on 1,000 simulations; for the case with 15 assets, we rely on 100 simulations only due to the computational cost of the 1-step procedure.

Results are reported in Table A1 for the case in which $T=500$ (i.e., approximately 2 years of daily observations). We observe that the 2-step approach is significantly more efficient in terms of computational time. Moreover, for $N=5$, the average mean-square errors attained with the 2-step approach are lower than those given by the 1-step procedure (8.5% vs. 14% for the all-NIG model, 0.14% vs. 0.28% for the all-MJD model), whereas they are almost the same for $N=15$ (approximately 10% for the all-NIG model, 0.16% for all-MJD). According to the efficiency index (A-1), in our experiment, the 2-step procedure performed 3,496 times more efficiently than the 1-step approach in the worst case (all-MJD, $N=15$) and 8,139 times more efficiently in the best one (all-NIG, $N=5$).

2. Finite-Sample Properties

Our second objective is to assess the error, bias, and inefficiency of the parameter estimates obtained by adopting the described estimation procedures. We assess the estimation procedure for four different sample sizes, varying the length of the simulated series from 1 year up to 4 years of daily observations and varying the number of components, considering up to 30 assets in the simulated markets.

TABLE A1
Computation-Efficiency Gains

We report the average mean-square error (MSE), computation times (measured in seconds), and efficiency gains of the 2-step approach to the 1-step maximum likelihood. $T=500$ days, and N is the number of assets in the portfolio. Efficiency gains are $E_{21}=\overline{\text{MSE}}_1\tau_1/(\overline{\text{MSE}}_2\tau_2)$, and $\bar{k}$ is the number of model parameters. (Processor: Intel Core i5-2400 CPU @ 3.10 GHz; RAM: 4.00 GB).

			2-Step		1-Step		
N	Model	$\bar{k}$	$\overline{\text{MSE}}$	Time	$\overline{\text{MSE}}$	Time	E_{21}
5	All-NIG	29	0.0857	0.7	0.1407	3,668.5	8,139
	All-MJD	35	0.0014	1.0	0.0028	3,756.8	7,295
15	All-NIG	79	0.1043	1.9	0.0973	10,723.3	5,350
	All-MJD	95	0.0016	3.4	0.0017	11,087.5	3,496

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

The assessment is made in terms of RMSE, bias, and inefficiency, respectively defined as follows:

$$\text{(A-2)} \qquad \text{RMSE}(\hat{\theta}) = \sqrt{\frac{1}{S}\sum_{s=1}^{S}\left(\hat{\theta}_s - \theta\right)^2},$$

$$\text{(A-3)} \qquad \text{bias}(\hat{\theta}) = \left|\text{E}[\hat{\theta}] - \theta\right|,$$

$$\text{(A-4)} \qquad \text{ineff}(\hat{\theta}) = \sqrt{\frac{1}{S}\sum_{s=1}^{S}\left\{\left(\hat{\theta}_s - \text{E}[\hat{\theta}]\right)^2\right\}},$$

where $\hat{\theta}$ indicates the estimates of the true parameter set θ used in the simulation step, and $\text{E}[\hat{\theta}] = \sum_{s=1}^{S}\hat{\theta}_s/S$.

We start by analyzing the finite-sample properties of the 2-step estimation procedure proposed in Section II.B.

We first estimate the chosen multivariate model using the index log-returns as a proxy for the systematic process Z. Next, we use the estimated parameters to simulate series of the returns of the assets under consideration, which the estimation procedure is then reapplied to. This allows us to recover the distribution of each parameter. We assess the estimation procedure in several cases, varying the length of the simulated series from 1 year up to 4 years of daily observations, $T = [250, 500, 750, 1{,}000]$, and varying the number of components, considering up to 30 assets in the simulated markets (N). For each of the 16 cases taken into account, we repeat the simulation and estimation $S = 10{,}000$ times, obtaining 10,000 sets of parameters, denoted by $\hat{\boldsymbol{\theta}}_s, s = 1, \ldots, S$.

Given the large number of parameters (if $N = 5$, the total number of parameters is 29 for the all-NIG model and 35 for the all-MJD model; if $N = 30$, there are 154 parameters for the all-NIG model and 185 parameters for the all-MJD model), we cannot display detailed results for each asset; hence, for illustrative purposes, we show only the assessment results for the estimation of the common factor Z, the first idiosyncratic factor $Y^{(1)}$, and average results relative to the loadings $a_n, n = 1, \ldots, N$. Complete results are available from the authors. We stress that the focus of our simulation study is on investigating the effectiveness of splitting the estimation procedure of the multivariate model in the 2 steps presented in Section II.B.

