

POLITECNICO

MILANO 1863

Financial Engineering

The Big Lévy: Multivariate Pricing

2023-2024

Manfredi Giacomo	CP: 10776946	Mat: 247438
Stillo Francesco	CP: 10698518	Mat: 247230
Toia Nicolò	CP: 10628899	Mat: 247208

Contents

1 Introduction	1
2 Martingale Condition for Forward Exponent Processes	2
2.1 Theoretical Framework	2
2.2 Proof of Martingality and Compensator	3
3 NIG Marginal Characteristics and Parameter Ratio Proof	5
4 Linear Correlation	7
4.1 Theoretical Framework	7
4.2 Problem & Solution	7
5 Dataset description	9
6 Synthetic forwards and Market term structure	10
6.1 S&P 500 and EURO STOXX 50 implicit term structure	10
6.2 Computation and Results	11
6.3 Data Processing: The Volatility Surface	13
7 Historical Correlation	17
8 Calibration of Multi-Variate Processes	18
8.1 Unweighted RMSE	20
8.2 Time Weighted RMSE	21
8.2.1 'Linear' weights	21
8.2.2 'Triangular' weights	22
8.3 Volume Weighted RMSE	23
8.4 Filtered Root Mean Square Error	24
8.5 Variance Gamma	26
8.5.1 A comparison between two subordinators: NIG vs VG	27
8.6 Alternative Calibrations	27
8.6.1 Single Maturity Calibration	27
8.6.2 Maturity Window Calibration	28
8.7 Final Consideration on the Calibration	29
9 Calibration of the Idiosyncratic and Systematic Factors	30
10 Black model	33
10.1 Calibration	33
11 Derivative Pricing	35
11.1 Lévy model	35
11.2 Black model	36
11.3 Semi-closed formula	37
11.4 Comments and Results	38
12 Final consideration on the implemented models	41
13 Appendix A: L. Ballotta & E. Bonfiglioli (2014)	42
14 Appendix B: Some definitions & computations	43
15 Appendix C: Calibrated parameters idiosyncratic and systematic NIG and VG	44
16 Appendix D: Proof of the semi-closed formula for the derivative	45
17 Appendix E: Python Code criticalities	49
18 References	50

1. Introduction

The aim of this Case Study is to price a financial derivative contract written on two underlying assets describing the dynamics of two indexes: S&P 500 and EURO-STOXX 50. In order to accomplish this objective a robust method for Multivariate Lévy Processes with dependence between components is introduced. The relevance of considering a multidimensional asset model based on Lévy processes is identified in its capability, outperforming the standard Gaussian framework, of detecting market shocks.

This project is conducted by exploring the studies conducted by L. Ballotta & E. Bonfiglioli (2014) [B&B] regarding details on the implementation of a multivariate pricing model based on Lévy processes and its calibration to match market data.

Moreover, to analyze the importance of the term structure of interest rates in derivative pricing, a model based on the liquid exchange traded derivatives is introduced to better capture the funding risk and the risk of default. The idea, explained by M. Azzone & R. Baviera (2020), is to resort on the use of a synthetic forward to compute market discount factors and forward prices (constructed to be independent w.r.t the strikes).

In the first part of this letter a theoretical argumentation is provided in order to describe the framework starting from proof of the martingality property of the forward exponent of both underlying assets. Subsequently a more specific framework is reached by exploiting the characteristic function of a Lévy NIG (Normal Inverse Gaussian) process and deriving the formula of the correlation between the two assets.

Then, after a brief data exploration to better understand the dataset composition, the focus shifts to calculating discount factors and interest rates from market data.

One more ingredient is needed in order to proceed with model calibration: this is the historical correlation between the annual returns of the two markets. After computing it, having all the necessary ingredients, the issue of model calibration is addressed. Various approaches to the objective function in the optimization problem are investigated to determine the parameters of the NIG distributions that characterize the two assets: different alternatives on weights and error calculation are tried with the aim of finding the best performing and most robust system among all.

In addition, the performance of the model with NIG subordinator is compared with that of a model with Variance Gamma (VG) subordinator, to get an indication of its goodness-of-fit.

Then, strong in the parameters found, the calibration of the parameters constituting the idiosyncratic and systemic components of the two assets is carried out.

In the last part of the paper, a new framework relying on the Black dynamics of the two assets is introduced and again the parameters of that model are calibrated.

Finally, the Lévy model with NIG subordinator, the Lévy model with VG subordinator and the Black model are compared and tested to price the derivative. The derivation of a semi-closed formula for the price of the derivative allows confirmation on what has been worked out.

Disclaimer: After checking historical values of the EURO-STOXX 50, we assume prices for the two indexes are in €. However, the validity of the discussion is not affected by the choice of the currency.

2. Martingale Condition for Forward Exponent Processes

In this passage, it is asked to find the constant p_i such that the processes $e^{X_i(t)+p_i t}$ are martingales, assuming that $X_i(t) = Y_i(t) + a_i Z(t)$ with $Y_i(t)$ and $Z(t)$ characterized by Normal Inverse Gaussian marginals.

2.1. Theoretical Framework

To conduct the study it is crucial to describe both assets with respect to "Multivariate Lévy processes" framework. Lévy processes can be defined as stochastic processes characterized by independent and stationary increments. They can be fully described by their characteristic function:

$$\begin{aligned}\phi(u; t) &= e^{t\varphi(u)}, \quad u \in \mathbb{R} \\ \varphi(u) &= iu\alpha - \frac{u^2\sigma^2}{2} + \int_{\mathbb{R}} (e^{iux} - 1 - iux\mathbb{1}_{(|x|<1)})\Pi(dx)\end{aligned}$$

The triple denoted as (α, σ, Π) , known also as characteristic triple of Lévy, holds significant implications for understanding the dynamics of these stochastic processes.

Firstly, $\alpha \in \mathbb{R}$, describes the drift of the process. Secondly $\sigma > 0$ akin to the notion of volatility in traditional financial models, quantifies the extent of dispersion, offering insights into the degree of unpredictability inherent in the process. Lastly, the characteristic triple is completed by the Lévy measure which describes the distribution of the process' jumps.

Describing the two assets using Lévy theory becomes more effective if some dependant components are introduced. In order to do that, the property of in-variance under linear transformation, which holds for this type of processes, is applied. The main result is explained in Proposition 1 of **B&B** [13.1]. Here it is important to mark that it is possible to decompose a Lévy process in a linear combination of independent Lévy processes

$$\mathbf{X}(t) = (X_1(t), \dots, X_n(t))^T = (Y_1(t) + a_1 Z(t), \dots, Y_n(t) + a_n Z(t))^T$$

bringing to the following characteristic function:

$$\phi_{\mathbf{X}}(\mathbf{u}; t) = \phi_Z \left(\sum_{j=1}^n a_j u_j; t \right) \prod_{j=1}^n \phi_{Y_j}(u_j; t), \quad \mathbf{u} \in \mathbb{R}^n$$

Describing the forward prices for the two assets as

$$F_i(t, T) = F_i(0, T)e^{(X_i(t)+p_i t)} \quad i = 1, 2$$

where $X_i(t)$ is a 2-dimensional Lévy process defined according to the above proposition, brings to an interpretation that is both simple and essential: by decomposing each asset $X_i(t)$ into two components, $Z(t)$ and $Y_i(t)$, Z can be considered as the part representing the common systemic risk inherent in the market, while Y_i can be seen as identifying the idiosyncratic shock specific to the individual asset.

Moreover, this procedure allows to recover, via numerical approach based on the knowledge of the characteristic function (see Prop1 **B&B** 2014), the jointly distribution of $X(t)$ if this is not known analytically. So, the choice of the distributions of $Y_i(t)$ and $Z(t)$ can affect significantly the model dependence structure.

The multidimensional modelling approach explained in the **B&B** letter makes it feasible to specify any uni-variate distribution for $Y_i(t)$ and $Z(t)$. On the other hand, for any chosen marginal distribution of the starting process $X_i(t)$, convolution conditions can be applied to the processes $Y_i(t)$ and $Z(t)$ so that their linear combination, $Y_i(t) + a_i Z(t)$, matches the distribution of $X(t)$. This method is effective for modeling financial assets using data from traded vanilla options. Since Multi-name derivatives contracts may present low liquidity, convolution enables the independent calibration of the marginal distribution without relying on the correlation matrix. This means the parameters for the idiosyncratic and systematic processes can be determined separately from the correlation fitting.

To address the problems described in this project we follow the line described in **B&B**, in which processes $X_i(t)$ (and consequently $Y_i(t)$ and $Z(t)$ thanks to the matching of distributions) are described as subordinated BM's. By recalling the definition illustrated in section 2.2 of **B&B** [13.2] it is possible to Explicit $X_i(t)$ as follows:

$$X_i(t) = \begin{cases} \xrightarrow{\text{Def.}} Y_i(t) + a_i Z(t), & a_i \in \mathbb{R} \\ \xrightarrow{sBM} \theta_i G_i(t) + \sigma_i W(G_i(t)) & \text{where } W = \{W(t)\}_t \text{ is a BM and } G_i(t) \text{ is the subordinator} \end{cases}$$

and consequently:

$$\begin{cases} Y_i(t) = \beta_i G_{Y_i}(t) + \gamma_i W(G_{Y_i}(t)) & \text{where } \beta_i \in \mathbb{R}, \gamma_i > 0, i = 1, 2 \\ Z(t) = \beta_Z G_{Y_Z}(t) + \gamma_Z W(G_{Y_Z}(t)) & \text{where } \beta_Z \in \mathbb{R}, \gamma_Z > 0 \end{cases}$$

The importance of building subordinated Lévy processes lies mainly in two factors: the uneven cadence with which the stock market is timed and the influence that news (which comes out at unpredictable time intervals) has on stock prices.

A class of subordinators that is widely used in financial modeling and that are consequently explored in this discussion are the so-called tempered stable subordinators, which have characteristic exponent (expressed in terms of the subordinated BM $X_i(t)$):

$$\varphi_G(u) = \frac{\alpha - 1}{\alpha k_i} \left[\left(1 - \frac{i u k_i}{1 - \alpha} \right)^\alpha - 1 \right], \quad \alpha \in [0, 1), \quad u \in \mathbb{R}$$

(Similarly for $Y_i(t)$, $Z(t)$ taking ν_i and ν_Z instead of k_i).