First, we analyze the finite-sample properties of the estimated systematic component in our model.

Table A2 displays the RMSE, bias (expressed in percentage terms with respect to the true parameter value), and inefficiency of the estimators for the all-NIG and all-MJD models as the length of the simulated series varies in $T = [250, 500, 750, 1{,}000]$. We observe, in general, a low level of bias for all of the estimators, meaning that the ML estimators are suitable for the first step of our procedure. Analogous considerations hold in the all-MJD case, with estimators obtained by EM. Hence, errors and inefficiency levels can be used as benchmarks to evaluate step 2.

In more detail, the bias for the all-NIG model is generally lower than 2% for sample sizes larger than 500. For smaller sample sizes, we observe some problems only in the estimation of the μ and θ parameters. As previously observed, these parameters control the mean of the process, which is well known to be very difficult to estimate in a reliable way. Concerning the all-MJD model, the bias appears to be larger, although still acceptable; the main issues are related to the estimation of the intensity and mean of the jump severities. However, the RMSEs are reasonably small for all parameters.

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

TABLE A2

2-Step Procedure Assessment: Common Factor

The estimation in Table A2 is based on Standard & Poor (S&P) daily log-returns. The observation period is Sept. 10, 2007–May 20, 2013. Source: Bloomberg. Estimation errors are expressed in absolute terms. The root-mean-square error (RMSE), bias, and inefficiency indices are computed according to equations (A-2)–(A-4). Bias is expressed in percentage terms with respect to the true parameter value.

Z	$T=250$	$T=500$	$T=750$	$T=1{,}000$
Panel A. All-NIG Model				
$\mu=0.0014$				
RMSE	9.85E−04	6.72E−04	5.42E−04	4.65E−04
Bias	3.10%	0.86%	1.32%	0.48%
Inefficiency	9.84E−04	6.71E−04	5.41E−04	4.65E−04
$\theta=-0.0014$				
RMSE	1.47E−03	1.02E−03	8.20E−04	7.12E−04
Bias	2.21%	1.73%	1.33%	0.35%
Inefficiency	1.47E−03	1.02E−03	8.20E−04	7.12E−04
$\sigma=0.0168$				
RMSE	1.76E−03	1.23E−03	1.01E−03	8.77E−04
Bias	1.05%	0.51%	0.38%	0.28%
Inefficiency	1.75E−03	1.22E−03	1.01E−03	8.75E−04
$k=3.32$				
RMSE	1.30E+00	8.97E−01	7.26E−01	6.32E−01
Bias	0.58%	0.25%	0.02%	0.17%
Inefficiency	1.30E+00	8.97E−01	7.26E−01	6.32E−01
Panel B. All-MJD Model				
$\mu=0.0012$				
RMSE	8.24E−04	5.83E−04	4.66E−04	4.05E−04
Bias	2.34%	1.52%	2.07%	2.22%
Inefficiency	8.23E−04	5.83E−04	4.65E−04	4.04E−04
$\sigma=0.0075$				
RMSE	1.17E−03	8.90E−04	7.41E−04	7.41E−04
Bias	0.97%	1.75%	1.71%	1.96%
Inefficiency	1.17E−03	8.80E−04	7.30E−04	7.27E−04
$\nu=-0.0025$				
RMSE	3.14E−03	1.90E−03	1.51E−03	1.28E−03
Bias	5.39%	6.87%	2.64%	2.65%
Inefficiency	3.13E−03	1.89E−03	1.51E−03	1.28E−03
$\tau=0.0210$				
RMSE	3.56E−03	2.82E−03	2.39E−03	2.36E−03
Bias	2.53%	2.80%	2.44%	2.66%
Inefficiency	3.52E−03	2.76E−03	2.33E−03	2.30E−03
$\lambda=0.47$				
RMSE	1.50E−01	1.07E−01	8.72E−02	7.88E−02
Bias	3.14%	4.13%	4.95%	4.28%
Inefficiency	1.49E−01	1.05E−01	8.40E−02	7.62E−02

In general, consistent with the literature (see, e.g., Aït-Sahalia and Jacod (2011) and references therein), infinite-activity processes, like the NIG, can be estimated in a more reliable manner than finite-activity processes, such as the MJD process.

Then, we focus on the finite-sample properties of the estimated idiosyncratic component in our model.

We implement step 2 by, first, solving the minimization problem in equation (8) with respect to the loadings $\boldsymbol{a}$; second, we use the estimated loadings as starting values to solve the N maximization problems in equation (7) with respect to $\boldsymbol{\theta}_{Y^{(n)}}$ for all $n=1,\ldots,N$. The minimization procedure in equation (8) is performed by fixing the variance of the common factor equal to the sample variance of the simulated series of the process Z; in this way, the assessment of this step turns out to be independent of the model specification.

The results of the estimation of the idiosyncratic process are presented in Table A3 for the case of the first instrument. Results relative to the other assets are available from

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

TABLE A3

2-Step Procedure Assessment: First Idiosyncratic Component

In Table A3, the observation period is Sept. 10, 2007–May 20, 2013. Source: Bloomberg. Estimation errors expressed in absolute terms. The root-mean-square error (RMSE), bias, and inefficiency indices are computed according to equations (A-2)–(A-4). Bias is expressed in percentage terms with respect to the true parameter value.

	$N=30$				$T=500$			
Y_1	$T=250$	$T=500$	$T=750$	$T=1{,}000$	$N=5$	$N=10$	$N=15$	$N=30$
Panel A. All-NIG Model								
$\mu=9.92\text{E}{-04}$								
RMSE	2.17E−03	1.13E−03	9.00E−04	7.69E−04	1.13E−03	1.12E−03	1.14E−03	1.13E−03
Bias	1.10%	0.28%	2.60%	0.47%	0.58%	0.01%	1.19%	0.28%
Inefficiency	2.17E−03	1.13E−03	9.00E−04	7.69E−04	1.13E−03	1.12E−03	1.14E−03	1.13E−03
$\theta=2.15\text{E}{-04}$								
RMSE	2.45E−03	1.37E−03	1.09E−03	9.40E−04	1.39E−03	1.37E−03	1.40E−03	1.37E−03
Bias	4.06%	4.42%	16.08%	4.22%	5.44%	4.08%	8.40%	4.42%
Inefficiency	2.45E−03	1.37E−03	1.09E−03	9.40E−04	1.39E−03	1.37E−03	1.40E−03	1.37E−03
$\sigma=0.0173$								
RMSE	1.39E−03	9.71E−04	7.97E−04	6.74E−04	9.65E−04	9.60E−04	9.61E−04	9.71E−04
Bias	1.20%	0.59%	0.46%	0.37%	0.61%	0.61%	0.54%	0.59%
Inefficiency	1.37E−03	9.66E−04	7.93E−04	6.71E−04	9.60E−04	9.54E−04	9.56E−04	9.66E−04
$k=1.483$								
RMSE	6.19E−01	4.31E−01	3.48E−01	3.02E−01	4.29E−01	4.28E−01	4.28E−01	4.31E−01
Bias	1.37%	0.50%	1.01%	0.48%	0.95%	1.03%	0.75%	0.50%
Inefficiency	6.19E−01	4.31E−01	3.47E−01	3.02E−01	4.29E−01	4.27E−01	4.28E−01	4.31E−01
Panel B. All-MJD Model								
$\mu=0.00133$								
RMSE	1.10E−03	7.61E−04	6.12E−04	5.33E−04	7.55E−04	7.63E−04	7.57E−04	7.61E−04
Bias	0.70%	0.28%	0.57%	0.01%	0.12%	0.96%	0.18%	0.28%
Inefficiency	1.10E−03	7.61E−04	6.12E−04	5.33E−04	7.55E−04	7.63E−04	7.57E−04	7.61E−04
$\sigma=0.01113$								
RMSE	1.36E−03	1.01E−03	8.76E−04	8.24E−04	1.03E−03	1.02E−03	1.01E−03	1.01E−03
Bias	0.70%	0.62%	0.44%	0.04%	0.41%	0.50%	0.68%	0.62%
Inefficiency	1.36E−03	1.01E−03	8.74E−04	8.24E−04	1.02E−03	1.02E−03	1.01E−03	1.01E−03
$\nu=-0.0004$								
RMSE	7.78E−03	3.26E−03	2.43E−03	2.03E−03	3.25E−03	3.18E−03	3.09E−03	3.26E−03
Bias	40.44%	8.25%	0.96%	1.40%	10.60%	11.05%	2.75%	8.25%
Inefficiency	7.78E−03	3.26E−03	2.43E−03	2.03E−03	3.25E−03	3.18E−03	3.09E−03	3.26E−03
$\tau=0.02429$								
RMSE	6.20E−03	4.40E−03	3.81E−03	3.62E−03	4.46E−03	4.40E−03	4.38E−03	4.40E−03
Bias	0.82%	0.93%	0.93%	1.35%	1.14%	1.27%	0.91%	0.93%
Inefficiency	6.20E−03	4.39E−03	3.81E−03	3.60E−03	4.45E−03	4.39E−03	4.38E−03	4.39E−03
$\lambda=0.29214$								
RMSE	1.61E−01	1.21E−01	1.03E−01	9.06E−02	1.21E−01	1.20E−01	1.21E−01	1.21E−01
Bias	5.53%	5.33%	4.02%	1.86%	5.28%	4.77%	5.55%	5.33%
Inefficiency	1.60E−01	1.20E−01	1.02E−01	9.04E−02	1.20E−01	1.19E−01	1.19E−01	1.20E−01