So, $X_i(t)$ can be seen as function of $(\sigma_i, k_i, \theta_i)$ tern of parameters if expressed as a sBM, but also as function of $(\gamma_i, \nu_i, \beta_i, a_i, \gamma_Z, \nu_Z, \beta_Z)$ group of parameters if expressed as linear combination of two subordinated BMs.

2.2. Proof of Martingality and Compensator

It is asked to find the constant p_i such that the processes $e^{X_i(t)+p_i t}$ (part of the forward price formula for the asset i at time t) are martingales. This must be solved in a framework such that $Y_i(t)$ and $Z(t)$, the two independent processes that compose $X_i(t)$, have NIG marginals (i.e. the tempered stable parameter alpha in the characteristic exponent of the corresponding G's is set equal to $\frac{1}{2}$).

Requiring that the martingality property holds for $e^{X_i(t)+p_i t}$ is equivalent to ask that:

$$\mathbb{E}_0[e^{X_i(t)+p_i t}] = e^{X_i(0)+p_i 0} \stackrel{*}{=} 1 \quad \iff \quad e^{-p_i t} = \mathbb{E}_0[e^{X_i(t)}]$$

Note *: It is true since $X_0 = 0$ a.s. for Lévy processes.

In order to compute p_i it is necessary to start with the calculation of $\mathbb{E}_0[e^{X_i(t)}]$. Thanks to probability theory, the computation of such expected value can be easily translated into the evaluation of the characteristic function of $X_i(t)$ at a specific point:

$$\phi_{X_i}(u_i; t) = \mathbb{E}[e^{i \langle x, u_i \rangle}] \quad \Rightarrow \quad \phi_{X_i}(-i; t) = \mathbb{E}[e^x]$$

where by definition, the marginal characteristic function of the process $X_i(t)$ is equal to the characteristic function of the process $\mathbf{X}(t) = (X_1(t), X_2(t))$ evaluated in $\mathbf{u} = (u_1, 0)$ (for $X_1(t)$) or $\mathbf{u} = (0, u_2)$ (for $X_2(t)$). Therefore, it is important to derive the expression of the characteristic function of process $\mathbf{X}(t)$.

To do this, we can go back to what is stated in Proposition 1 of **B&B** [13.1]: given that $\mathbf{X}(t)$ depends on $X_i(t)$ consisting of the linear combination of $Y_i(t)$ and $Z(t)$ (which have NIG marginals), it is possible to derive the characteristic function of $\mathbf{X}(t)$ known the NIG marginal characteristic functions of $Y_i(t)$ and $Z(t)$. The passages follow:

By definition of subordinated BM's:

$$\begin{cases} \Phi_{Y_i}(u, t) = e^{t \varphi_{G_{Y_i}}(u \beta_i + i u^2 (\gamma_i^2 / 2))}, & u \in \mathbb{R}, \quad i = 1, 2 \\ \Phi_Z(u, t) = e^{t \varphi_Z(u \beta_Z + i u^2 (\gamma_Z^2 / 2))}, & u \in \mathbb{R} \end{cases}$$

with characteristic exponents following NIG distribution ($\alpha = \frac{1}{2}$):

$$\begin{cases} \varphi_{G_{Y_i}}(u) = \frac{1}{\nu_i} [1 - \sqrt{1 - 2 i u \nu_i}], & u \in \mathbb{R}, \quad i = 1, 2 \\ \varphi_Z(u) = \frac{1}{\nu_Z} [1 - \sqrt{1 - 2 i u \nu_Z}], & u \in \mathbb{R} \end{cases}$$

The characteristic function of $\mathbf{X}(t)$ becomes:

$$\phi_{\mathbf{X}}(\mathbf{u}; t) \stackrel{13.1}{=} e^{\frac{t}{\nu_Z} (1 - \sqrt{1 - 2 i \beta_Z \nu_Z (a_1 u_1 + a_2 u_2) + (a_1 u_1 + a_2 u_2)^2 \nu_Z \gamma_Z^2})} \phi_{Y_1}(u_1; t) \phi_{Y_2}(u_2; t), \quad \mathbf{u} = (u_1, u_2) \in \mathbb{R}^2$$

Now, since an explicit expression for the characteristic function of $\mathbf{X}(t)$ has been derived, it can be evaluated in one of the two input variables forcing the other one zero in order to find the marginal characteristic function of $X_i(t)$:

$$\phi_{X_i}(u_i; t) = e^{\frac{t}{\nu_Z}(1-\sqrt{1-2i\beta_Z\nu_Z(a_i u_i)+(a_i u_i)^2\nu_Z\gamma_Z^2})+\frac{t}{\nu_i}(1-\sqrt{1-2i\beta_i\nu_i u_i+u_i^2\nu_i\gamma_i^2})}, \quad u_i \in \mathbb{R}, \quad i = 1, 2$$

By evaluating the marginal characteristic function of $X_i(t)$ in "-i" it is possible to compute the needed expected value:

$$\begin{aligned} \mathbb{E}_0[e^{X_i(t)}] &= \phi_{X_i}(-i; t) = e^{\frac{t}{\nu_Z}(1-\sqrt{1-2\beta_Z\nu_Z a_i - a_i^2\nu_Z\gamma_Z^2})} \phi_{Y_i}(-i; t), \quad u_i \in \mathbb{R}, \\ &= e^{\frac{t}{\nu_Z}(1-\sqrt{1-2\beta_Z\nu_Z a_i - a_i^2\nu_Z\gamma_Z^2})} e^{\frac{t}{\nu_{Y_i}}(1-\sqrt{1-2\beta_{Y_i}\nu_{Y_i} - \nu_{Y_i}\gamma_{Y_i}^2})} \end{aligned}$$

and so the expression of p_i (i.e. the "drift compensator") that guarantees the martingality of the required process can be found as follow:

$$\begin{aligned} e^{-p_i t} &= \mathbb{E}_0[e^{X_i(t)}] = \phi_{X_i}(-i; t), \quad i = 1, 2 \\ \iff e^{-p_i t} &= e^{\frac{t}{\nu_Z}(1-\sqrt{1-2\beta_Z\nu_Z a_i - a_i^2\nu_Z\gamma_Z^2})} e^{\frac{t}{\nu_{Y_i}}(1-\sqrt{1-2\beta_{Y_i}\nu_{Y_i} - \nu_{Y_i}\gamma_{Y_i}^2})} \\ \xleftarrow{\ln(\cdot)} -p_i t &= \frac{t}{\nu_Z}(1 - \sqrt{1 - 2\beta_Z\nu_Z a_i - a_i^2\nu_Z\gamma_Z^2}) + \frac{t}{\nu_{Y_i}}(1 - \sqrt{1 - 2\beta_{Y_i}\nu_{Y_i} - \nu_{Y_i}\gamma_{Y_i}^2}) \\ \xrightarrow{-\frac{1}{t}} p_i &:= -\frac{1}{\nu_Z}(1 - \sqrt{1 - 2\beta_Z\nu_Z a_i - a_i^2\nu_Z\gamma_Z^2}) - \frac{1}{\nu_{Y_i}}(1 - \sqrt{1 - 2\beta_{Y_i}\nu_{Y_i} - \nu_{Y_i}\gamma_{Y_i}^2}), \quad i = 1, 2 \end{aligned}$$

3. NIG Marginal Characteristics and Parameter Ratio Proof

As recalled in the introductory paragraph of **B&B** and at the beginning of this document, the characteristic function of a Lévy process $X(t)$ can be written generically as:

$$\phi_X(u; t) = e^{t\varphi_X(u)}, \quad u \in \mathbb{R}$$

The explicit expression of the characteristic function of $X_i(t)$ was found in the previous section. Here it is simply recalled:

$$\phi_{X_i}(u; t) = e^{\frac{t}{\nu_Z} (1 - \sqrt{1 - 2i\beta_Z\nu_Z(a_i u) + (a_i u)^2\nu_Z\gamma_Z^2}) + \frac{t}{\nu_i} (1 - \sqrt{1 - 2i\beta_i\nu_i u + u^2\nu_i\gamma_i^2})}, \quad u \in \mathbb{R}, \quad i = 1, 2$$

While, as already expressed in the previous portion, starting from any distribution of $X(t)$ one can apply the convolution conditions and obtain a linear combination $Y(t) + aZ(t)$ that has the same distribution as X, the opposite cannot be stated with certainty. In fact, in the case considered, starting from two independent levy processes with NIG marginals, one obtains a linear combination $X_i(t) = Y_i(t) + aZ(t)$ with a characteristic function in general different from the NIG one. Is it possible to obtain the characteristic function of a NIG process starting from the general case? Let's try to force equivalence.

First, the expression of the characteristic function a process with marginal NIG is invoked:

$$\phi_{X_{NIG}}(u; t) = e^{\frac{t}{k_X} (1 - \sqrt{1 - 2i\theta_X k_X u + u^2 k_X \sigma_X^2})}$$

Since the characteristic function of the NIG has a single term of the form $1 - \sqrt{1 - (...)}$, it is necessary to assume that the terms in parentheses are equal in order to be able to collect:

$$\begin{aligned} \sqrt{1 - 2i\beta_Z\nu_Z(a_i u) + (a_i u)^2\nu_Z\gamma_Z^2} &= \sqrt{1 - 2i\beta_i\nu_i u + u^2\nu_i\gamma_i^2}, \quad i = 1, 2 \\ \Downarrow \\ \begin{cases} -2i\beta_Z\nu_Z a_i u = -2i\beta_i\nu_i u \\ a_i^2 u^2 \nu_Z \gamma_Z^2 = u^2 \nu_i \gamma_i^2 \end{cases} &\Downarrow \\ \begin{cases} \beta_Z \nu_Z a_i = \beta_i \nu_i \\ a_i^2 \nu_Z \gamma_Z^2 = \nu_i \gamma_i^2 \end{cases} & \end{aligned}$$

The characteristic function becomes:

$$\phi_{X_i}(u; t) = e^{\left(\frac{t}{\nu_Z} + \frac{t}{\nu_i}\right)(1 - \sqrt{1 - 2i\beta_Z\nu_Z(a_i u) + (a_i u)^2\nu_Z\gamma_Z^2})}, \quad u \in \mathbb{R}, \quad i = 1, 2$$

By matching this expression of the characteristic function of $X_i(t)$ with that feature of the NIG processes the following relations are obtained:

$$\begin{cases} \theta_i k_i = \beta_Z \nu_Z a_i \\ k_i \sigma_i^2 = a_i^2 \nu_Z \gamma_Z^2, \\ \frac{1}{k_i} = \frac{1}{\nu_Z} + \frac{1}{\nu_i} \implies k_i = \frac{\nu_Z \nu_i}{\nu_Z + \nu_i} \end{cases} \quad i = 1, 2$$

Note: The first two equations are exactly the ones called "Equations [9]" reported in section 2.2 on the **B&B** paper.

Thanks to this representation, it is possible to express the process $X(t)$, initially described by seven parameters ($\gamma_i, \nu_i, \beta_i, a_i, \gamma_Z, \nu_Z, \beta_Z$), as a function of only three typical parameters of the characteristic function NIG: (σ_i, k_i, θ_i).

It is possible to derive a relationship between the two sets of parameters by combining the expressions found above:

$$\begin{cases} a_i \beta_Z = \theta_i \frac{\nu_i}{\nu_i + \nu_Z} \implies \theta_i = (1 + \frac{\nu_Z}{\nu_i}) a_i \beta_Z = a_i \beta_Z + \underbrace{\frac{\nu_Z}{\nu_i} a_i \beta_Z}_{\beta_i} \\ a_i^2 \gamma_Z^2 = \frac{\nu_i}{\nu_i + \nu_Z} \sigma_i^2 \implies \sigma_i^2 = (1 + \frac{\nu_Z}{\nu_i}) a_i^2 \gamma_Z^2 = a_i^2 \gamma_Z^2 + \underbrace{\frac{\nu_Z}{\nu_i} a_i^2 \gamma_Z^2}_{\gamma_i^2} \\ \frac{1}{k_i} = \frac{1}{\nu_Z} + \frac{1}{\nu_i} \implies k_i = \frac{\nu_Z \nu_i}{\nu_Z + \nu_i} \end{cases}$$

Within the context of the previously mentioned "Equations 9", it is possible to introduce a new constraint that wouldn't be universally applicable. However, in this specific scenario, this constraint becomes a crucial factor and its impact will be noticed when facing the calibration:

$$\begin{aligned}
 \begin{cases} \theta_i k_i = \beta_Z \nu_Z a_i \\ k_i \sigma_i^2 = a_i^2 \nu_Z \gamma_Z^2 \end{cases} &\iff \begin{cases} a_i = \frac{k_i \theta_i}{\beta_Z \nu_Z} & (a) \\ \nu_Z \gamma_Z^2 = \frac{k_i \sigma_i^2}{a_i^2} & (b) \end{cases} \quad i = 1, 2 \\
 \stackrel{(b)}{\implies} \quad \frac{k_i \sigma_i^2}{a_i^2} &= \nu_Z \gamma_Z^2 \\
 \stackrel{(a)}{\implies} \quad \frac{k_i \sigma_i^2}{k_i^2 \theta_i^2} \beta_Z^2 \nu_Z^2 &= \nu_Z \gamma_Z^2 \\
 \implies \quad \frac{\sigma_i^2}{k_i^2 \theta_i^2} &= \frac{\gamma_Z^2}{\nu_Z^2 \beta_Z^2} = c \quad i = 1, 2
 \end{aligned}$$

Observation: Now that more stringent assumptions about the model considered have been made and new relationships have been found, it is possible to rewrite the "drift compensator" in a more compact form that takes into account what has just been discussed:

$$\begin{aligned}
 p_i &= -\frac{1}{\nu_Z} (1 - \sqrt{1 - 2\beta_Z \nu_Z a_i - a_i^2 \nu_Z \gamma_Z^2}) - \frac{1}{\nu_{Y_i}} (1 - \sqrt{1 - 2\beta_{Y_i} \nu_{Y_i} - \nu_{Y_i} \gamma_{Y_i}^2}), \quad i = 1, 2 \\
 &= -\frac{1}{\nu_Z} (1 - \sqrt{1 - 2\beta_Z \nu_Z a_i - a_i^2 \nu_Z \gamma_Z^2}) - \frac{1}{\nu_{Y_i}} (1 - \sqrt{1 - 2\beta_Z \nu_Z a_i - a_i^2 \nu_Z \gamma_Z^2}) \\
 &= -\left(\frac{1}{\nu_Z} + \frac{1}{\nu_{Y_i}}\right) (1 - \sqrt{1 - 2\beta_Z \nu_Z a_i - a_i^2 \nu_Z \gamma_Z^2}) \\
 &= -\frac{1}{k_i} (1 - \sqrt{1 - 2\theta_i k_i - k_i \sigma_i^2})
 \end{aligned}$$

4. Linear Correlation

The following section explores the linear correlation between the two processes, following the paper by **B&B**. Moreover, a discussion about the maximum correlation obtainable within this framework is provided.

4.1. Theoretical Framework

By exploring the **B&B** paper it is possible to find the expression for the correlation coefficient of the multivariate Lévy process $\mathbf{X}(t) = (X_1(t), \dots, X_n(t))^T = (Y_1(t) + a_1 Z(t), \dots, Y_n(t) + a_n Z(t))$:

$$\rho_{jl}^{\mathbf{X}} = \text{Corr}(X_j(t), X_l(t)) = \frac{a_j a_l \text{Var}(Z(1))}{\sqrt{\text{Var}(X_j(1))} \sqrt{\text{Var}(X_l(1))}}$$

This expression, dated 2014, has several new and improved features compared to studies in the same area (Multivariate asset models using Lévy processes) prior to it. For instance, in the article by Luciano & Semeraro (2010), it is constructed a model that gives birth to a formula for the correlation index characterized by some limitations such as:

1. The model's ability to achieve high correlation is limited by constraints dependent on marginal parameters.
2. The correlation of the multivariate process is capped by the correlation of the subordinators.
3. The correlation index may be zero even if the processes are correlated.
4. The sign of the correlation index between the j-th and the l-th processes depends solely on the product of the averages, so it is not possible to capture negative and positive correlations simultaneously for given margins.

First, the **B&B** formula puts a remedy to the limitation 3. expressed above. In fact, the correlation index $\rho_{jl}^{\mathbf{X}}$ can be worth zero if and only if $a_j a_l = 0$ i.e. X_j and X_l are independent because at least one of them has no Z component, or if $\text{Var}(Z(1)) = 0$ i.e. $Z(t)$ follows a degenerate distribution [14.1]. Second, a remedy is also made for the limits 1. and 2.. Indeed, in the case where Y_l and Y_j have degenerate distributions [$\text{Var}(Y_l(1)) = 0$ and $\text{Var}(Y_j(1)) = 0$] it can be observed that the correlation index takes value 1:

$$\begin{aligned} \rho_{jl}^{\mathbf{X}} &= \text{Corr}(X_j(t), X_l(t)) = \\ &= \frac{a_j a_l \text{Var}(Z(1))}{\sqrt{\text{Var}(Y_j(1) + a_j Z(1))} \sqrt{\text{Var}(Y_l(1) + a_l Z(1))}} = \\ &= \frac{a_j a_l \text{Var}(Z(1))}{\sqrt{a_j^2 \text{Var}(Z(1)) a_l^2 \text{Var}(Z(1))}} = \\ &= 1 * \text{sign}(a_j a_l) \end{aligned}$$

Finally, the fourth limit is also bypassed. In fact, $\text{sign}(\rho_{jl}^{\mathbf{X}}) = \text{sign}(a_j a_l)$. The value of the correlation index is therefore free from sign and can take either positive or negative values.

4.2. Problem & Solution

The goal is to calculate the linear correlation between the two Lévy processes $X_1(t) = Y_1(t) + a_1 Z(t)$ and $X_2(t) = Y_2(t) + a_2 Z(t)$ with properties and parameters consistent with what has been described above. To apply the correlation index formula, it is necessary to calculate the variances of processes $X_1(t)$, $X_2(t)$ and $Z(t)$ evaluated at $t = 1$. Recall that the three processes follow the NIG law, distribution for which the variance is known (see Financial modelling for jump processes, Cont-Tankov, TABLE 4.5):

$$\text{Given } X_t \sim \text{NIG}, \quad \text{Var}(X_t) = \sigma^2 t + \theta^2 k t$$

It is therefore possible to easily calculate the desired variances:

$$\begin{cases} \text{Var}(X_1(1)) = [\sigma_1^2 t + \theta_1^2 k_1 t] \Big|_{t=1} = \sigma_1^2 + \theta_1^2 k_1 \\ \text{Var}(X_2(1)) = [\sigma_2^2 t + \theta_2^2 k_2 t] \Big|_{t=1} = \sigma_2^2 + \theta_2^2 k_2 \\ \text{Var}(Z(1)) = [\gamma_Z^2 t + \beta_Z^2 \nu_Z t] \Big|_{t=1} = \gamma_Z^2 + \beta_Z^2 \nu_Z \end{cases}$$

Using these expressions, the correlation index formula between the two processes $X_1(t)$ and $X_2(t)$ can be derived:

$$\rho_{12}^{\mathbf{X}} = \frac{a_1 a_2 \text{Var}(Z(1))}{\sqrt{\text{Var}(X_1(1))} \sqrt{\text{Var}(X_2(1))}} \stackrel{[14.2]}{=} \text{sign}(a_1 a_2) * \sqrt{\frac{\nu_1 \nu_2}{(\nu_1 + \nu_Z)(\nu_2 + \nu_Z)}}$$

The formula found is perfectly consistent with what might be expected from the problem described. In fact, the correlation of the two assets $X_1(t)$ and $X_2(t)$, both described by an idiosyncratic component Y and a systemic component Z, changes its behavior according to the value taken by ν i.e., the variances of the subordinators. When $\nu_Z = 0$, i.e. the variance of the subordinator of the systemic component becomes deterministic, the correlation between the two assets is maximum ($\rho = |1|$) since both are exposed to the same fixed market risk. The sign of the correlation depends on the product of the coefficients a_1 and a_2 . Conversely, when either $\nu_Y = 0$, this means that the corresponding asset is deterministically exposed in its idiosyncratic component, so its exposition to the market risk, and thus its correlation with the other asset, becomes irrelevant.