the authors. In particular, the left-hand side of Table A3 displays the RMSE, bias, and inefficiency of the estimators when the total number of assets is fixed ($N=30$) and the length of the simulated series varies in $T=[250,500,750,1{,}000]$. On the right-hand side of Table A3, we show the assessment results for a fixed $T=500$, varying the number of assets. Consistent with the results shown previously, Table A3 reveals similar estimation errors for $N=[5,10,15,30]$, showing that the number of assets has only a minimal impact on the estimation errors of the idiosyncratic terms for both the specifications we tested. Further, the results in Table A3 reveal very little bias, implying that our estimation procedure performs as expected. Moreover, we observe estimation errors and inefficiency levels in line with those obtained in step 1; therefore, splitting the estimation procedure into 2 steps, ease of implementation aside, proves to be effective.

As noted earlier, in this section, we only discuss results relative to the first instrument; similar conclusions hold for all assets considered.

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

We conclude our simulation exercise by comparing the finite-sample properties of the 2-step estimation versus the 1-step estimation procedures. Thus, we repeat the simulation study for the estimation of the all-NIG and all-MJD models' parameters using the 1-step ML approach discussed in Section II.B, which represents a useful term of comparison to evaluate the results obtained from the 2-step procedure presented previously. Hence, we use the same data set as before, but we relax the assumption that the systematic risk factor $\boldsymbol{Z}$ is observable.

The ML estimation consists of maximizing the likelihood function in equation (4); the quadrature of the integral in equation (4) is performed via the trapezoidal rule.

Due to the computational cost of the procedure just highlighted, we consider a small number of assets ($N=5$, i.e., 29 parameters to be estimated for the all-NIG model, 35 for the all-MJD model), repeating the simulation 1,000 times; we then perform 100 simulations to evaluate the estimation for $N=15$ assets (i.e., 79 parameters for the all-NIG model, 95 for the all-MJD model). Results relative to the common factor $\boldsymbol{Z}$, the first idiosyncratic component $Y^{(1)}$, and the first loading a_1 are displayed in Table A4. Complete results are available from the authors.

Bearing in mind the different number of simulations performed, we can compare the results of the 2-step procedure assessment with those presented in this section. In particular, for both the all-NIG and all-MJD models, the results relative to the common factor Z can be compared with those displayed in the second column of Table A2, corresponding to estimates based on $T=500$ observations, whereas the results relative to the first

TABLE A4

1-Step Approach Assessment

In Table A4, the observation period is Sept. 10, 2007–May 20, 2013 (source: Bloomberg). Estimation errors are expressed in absolute terms. The root-mean-square error (RMSE), bias, and inefficiency indices are computed according to equations (A-2)–(A-4). Bias is expressed in percentage terms with respect to the true parameter value.