Note: Although theoretically the correlation index can vary between -1 and 1 with appropriate parameter choices, in the system considered there are limits imposed by the calibration of the model and thus by the parameters that are found. Without going into the details of the implementation of the calibration (which will be addressed later) it is sufficient to know now that through it the following parameters are found: $(\sigma_1.k_1.\theta_1, \sigma_2.k_2.\theta_2)$. Recalling one of the possible forms of the correlation index formula obtained during the derivation of the final formula, it is possible to express ρ as a function of the calibrated parameters:

$$\rho_{12}^X = \frac{\sqrt{k_1 k_2}}{\nu_Z} * sign(a_1 a_2)$$

It is important to note that in the model even the choice of ν_Z is not free, in fact it is calibrated with the objective of matching the historical correlation between the two considered markets.

As a result, the maximum ρ value obtainable in practice is strictly a function of model calibration

5. Dataset description

In this study, we analyze a dataset containing options data for the *S&P500* and *EURO – STOXX50* indices, specifically focusing on the volatility surface as of July 9, 2023. The dataset includes expiry dates, bid and ask prices for both call and put options, and strike prices for all liquid options. Additionally, the dataset contains volumes and spot prices for both markets.

An additional dataset provides the returns of both indices over the past 10 years, with various time intervals. This data is useful for estimating the historical correlation between the two markets.¹

Our pre-processing criteria are based on liquidity measures. We ensure that the data includes maturities with at least three strikes and that “penny options”, which are options priced very low, are filtered out. Additionally, we verify that options with wide bid-ask spreads, indicative of low liquidity, are excluded.

In order to provide a comprehensive understanding of the dataset, we present a simple overview of key metrics for both the EURO-STOXX 50 Index Options and the S&P 500 Index Options. These metrics include the mean, median, standard deviation, and selected quantiles for strikes, call prices, and put prices (mid prices are taken into consideration).

Metric	Strikes	CallPrices	PutPrices
Mean	3789.6	259.96	45.545
Median	3825	219.5	3.9
Standard Deviation	723.08	231.64	77.557
Quantile 0.05	2500	0.7075	0.6
Quantile 0.95	4850	654.67	223.38

Table 1: Dataset Overview for EURO-STOXX 50 Index Options

Metric	Strikes	CallPrices	PutPrices
Mean	3906.7	399.28	27.865
Median	4000	363.6	1.2125
Standard Deviation	1075.4	328.65	57.568
Quantile 0.05	2000	0.555	0.24
Quantile 0.95	5400	973.68	180.93

Table 2: Dataset Overview for S&P 500 Index Options

¹We report a possible oversight in the Semianually data provided in the returns dataset. Indeed, the correlation returns 1, perfectly correlated markets.

6. Synthetic forwards and Market term structure

In this section, we delve into the significance of the term structure of interest rates within the derivative market, which determines the discount rate for expected payoffs. Our focus centers on the equity market and we ask which is the term structure utilized when dealing with liquid exchange-traded derivatives.

To address this query, we draw upon the insights of R. Baviera & M. Azzone (2020), who introduced the concept of an implied discount factor. It is pertinent to note that in the equity market, interest rates are typically not risk-free. Two fundamental risks arise when handling contingent claims: funding risk and counterparty default risk. However, when dealing with exchange-traded derivatives, the latter risk can often be disregarded due to the presence of a clearinghouse with margin calls.

Presently, the Overnight Index Swap (OIS) has emerged as a potential candidate for the risk-free curve used in derivative discounting. The construction of this curve follows the classical pre-crisis methodology (Ron, 2000). Market makers in exchange-traded derivative markets rely on the OIS curve as a benchmark for discounting future cash flows. Unlike traditional LIBOR-based curves, the OIS curve is derived from overnight index swaps, reflecting the market's expectations of future overnight interest rates. Nevertheless, adjustments are made to incorporate spreads, accounting for additional risks and costs not captured by the "risk-free" rate implied by the OIS curve. By dynamically adjusting spreads based on market conditions, market makers aim to ensure profitability while maintaining liquidity and efficiency in derivative markets.

Therefore, the construction of a measure for option prices to assess funding costs is crucial for market makers and higher-level managerial decision-making within financial firms. This can be achieved by considering fundamental concepts such as put-call parity, enabling the derivation of implicit interest rates.

6.1. S&P 500 and EURO STOXX 50 implicit term structure

In this section, we briefly introduce the methodology applied to compute the implicit discount factors using only call and put prices. First, we recall the put-call parity for European options, which relies on the non-arbitrage condition, at value date t_0 and a fixed maturity T :

$$C(K) - P(K) = \bar{B}(t_0, T) \cdot (F - K)$$

where $C(K)$ and $P(K)$ are respectively the European call and put options for a given strike K , while F is the forward price and $\bar{B}(t_0, T)$ is the market discount factor between t_0 and T .

We can now imitate a forward position by using call and put options with the same strike and maturity, the so-called synthetic forward $\mathcal{G}(K)$. To compute the synthetic forward, we rely on bid and ask option prices for each strike and maturity, and then we find:

$$\begin{cases} \mathcal{G}^{\text{bid}}(K) = C^{\text{bid}}(K) - P^{\text{ask}}(K) \\ \mathcal{G}^{\text{ask}}(K) = C^{\text{ask}}(K) - P^{\text{bid}}(K) \\ \mathcal{G}(K) = \frac{\mathcal{G}^{\text{bid}}(K) + \mathcal{G}^{\text{ask}}(K)}{2} \end{cases}$$

We can estimate the market discount factor by considering the angular coefficient in the linear regression:

$$\mathcal{G}_i = -\bar{B}(t_0, T)K_i + \bar{B}(t_0, T)F + \epsilon_i \quad i = 1, \dots, N$$

Hence, we can introduce its least squares estimator:

$$\bar{B}(t_0, T) = -\frac{\sum_{i=1}^N (K_i - \hat{K}) \cdot (\mathcal{G}_i - \hat{\mathcal{G}})}{\sum_{i=1}^N (K_i - \hat{K})^2}$$

where $\hat{\mathcal{G}}$ and \hat{K} are the respective means.

Finally, we can compute the forward prices by leveraging the synthetic forward and the market discounts:

$$F = \frac{\mathcal{G}(K)}{\bar{B}(t_0, T)} + K$$

The main idea expressed is that the market-implied discount factor $\bar{B}(t_0, T)$ is the unique factor that ensures the forward price remains constant regardless of the strike K .

6.2. Computation and Results

Following the technique outlined above, we proceeded to compute the interest rate term structure. As a first step, we computed the synthetic forward using the available bid and ask data for call and put options. Hence, we derived the market-implied discount factors via the formula provided above. Consequently, we computed the forward prices (bid, ask, and mid prices) and plotted them for each strike, against the strikes, and checked for their independence from the strikes.

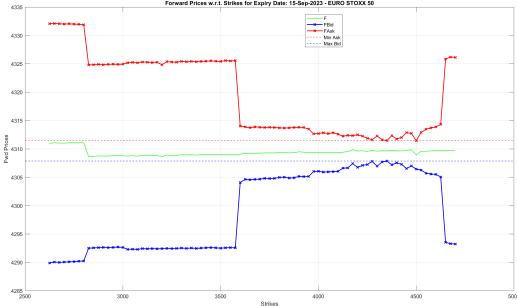


Figure 1: EURO STOXX 50

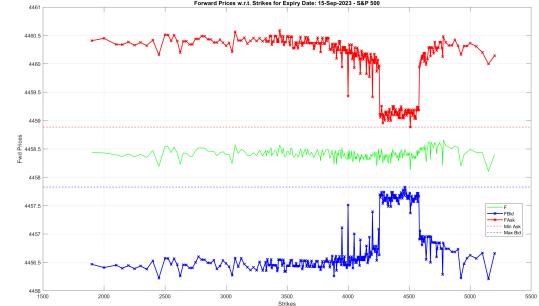


Figure 2: S&P 500

As we can see, the forward prices are independent of the strike K . Moreover, we notice that the length of the bid-ask interval changes with the strikes, highlighting different liquidity of the derivatives for different strikes. We note that for this particular maturity, both markets present a window where the bid-ask forward price is narrower around the respective spot price ($\text{Spot}_{EU} = 4286.6\text{\euro}$ & $\text{Spot}_{US} = 4424.5\text{\euro}$). Here, the interval of the European market seems to be more liquid, although analyzing the dataset, we see that the American market presents more strikes for this maturity, leading the graph to have several spikes.

Moreover, we can analyze the term structure of the two markets. As shown in the figure below, we observe a monotone decreasing curve for the market discounts in both cases. Examining the time window for which we have data for both markets, we can see that the zero rate curve has a similar shape, highlighting similar expectations for future rates movements. However, the US market presents an upward-shifted curve compared to the European market. This is due to the current monetary policies implemented by the two central banks. Indeed, the more resilient economy of the American market has led the Federal Reserve to raise rates more aggressively to tame inflation than the ECB has had to do. Additionally, thanks to the longer maturity provided in the S&P dataset, we can see that the market is expecting some rate cuts from the Federal Reserve, as the curve tends to drop relatively quickly after mid 2024.

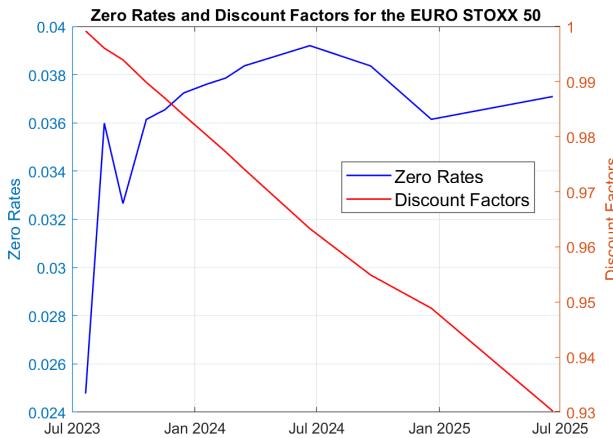


Figure 3: EURO STOXX 50

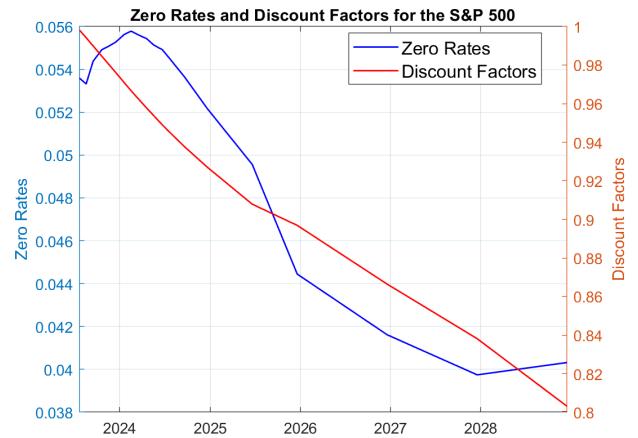


Figure 4: S&P 500

Empirically, we observe that the market-implied curve seems relatively close to the one provided by the ECB and the FED (see references for links). However, observing more closely, we can see that the market-implied interest

rate curve presents a spread over the "risk-free" rate, which in our case is around 50-100 basis points. The spread is slightly higher than the one found in the study by R. Baviera and M. Azzzone. This can be explained by looking at the 2023 interest rate market. Indeed, due to inflation, interest rates have increased in recent years, likely bringing some additional risk into the equity market. The risk we are referring to is the funding risk. Higher interest rates mean higher yields in short-term treasury bonds, so the equity market needs to outperform, increasing the funding risk and leading to a bigger spread.

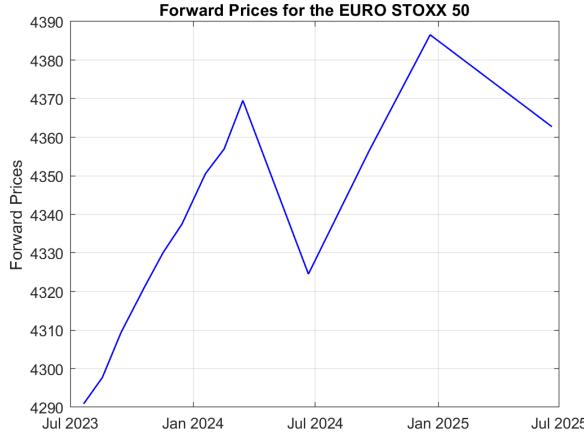


Figure 5: EURO STOXX 50

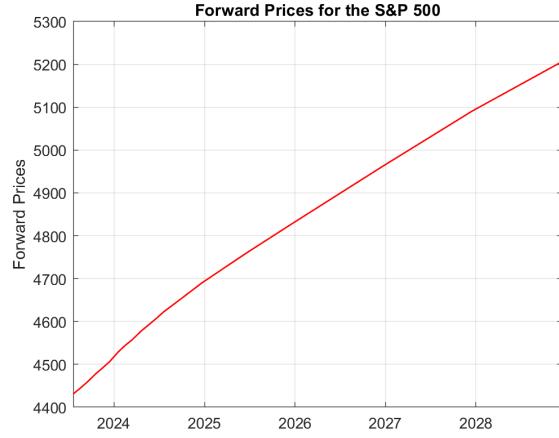


Figure 6: S&P 500

Regarding the forward prices, we observe that the US market has a monotone growing curve, while the European market presents some peaks and valleys. This behavior vanishes when considering the theoretical curve following the Garman-Kohlhagen model as presented below.

$$F(t_0, T) = \frac{S_0}{\bar{B}(t_0, T)}$$

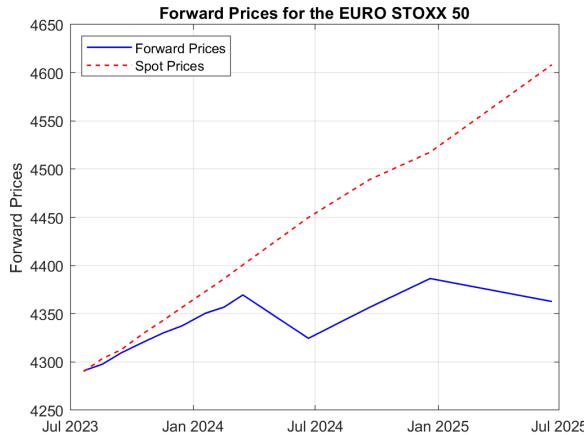


Figure 7: EURO STOXX 50

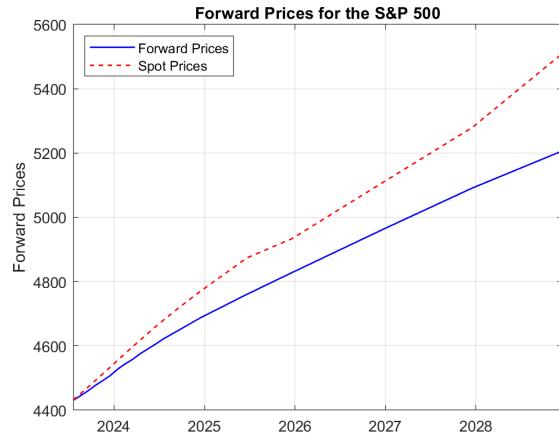


Figure 8: S&P 500

We can observe that the theoretical curve leveraging the market discount has a smoother shape, and the European market also presents a monotone behavior. Hence, we can compare liquid traded forwards simulated by the synthetic forward and the theoretical curve from the market discounts.

6.3. Data Processing: The Volatility Surface

The next step was to process the data framework in order to build a robust volatility surface and consequently calibrate the model on this new filtered dataset.

First of all, we computed the implied volatilities by inverting the Black formula using the Matlab function `blkimpv`, and we selected out-of-the-money (OTM) options. We chose to select only OTM options since the price of those particular products is mostly related to the stochasticity of the price and volatility, whereas in-the-money (ITM) options have an additional term due to the net "gain" at t_0 from the difference between the spot price and the strike.

As an example, consider a long position on two European call options and assume that all parameters, except the strike, are the same. One of the two calls has a strike lower than the current spot (ITM), while the other call has a strike above the spot (OTM). Suppose that the two strikes have the exact same distance from the at-the-money option (ATM), then the price of the ITM option will be higher due to the positive initial delta discounted at t_0 , that lead to inaccuracies when calibrating. Since we would like to capture the stochasticity of the market, we opted to select OTM options for the whole dataset, i.e., we select:

- Put options if $K < ATM$
- Call options if $K > ATM$

Then we plot the smiles for each maturity and check for any discontinuities in the curve.

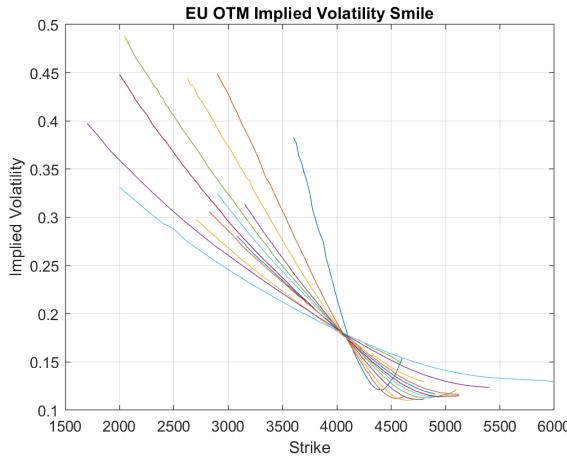


Figure 9: EURO STOXX 50

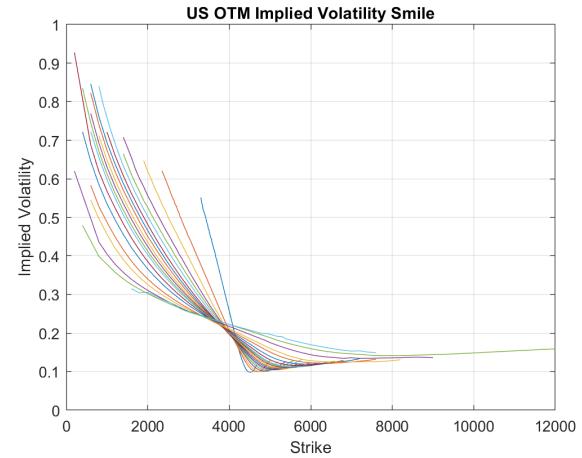


Figure 10: S&P 500

In the second step, we compute the Delta sensitivity for the OTM options and filter out all options that have a Delta smaller than 10% or bigger than 90%. Options with a Delta close to 0% (deep out-of-the-money puts or deep out-of-the-money calls) or close to 100% (deep in-the-money puts or deep in-the-money calls) have extreme sensitivity to the underlying asset's price movements. This can lead to large variations in option prices with small changes in the underlying asset price, making the calibration less stable and reliable.

In addition, these options correspond to lower liquidity levels and hence a wider bid-ask spread. Filtering out these options can improve the accuracy in reflecting the true market value. Summing up, by excluding options with Deltas less than 10% or greater than 90%, the calibration process focuses on options that are more representative of typical market conditions. This leads to a better fit of the model to the market data and more robust calibration results.

Finally, we plot the volatility smile for each maturity again. The resulting smiles are now narrower, and we check again for potential jumps or discontinuities.