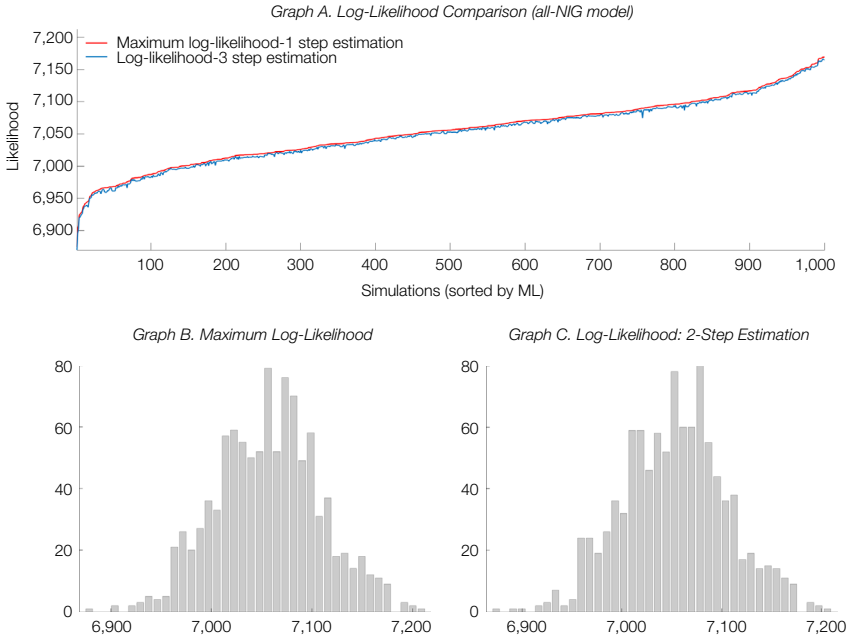
	$N=5$ (1,000 simulations)	$N=15$ (100 simulations)		$N=5$ (1,000 simulations)	$N=15$ (100 simulations)		$N=5$ (1,000 simulations)	$N=15$ (100 simulations)
Z			Y_1			First Loading		
Panel A. All-NIG Model								
$\mu=0.0014$			$\mu=9.92\text{E}{-04}$			$a_1=0.8898$		
RMSE	7.07E−04	5.67E−04		1.19E−03	1.11E−03		4.72E−02	3.85E−02
Bias	4.16%	2.98%		2.53%	15.88%		0.03%	0.18%
Inefficiency	7.04E−04	5.65E−04		1.19E−03	1.10E−03		4.72E−02	3.85E−02
$\theta=-0.0014$			$\theta=2.15\text{E}{-04}$					
RMSE	1.11E−03	8.64E−04		1.47E−03	1.52E−03			
Bias	6.89%	4.02%		7.49%	66.61%			
Inefficiency	1.10E−03	8.62E−04		1.47E−03	1.52E−03			
$\sigma=0.0168$			$\sigma=0.0173$					
RMSE	1.24E−03	1.25E−03		9.95E−04	1.12E−03			
Bias	0.21%	0.43%		0.60%	1.14%			
Inefficiency	1.24E−03	1.25E−03		9.90E−04	1.10E−03			
$k=3.32$			$k=1.483$					
RMSE	1.15E+00	8.35E−01		4.87E−01	4.59E−01			
Bias	2.68%	3.89%		0.82%	3.07%			
Inefficiency	1.14E+00	8.25E−01		4.87E−01	4.57E−01			
Panel B. All-MJD Model								
$\mu=0.0012$			$\mu=0.00133$			$a_1=0.8898$		
RMSE	6.24E−04	6.23E−04		7.85E−04	7.68E−04		5.44E−02	3.98E−02
Bias	8.04%	8.85%		3.41%	4.60%		1.21%	0.51%
Inefficiency	6.16E−04	6.13E−04		7.84E−04	7.65E−04		5.33E−02	3.95E−02
$\sigma=0.0075$			$\sigma=0.01113$					
RMSE	1.14E−03	6.80E−04		1.12E−03	9.61E−04			
Bias	4.49%	3.16%		0.92%	0.48%			
Inefficiency	1.09E−03	6.38E−04		1.11E−03	9.59E−04			
$\nu=-0.0025$			$\nu=-0.0004$					
RMSE	1.69E−03	1.85E−03		3.03E−03	3.19E−03			
Bias	3.27%	10.91%		23.86%	17.89%			
Inefficiency	1.69E−03	1.83E−03		3.03E−03	3.19E−03			
$\tau=0.0210$			$\tau=0.02429$					
RMSE	2.72E−03	3.26E−03		4.60E−03	4.27E−03			
Bias	3.73%	7.95%		1.28%	3.59%			
Inefficiency	2.60E−03	2.80E−03		4.59E−03	4.18E−03			
$\lambda=0.47$			$\lambda=0.29214$					
RMSE	2.33E−01	1.49E−01		1.36E−01	1.15E−01			
Bias	34.63%	18.80%		7.35%	4.12%			
Inefficiency	1.66E−01	1.20E−01		1.35E−01	1.14E−01			