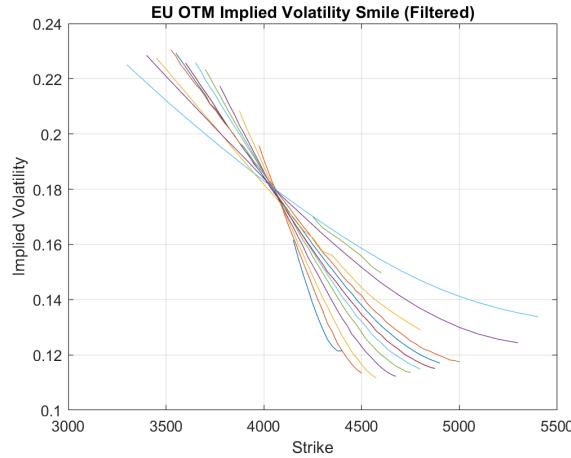


Figure 11: EURO STOXX 50

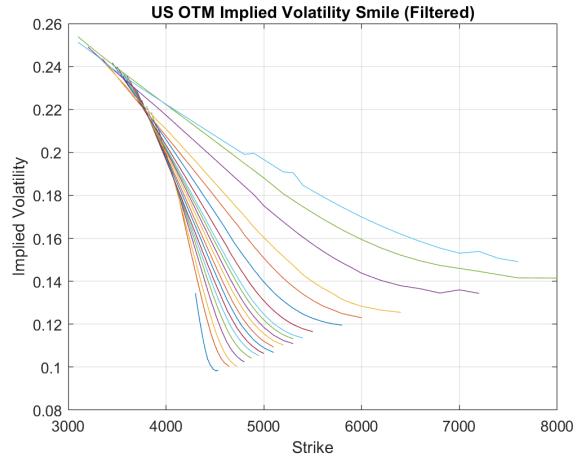


Figure 12: S&P 500

As we can notice, the American scenario presents a volatility smile that incorporates some jumps around the strike $K = 5000$, and also some smaller discrepancies in the tail of the smile for deep in-the-money calls. We decide to further investigate this maturity and apply several techniques to make it smoother, avoiding the loss of information from market data.

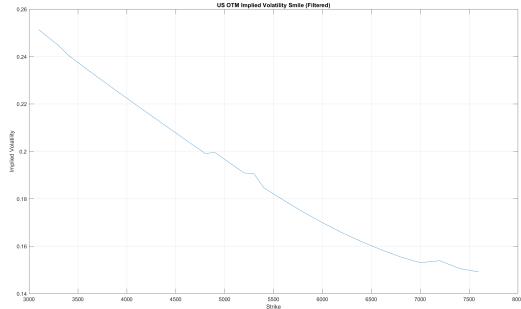


Figure 13: Original

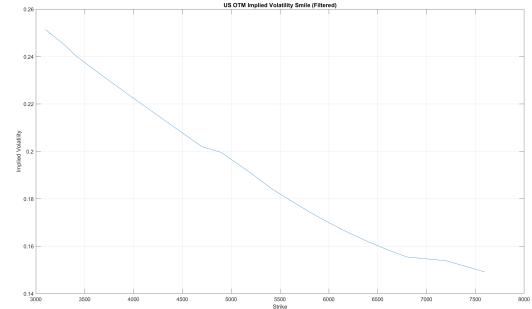


Figure 14: Smooth

In detail, we investigate the last maturity of the S&P 500 index. The presence of jumps in the volatility smile can lead to inaccuracies in the market calibration. Hence, we exploit different techniques to flatten the smile. Our first attempt was to use put-call parity to flatten the curve if the selection of OTM options performed above has an influence on the smile. However, this hypothesis led to no conclusion, so we checked that the relation between market prices is satisfied and took no further actions in this sense.

Subsequently, we decide to filter out from the dataset the strikes for which the implied volatility is higher than the previous one. Indeed, the presence of these discrepancies in the volatility smile can lead to arbitrage opportunities, violating the non-arbitrage condition. For instance, we can take a long position on the smaller strike with smaller volatility and a short position on the higher strike with higher volatility, guaranteeing a free lunch. For more details see M.Roper(2010). From this deduction, we investigate the call and put prices for the US Market at the last maturity, the one where the jumps in the implied volatility are more prominent. The following plots represent the bid, ask and mid prices for that maturity.

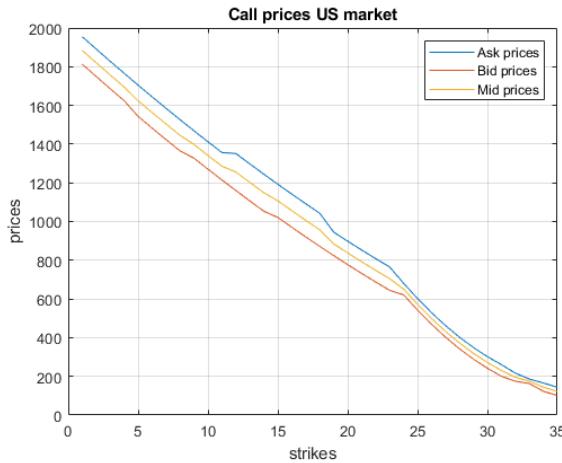


Figure 15: Call prices for the US market at last expiry

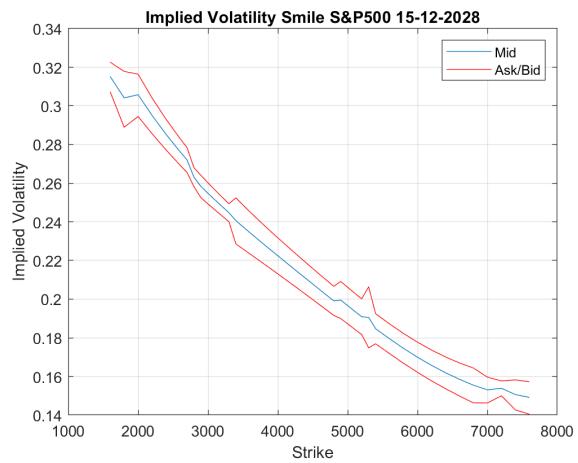


Figure 16: Implied volatilities for the US market at last expiry

As notable in the plot, the prices shown above do not seem to present arbitrage opportunities since options are not traded on the same curve when sold or bought. In this scenario, the bid-ask spread does not leave space for arbitrage opportunities. However, in our simplified model, the arbitrage opportunity may exist since we consider options traded at mid-prices.

For completeness, we report the plot of the implied volatilities obtained by inverting the call and put option prices for both the bid-ask and the mid prices. As we can see, the smile is far from smooth and presents several spikes. This is likely due to numerical approximation of the price curve, which was not completely smooth either. However, the bid-ask spread also reduces arbitrage opportunities. This shape of the curve is probably due to liquidity issues for such a long maturity. Indeed, the volumes present some spikes, meaning for two strikes across the entire curve, the volume is more than triple the others. Hence, in our model simplification, considering mid prices instead of the real traded ones can lead to some inaccuracies.

Furthermore, the volatility smile corresponding to the "2026/12/18" maturity has been slightly modified to flatten the tail of the curve.

To conclude this section, the updated implied volatility smiles and the respective volatility surfaces are provided:

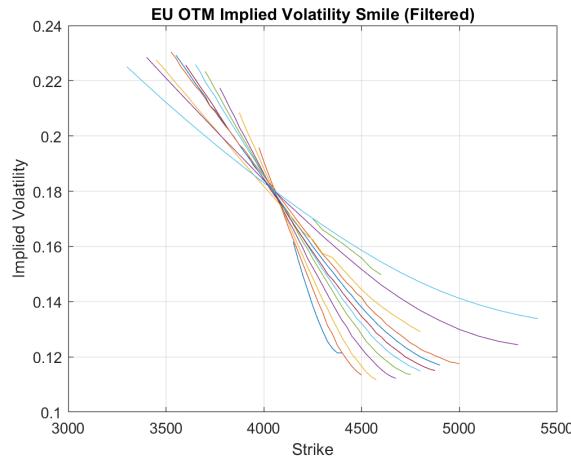


Figure 17: EURO STOXX 50

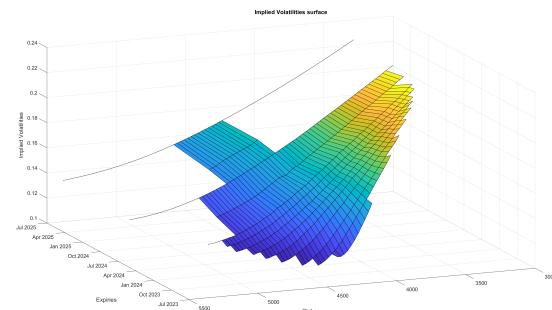


Figure 18: EURO STOXX 50 implied volatility surface

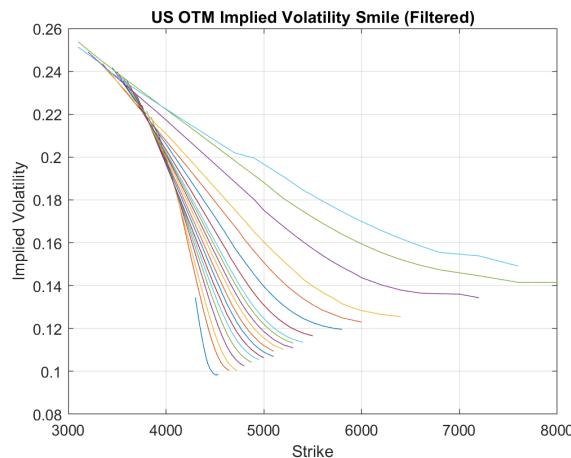


Figure 19: EURO STOXX 50

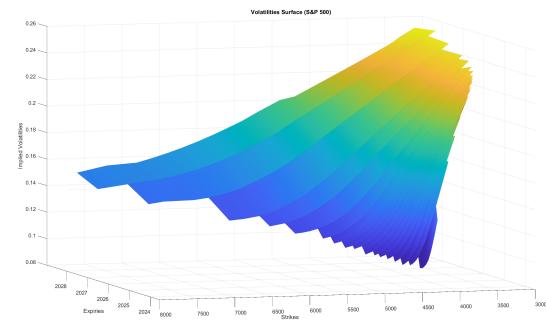


Figure 20: S&P 500

We can observe that the volatility surface of the American market is very steep in the first maturities, indicating that the prices are changing rapidly with respect to the given strikes. On the other hand, the European surface appears to be more relaxed, although for certain maturities, the range of strikes in the dataset is relatively small.

7. Historical Correlation

In this section, we focus on the historical returns contained in the dataset. The objective of the present discussion is to estimate the historical correlation between the two indexes using the yearly returns available in the struct "SPXSX5Ereturns.mat".

Before implementing any computations, we considered what to expect from the results. As long as S&P500 and EURO-STOXX 50 both represent developed markets and are influenced by similar economic factors, such as central bank policies, global economic conditions and geopolitical events we could expect high correlation. Although generally high, we suppose correlation varies over time. In particular, we assume that in case of global crises such as the pandemic the indexes may increase their correlation since they react to external global factors. In stable periods, the correlation may slightly decrease.

Indexes like the ones taken into consideration, generally show a positive trend in the long term as they are a representation of the economic growth and technological evolution. Correlation may significantly decrease in case of events that affect a specific area without massive global implications.

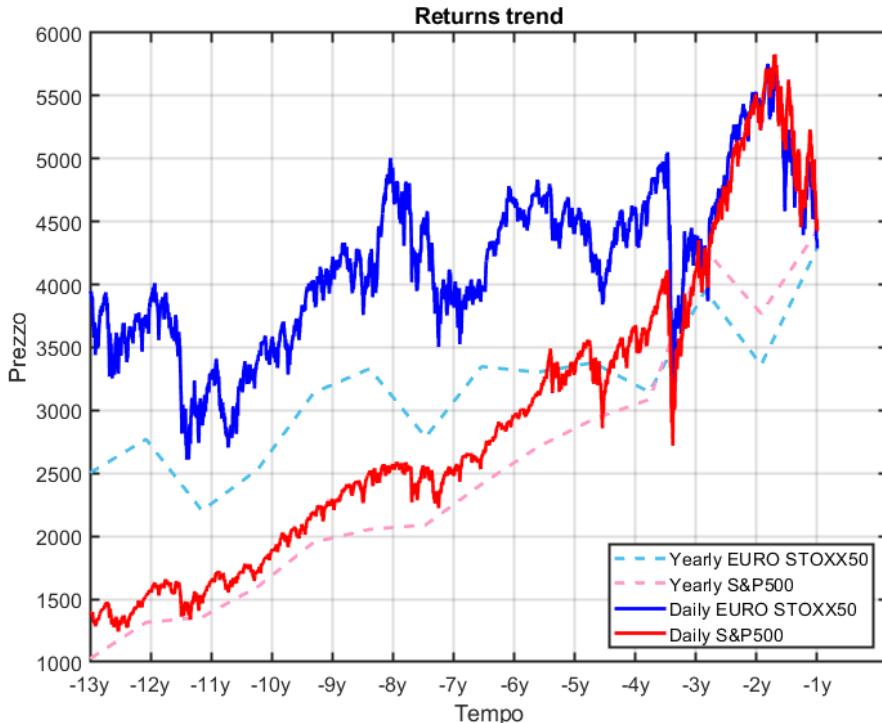


Figure 21: Historical returns

The plot above may help with some further comments. In fact, we observe that S&P 500 tends to be more performing with respect to EURO-STOXX50. Moreover, as yearly returns catch long term trends better than daily ones, we expect to observe a higher correlation for that kind of return. Computations confirm the previous proposition as correlation increases with the returns period. In the following table we present the results.

Returns	Correlation
Daily returns	0.6230
Monthly returns	0.7996
Yearly returns	0.8010

Table 3: Correlation values for different returns

8. Calibration of Multi-Variate Processes

In this section, we explore the calibration of a multivariate Lévy process model to market data. We consider a frictionless equity market where the log-returns are modeled via a multivariate Lévy process as defined above. Hence, under a risk-neutral martingale measure, the dynamics for each asset are given by:

$$S_j(t) = S_j(0)e^{(r-q_j-\varphi_{X_j}(-i))t+X_j(t)}, \quad j = 1, \dots, n$$

where $r > 0$ is the risk-free interest rate, $S_j(0)$ and q_j correspond, respectively, to the spot price and the dividend yield of the j -th asset, and $\varphi_{X_j}(-i)$ is the exponential compensator of the j -th component of the multivariate Lévy process, $X_j(t)$.

As highlighted in the **B&B** study, to perform a full calibration, both single-name and multi-name products are required to access full information on the correlation matrix. However, not all products are sufficiently liquid, hence we simplify our model by considering only liquid European options as described above.

The calibration process focuses on extracting the parameters of the marginal process $X(t)$, which is a subordinated Brownian Motion driven by a subordinator G , independent of the Brownian Motion. We explore the family of tempered stable subordinators, focusing on the Normal Inverse Gaussian (NIG), with an index of stability $\alpha = 1/2$, and the Variance Gamma (VG), with $\alpha = 0$. We select these subordinators due to their ability to model heavy-tailed distributions, which are commonly observed in financial asset returns.

The calibration will result in six parameters, three for each market. The NMVM parameters are σ , κ , and θ , which respectively describe the average volatility, the volatility of volatility, and the skewness. Specifically, σ approximate the market volatility, while κ highlights the variance rate, and θ describes the symmetry of the volatility smile (or surface).

To perform the calibration of the proposed multivariate asset model, we set up a minimization problem by minimizing the weighted root mean square error (RMSE). We build an objective function by considering the RMSE for both markets across all maturities and then weight them to normalize the impact of the two markets. The weights are computed considering the spot prices of the two indices:

$$w_{EU} = \frac{\text{Spot}_{EU}}{\text{Spot}_{EU} + \text{Spot}_{US}}, \quad w_{US} = \frac{\text{Spot}_{US}}{\text{Spot}_{EU} + \text{Spot}_{US}}$$

Even though these weights are close to 0.5 since the spot prices of the two indices are similar, they proved to be consistent and robust. Hence, we opted for them as they present a more interpretable objective function.

The RMSE is computed using the built-in function of MATLAB (MATLAB 2024a), which considers the difference between market and model option prices for all selected strikes and maturities. Hence, for each maturity, we compute the RMSE as:

$$\text{RMSE}_i = \sqrt{\frac{\sum_{j=1}^{N_i} (\text{price}_{\text{OTM},j}^{\text{mkt}} - \text{price}_{\text{OTM},j}^{\text{mod}})^2}{N_i}}$$

where N_i is the number of strikes for the i -th maturity. We then sum all the RMSE values for different maturities to find the total error for each index:

$$\text{total error}_k = \sum_{i=1}^{\#\text{maturities}} \text{RMSE}_{k,i}, \quad k = \text{US, EU}$$

Finally, we construct the objective function using the weights presented above:

$$\text{objective function} = w_{US} \cdot \text{total error}_{\text{US}} + w_{EU} \cdot \text{total error}_{\text{EU}}$$

In order to compute the RMSE we need a method to compute model prices, according with the NMVM framework we leverage the Lewis formula. In addition, we approximate the integral of the Lewis formula via fast Fourier transform (FFT). Let us recall the Lewis formula:

$$\frac{c(x)}{B(t_0, t)F_0} = 1 - e^{-x/2} \underbrace{\int_{-\infty}^{\infty} \frac{e^{-i\xi x}}{2\pi} \phi\left(-\xi - \frac{i}{2}\right) \frac{1}{\xi^2 + \frac{1}{4}} d\xi}_{I(x)}$$

where $\phi(\xi)$ is the characteristic function of our chosen model. In particular, we can write:

NIG Case

$$\text{Drift compensator (NIG)} = -\frac{t}{k} \left(1 - \sqrt{1 - 2k\theta - k\sigma^2} \right)$$

$$\phi(\xi) = \exp \left(\frac{t}{k} \left(1 - \sqrt{1 - 2i\xi k\theta + \xi^2 k\sigma^2} \right) \right) \cdot \exp(i\xi \cdot \text{drift compensator (NIG)})$$

VG Case

$$\text{Drift compensator (VG)} = \frac{t}{k} \log \left(1 - \theta k - \frac{1}{2} k\sigma^2 \right)$$

$$\phi(\xi) = \left(1 - i\xi\theta k + \frac{1}{2}\xi^2\sigma^2 k \right)^{-\frac{t}{k}} \cdot \exp(i\xi \cdot \text{drift compensator (VG)})$$

Furthermore, the integral can be written as a Fourier transform. This is true if we perform the following:

$$I(x) = \mathcal{F}[f(\xi)](x) = \int_{-\infty}^{\infty} f(\xi) e^{-i\xi x} d\xi \text{ if we choose } f(\xi) = \frac{1}{2\pi} \phi \left(-\xi - \frac{i}{2} \right) \frac{1}{\xi^2 + \frac{1}{4}}$$

We also need to choose parameters such that the FFT constraints are satisfied:

$$\begin{cases} z_1 &= -z_N \\ \xi_1 &= -\xi_N \\ dz &= \frac{z_N - z_1}{N} \\ d\xi &= \frac{\xi_N - \xi_1}{N} \\ d\xi \cdot dz &= \frac{2\pi}{N} \end{cases}$$

Hence, out of the original seven hyperparameters, we only need to specify two. As a rule of thumb, the first will always be the number of points to use in the grid to perform the discretization step of the algorithm (we chose $N = 2^M$). The other can be chosen between either ξ_1 or dz . In our case, we chose to model our FFT using M and dz , as modeling the FFT using the log-moneyness step lets us anchor our model to real financial quantities.

We set our hyperparameters to be $M = 15$ and $dz = 0.0025$. These two hyperparameters seem to be a good trade-off: M is large enough to let the algorithm work but at the same time does not affect the computational time too much, dz is enough to get out of local minima and not too large to overcome critical points.

We implement the minimization via the MATLAB function *fmincon*. In the minimization problem, we impose constraints to ensure the parameters' properties are respected:

$$\sigma_i \geq 0 \quad i = \text{US, EU}, \quad k_i \geq 0 \quad i = \text{US, EU}$$

We also impose the non-linear constraint derived in Section 2 from the convolution condition:

$$\frac{\sigma_{\text{US}}^2}{k_{\text{US}} \theta_{\text{US}}^2} = \frac{\sigma_{\text{EU}}^2}{k_{\text{EU}} \theta_{\text{EU}}^2}$$

However, *fmincon* presents some limitations. Indeed, it is very sensitive to the initial guess and it can happen that not all constraints are satisfied without reaching any valuable minimum. This is probably due to the complexity of the surface we are minimizing. We tried to solve these issues by considering different algorithms implemented in *fmincon*, such as Sequential Quadratic Programming (sqp), which is optimal for mid-size datasets and high precision, but this does not bring any consistent advantages.

Moreover, we explored different algorithms such as the Genetic Algorithm (*ga*), which aims to reach the global minimum of the objective function. Even though this method proved more robust and consistently produced reliable results, its drawback is the computational time. To reach the global minimum, the population size and the maximum generation must be relatively high, increasing the computational time exponentially.