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

FIGURE A1

Maximum Likelihood Comparison: 1-Step versus 2-Step Approach (all-NIG model)

Graph A of Figure A1 illustrates the maximum log-likelihood for each simulation (simulations sorted by increasing values of the maximum likelihood for better clarity). Graphs B and C display the histograms of the two log-likelihood distributions obtained by the 1-step (Graph B) and the 2-step (Graph C) estimation approaches. Plots are obtained by the simulation of 1,000 samples, each made of 500 observations, for the all-NIG model with 5 components.



idiosyncratic factor can be compared with those in the fifth and seventh columns of Table A3. In particular, we note that in the case of the all-NIG model, the errors obtained with the 2-step procedure, using ML estimation, are in line with those obtained with the 1-step ML approach, which, in principle, computational issues aside, should be the preferred method, exploiting at once all the information contained in the data. However, in the case of the all-MJD model, we observe that the errors of the 2-step procedure are just slightly larger than those obtained with the 1-step ML approach due to the fact that in the 2-step procedure, the univariate estimations are performed via the less efficient EM algorithm.

This is visually confirmed in figures A1 and A2, which report the maximum log-likelihood for each simulation (Graph A), sorting the simulations by increasing values of the ML for better clarity, and the histograms (Graph B) of the two log-likelihood distributions generated by the two estimation methods. From Figure A1, we note that in the all-NIG case, the estimates obtained by means of the 1-step and 2-step procedures lead to very close maximum log-likelihoods. Conversely, Figure A2 shows that for the all-MJD case, the log-likelihoods resulting from the 2-step routine, where the univariate estimations are performed by the EM algorithm, are less close to the ones from the 1-step procedure (i.e., the actual maximum one).

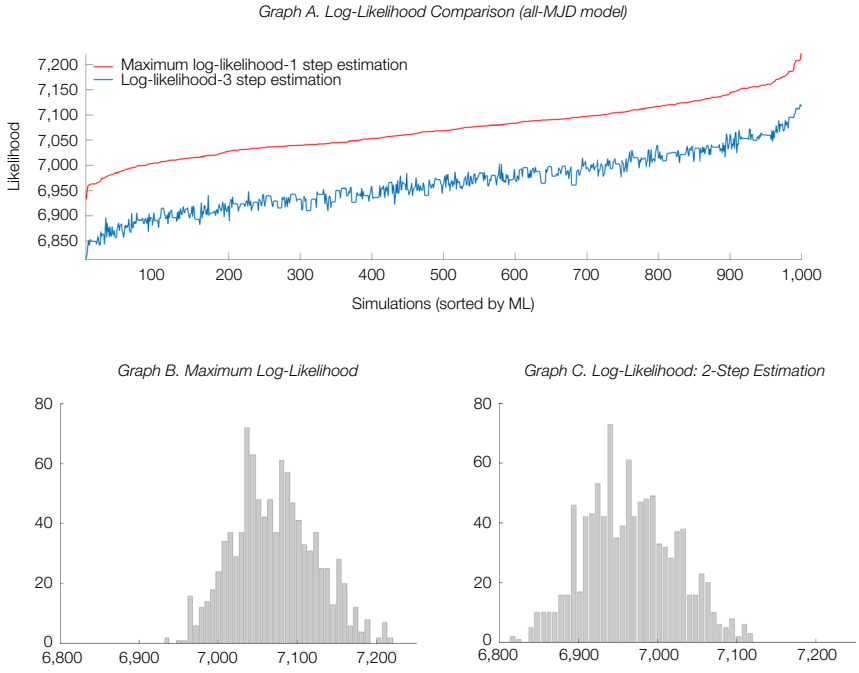
Once more, it is important to notice that the comparison between the 1- and 2-step approaches can only be made for a small number of assets ($N \leq 15$) because the 1-step approach becomes imprecise and noncomputable for larger N.

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

FIGURE A2

Maximum Likelihood Comparison: 1-Step versus 2-Step Approach (all-MJD model)

Graph A of Figure A2 illustrates the maximum log-likelihood for each simulation (simulations sorted by increasing values of the maximum likelihood for better clarity). Graphs B and C display the histograms of the two log-likelihood distributions obtained by the 1-step (Graph B) and the 2-step (Graph C) estimation approaches. Plots are obtained by the simulation of 1,000 samples, each made of 500 observations, for the all-MJD model with 5 components.



Appendix B. Brief Reference to the Literature

In Appendix B, we present a brief reference to the literature closely related to this article. Evidence of pure Lévy jump risk representing a large share of uncertainty in stock returns has been put forward, for example, by Lee and Hannig (2010) and Ornthanalai (2014), among others. In particular, both Lee and Hannig (2010) and Ornthanalai (2014) highlight the role of infinite-activity jumps, that is, jumps of small size occurring with high frequency, which, in principle, could be misidentified as diffusions. These findings stress the need for hedging and risk management strategies equipped to face not just (rare) crash risk alone but also, and most importantly, risks associated with small- and intermediate-sized jumps. This issue is particularly relevant when risk management is targeted for short horizons, such as the ones applied in the current regulatory risk management framework (10 days; see, e.g., Basel (2010)). Indeed, over such short time horizons, the effects of stochastic volatility are in general negligible (mainly due to the diffusive nature of the processes used for the modeling of volatility trends); thus, the jump component of the (log-)returns is relatively more important, as discussed by Aït-Sahalia (2004), for example. This explains our focus on risk management applications. Factor constructions such as the one proposed by Ballotta and Bonfiglioli (2016) have attracted attention mainly due to their simplicity and analytical tractability, which makes them particularly intuitive. Contributions in this direction started with Vašíček (1987) for the case of Brownian motion;

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

alternative constructions for multivariate Lévy processes based on linear transformations have also been put forward by Luciano and Semeraro (2010) and extended in Luciano et al. (2016). However, the focus in these latter contributions is on Lévy processes with explicit representations in terms of subordinated Brownian motions, which are not always available. For a complete literature review, refer to Ballotta and Bonfiglioli (2016), Luciano et al. (2016), and references therein.

References

Ahn, S. C., and A. R. Horenstein. "Eigenvalue Ratio Test for the Number of Factors." *Econometrica*, 81 (2013), 1203–1227.

Ahn, S. C., and M. F. Perez. "GMM Estimation of the Number of Latent Factors: With Application to International Stock Markets." *Journal of Empirical Finance*, 17 (2010), 783–802l.

Aït-Sahalia, Y. "Disentangling Diffusion from Jumps." *Journal of Financial Economics*, 74 (2004), 487–528.

Aït-Sahalia, Y., and J. Jacod. "Testing Whether Jumps Have Finite or Infinite Activity." *Annals of Statistics*, 39 (2011), 1689–1719.

Anderson, T. W. "Maximum Likelihood Estimates for a Multivariate Normal Distribution When Some Observations Are Missing." *Journal of the American Statistical Association*, 52 (1957), 200–203.

Bai, J. "Inferential Theory for Factor Models of Large Dimensions." *Econometrica*, 71 (2003), 135–171.

Bai, J., and S. Ng. "Determining the Number of Factors in Approximate Factor Models." *Econometrica*, 70 (2002), 191–221.

Bai, J., and S. Ng. "Extremum Estimation When the Predictors Are Estimated from Large Panels." *Annals of Economics and Finance*, 9 (2008), 201–222.

Bakshi, G., and G. Panayotov. "First-Passage Probability, Jump Models, and Intra-Horizon Risk." *Journal of Financial Economics*, 95 (2010), 20–40.

Ballotta, L., and E. Bonfiglioli. "Multivariate Asset Models Using Lévy Processes and Applications." *European Journal of Finance*, 22 (2016), 1320–1350.

Ballotta, L.; G. Deelstra; and G. Rayée. "Multivariate FX Models with Jumps: Triangles, Quantos and Implied Correlation." *European Journal of Operational Research*, 260 (2017), 1181–1199.

Ballotta, L., and G. Rayée. "Smiles & Smirks: A Tale of Factors." Working Paper, available at https://ssrn.com/abstract=2980349 (2017).

Barndorff-Nielsen, O. E. "Normal Inverse Gaussian Distributions and Stochastic Volatility Modelling." *Scandinavian Journal of Statistics*, 24 (1997), 1–13.

Basel. "Basel III: A Global Regulatory Framework for More Resilient Banks and Banking Systems." Technical Report, Bank for International Settlements, available at http://www.bis.org/publ/bcbs189.pdf (2010).

Basel. "Regulatory Consistency Assessment Programme (RCAP): Second Report on Risk-Weighted Assets for Market Risk in the Trading Book." Technical Report, Bank for International Settlements, available at http://www.bis.org/publ/bcbs267.pdf (2013).

Boudoukh, J.; M. Richardson; R. Stanton; and R. Whitelaw. "MaxVaR: Long Horizon Value at Risk in a Mark-to-Market Environment." *Journal of Investment Management*, 2 (2004), 1–6.

Carr, P., and L. Wu. "Time-Changed Lévy Processes and Option Pricing." *Journal of Financial Economics*, 71 (2004), 113–141.

Carr, P., and L. Wu. "Stochastic Skew in Currency Options." *Journal of Financial Economics*, 86 (2007), 213–247.

Cont, R., and P. Tankov. *Financial Modelling with Jump Processes*, 2nd ed. Boca Raton, FL: Chapman and Hall/CRC (2004).

Duncan, J.; J. Randal; and P. Thomson. "Fitting Jump Diffusion Processes Using the EM Algorithm." Contributed Talk at the Australasian Meeting of the Econometric Society, Canberra, Australia (2009).

Eberlein, E.; R. Frey; and E. A. von Hammerstein. "Advanced Credit Portfolio Modeling and CDO Pricing." In *Mathematics: Key Technology for the Future*, W. Jager and H. J. Krebs, eds. New York, NY: Springer (2008).

Eberlein, E., and D. Madan. "On Correlating Lévy Processes." *Journal of Risk*, 13 (2009), 3–16.

Efron, B. "Bootstrap Methods: Another Look at the Jackknife." *Annals of Statistics*, 7 (1979), 1–26.

Fang, F., and C. W. Oosterlee. "A Novel Pricing Method for European Options Based on Fourier-Cosine Series Expansions." *SIAM Journal on Scientific Computing*, 31 (2008), 826–848.

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press

Fusai, G.; I. Abrahams; and C. Sgarra. "An Exact Analytical Solution for Discrete Barrier Options." *Finance and Stochastics*, 10 (2006), 1–26.
Fusai, G.; G. Germano; and D. Marazzina. "Spitzer Identity, Wiener-Hopf Factorization and Pricing of Discretely Monitored Exotic Options." *European Journal of Operational Research*, 251 (2016), 124–134.
Glasserman, P. *Monte Carlo Methods in Financial Engineering*. New York, NY: Springer (2004).
Honoré, P. "Pitfalls in Estimating Jump-Diffusion Models." Working Paper, available at https://ssrn.com/abstract=61998 (1998).
Jackson, K.; S. Jaimungal; and V. Surkov. "Fourier Space Time-Stepping for Option Pricing with Lévy Models." *Journal of Computational Finance*, 12 (2008), 1–29.
Kritzman, M., and D. Rich. "The Mismeasurement of Risk." *Financial Analysts Journal*, 58 (2002), 91–99.
Lee, S. S., and J. Hannig. "Detecting Jumps from Lévy Jump Diffusion Processes." *Journal of Financial Economics*, 96 (2010), 271–290.
Loregian, A. "Multivariate Lévy Models: Estimation and Asset Allocation." Ph.D. Thesis, Università degli Studi di Milano-Bicocca (2013).
Luciano, E.; M. Marena; and P. Semeraro. "Dependence Calibration and Portfolio Fit with Factor-Based Time Changes." *Quantitative Finance*, 16 (2016), 1037–1052.
Luciano, E., and P. Semeraro. "Multivariate Time Changes for Lévy Asset Models: Characterization and Calibration." *Journal of Computational and Applied Mathematics*, 233 (2010), 1937–1953.
Merton, R. C. "Option Pricing When Underlying Stocks Are Discontinuous." *Journal of Financial Economics*, 3 (1976), 125–144.
Meucci, A. *Risk and Asset Allocation*. New York, NY: Springer (2005).
Ornthanalai, C. "Lévy Jump Risk: Evidence from Options and Returns." *Journal of Financial Economics*, 112 (2014), 69–90.
Stulz, R. M. "Rethinking Risk Management." *Journal of Applied Corporate Finance*, 9 (1996), 8–25.
Vašíček, O. "Probability of Loss on Loan Portfolio." Memo, KMV Corporation (1987).

https://doi.org/10.1017/S0022109018001321 Published online by Cambridge University Press