Hence, we opted to use *fmincon*, aware of its limitations, and tried to avoid numerical mistakes while building a model that is robust with reasonable computational time.

In the following paragraphs, we explore different calibrations leveraging different implementations of the RMSE computation and considering both NIG and VG models. Moreover, we will focus our attention on the issues emerging in the last maturity of the US market.

Before going into details, we can introduce some expectations. We expect the calibration to be consistent and robust, which is one of the strengths of considering such a parsimonious model. Looking at the market volatility surface, we expect the S&P 500 parameters to be more volatile than the corresponding ones for the EURO STOXX 50. Furthermore, the last maturity in the American index, as discussed above, can generate some issues in the calibration procedure.

8.1. Unweighted RMSE

We start with the naive idea described above, without any additional weights on the RMSE computation. Considering the NIG model, we calibrated on all the available maturities and immediately encountered an issue with the last maturity of the S&P 500. Indeed, the minimization was far from robust and often reached a local minimum, unable to overcome it regardless of the initial condition. This is probably due to the unconventional shape of the volatility smile discussed in the previous section. However, by tuning the initial guess, we reached a minimum that is quite satisfactory.

Parameter	σ_{EU}	κ_{EU}	θ_{EU}	σ_{US}	κ_{US}	θ_{US}
Value	0.1219	0.7839	-0.1679	0.1640	4.9808	-0.0896

Table 4: Calibrated parameters for EU and US Markets (no weights)

However, the lack of consistency makes this model unfeasible. Hence, we opted to calibrate the parameters over all the maturities except the last one of the American index. This modification led to slightly different results, but it gained in robustness. Indeed, the calibration reached the same minimum for almost every initial guess, such that the parameters are sufficiently small. For instance, generating six random numbers from a uniform distribution (*rand* in MATLAB) multiplied by a coefficient of 0.2 leads to consistent results. The calibrated parameters are presented in the following table:

Parameter	σ_{EU}	κ_{EU}	θ_{EU}	σ_{US}	κ_{US}	θ_{US}
Value	0.1246	0.8267	-0.1620	0.1559	3.8448	-0.0940

Table 5: Calibrated parameters for EU and US Markets (no weights)

Analyzing the parameters, we notice that the average calibrated volatility of the S&P 500 is higher by around 300 basis points. This is in accordance with the volatility surface shown above, together with the other parameters. The American calibrated parameter presents a rather high value of κ , unexpected at first sight. However, considering the entire volatility surface, we can see that the volatility changes rapidly and behaves differently for the longer maturities. Hence, this value becomes acceptable under these conditions.

To evaluate the goodness of the calibrated parameters, we compute the model prices for both call and put options, leveraging the calibrated parameters, and compare them with the market prices. The metric we used is the relative percentage error, which yields the following results with the calibrated parameters:

Metric	Average Percentage Error	
	EU Market	US Market
Market Price	3.0155%	3.3041%
Implied Volatility	5.6615%	6.0369%

Table 6: Relative Percentage Error for EU and US Markets

As noted, the error in the prices is relatively small for both markets, with slightly better precision in the European index. Moreover, we inverted the prices using the Black formula to find the implied volatilities and compare

them with the market ones as well. Even though the error is slightly larger, it is still a good fit for the market structure. To better understand the behavior of the model, we present a plot comparing the volatility surface and an example of the prices for a call option for a given maturity.

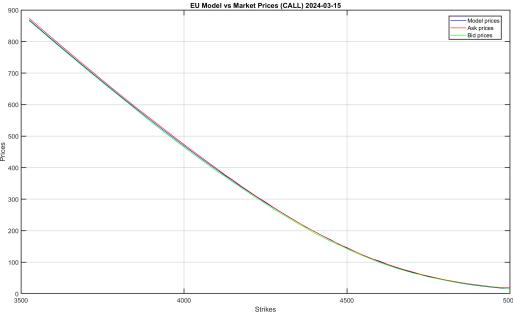


Figure 22: Call prices for the EU market

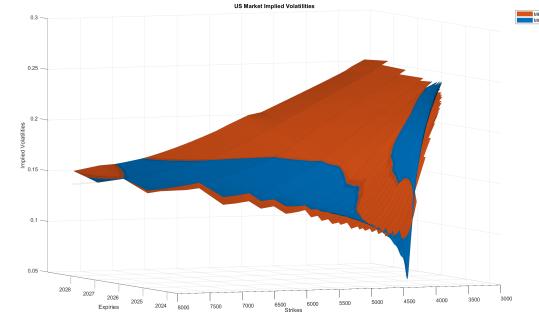


Figure 23: Implied volatility surface for the US market

As we can see, the model volatility surface captures the market data quite well, although it presents some inaccuracies for the shorter maturities. In contrast, the prices are well fitted for both call and put options and for all the maturities. It is worth noting that even though the parameters are calibrated excluding the last maturity of the American data, the model is able to capture the prices and volatility smile for this maturity as well.

8.2. Time Weighted RMSE

At this stage, we've implemented a couple of methods to enhance calibration performance, focusing on certain specific maturities by applying time-based weights. The idea behind introducing time weights stemmed from the need to alleviate the influence of longer maturities, to mitigate their impact by assigning lesser importance to distant maturities. These time weights complement the original weights derived from spot prices of the indexes, influencing the computation of total errors for both markets. Consequently, total errors are now calculated as a weighted sum across all available maturities, integrating both spot price-based and time-based considerations.

$$\text{total error}_k = \sum_{i=1}^{\#\text{maturities}} \tilde{w}_i \cdot \text{RMSE}_{k,i}, \quad k = \text{US, EU}$$

where \tilde{w}_i is the time weight for the i -th maturity. Despite accounting for all maturities, time weights may introduce inaccuracies. It's possible that they may overemphasize certain maturities, potentially overlooking valuable information from other expiries due to the weight construction.

We've implemented two time-weighting methods based on different lines of thought.

8.2.1 'Linear' weights

The first approach prioritizes the most immediate expiries, considering them to be the most precise, while assigning lesser importance to the more distant ones.

Here is the method we employed to derive the 'linear' weights:

$$\tilde{w}_{N-i} = \frac{ttm_i}{\sum_{j=1}^N ttm_j}$$

We normalized the weights in order to ensure they sum up to one.

Parameter	σ_{EU}	κ_{EU}	θ_{EU}	σ_{US}	κ_{US}	θ_{US}
Value	0.1288	0.5132	-0.1787	0.1443	1.9073	-0.1039

Table 7: Parameter Values for EU and US Markets (linear time weights)

The observation that the average calibrated volatility for the US market exceeds that of the EURO STOXX aligns with the surfaces depicted above.

Metric	Average Percentage Error	
	EU Market	US Market
Market Price	2.7914%	3.3516%
Implied Volatility	5.3822%	6.3330%

Table 8: Relative Percentage Error for EU and US Markets ('linear' time weights)

We maintain consistency by utilizing the same evaluation metrics throughout our discussion. While the calibrated parameters yield satisfactory results, with model prices and implied volatilities closely matching market values, consistency remains a concern due to parameter variations depending on the initial conditions.

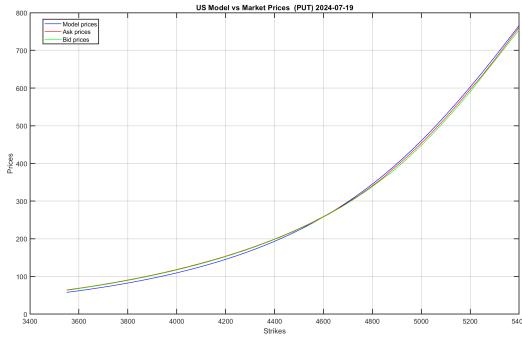


Figure 24: Put prices for the US market

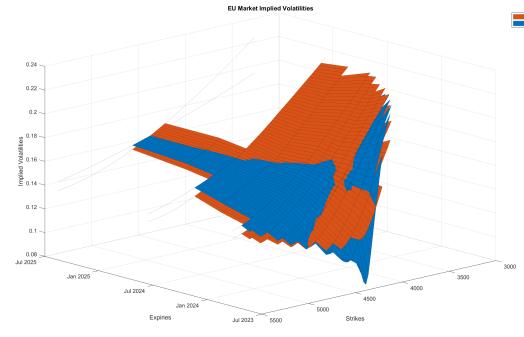


Figure 25: Implied volatility surface for the EU market

² The implied volatility surface overlaps pretty well the market curve in the central maturities but present some inaccuracies especially for late maturities.

8.2.2 'Triangular' weights

Given that the ultimate objective of the project is to price a one-year maturity derivative, we propose employing a different type of time-weighting strategy based on the proximity to the one-year maturity. As linear weights, triangular weights also tend to underestimate distant maturities. Consequently, we opt to calibrate on all available maturities.

Here's the algorithm we utilized to calculate the weights:

$$\bar{w}_i = \frac{e^{-0.01|ttm_i - mat_{1y}|}}{\sum_{j=1}^N e^{-0.01|ttm_j - mat_{1y}|}}$$

where \bar{w} are the 'triangular' time weights and the decay rate is 0.01.

Parameter	σ_{EU}	κ_{EU}	θ_{EU}	σ_{US}	κ_{US}	θ_{US}
Value	0.1312	1.2349	-0.1406	0.1387	2.3005	-0.1089

Table 9: Parameter Values for EU and US Markets ('triangular' time weights)

²The surface for the EU market shrinks when it gets closer to the last maturities due to few strikes available at that expiry. Hence, we built a mesh that takes into account all strikes and maturities available. Where the volatility surface data are not provided we simply plot the smile.

Again we could expect a higher volatility for the S&P 500 rather than the EURO-STOXX 50. The skew is negative, it means that OTM put options have a higher implied volatility than OTM call options and that could be expected looking at the unfiltered smiles. This is often seen in equity markets where investors are more concerned about price drops and hence are willing to pay more for put options to protect their investments.

Metric	Average Percentage Error	
	EU Market	US Market
Market Price	3.4482%	3.2065%
Implied Volatility	6.3026%	5.9599%

Table 10: Relative Percentage Error for EU and US Markets ('triangular' time weights)

With this kind of weights we are overestimate the 1 year expiries. As a result, we get precise prices and implied volatilities for those maturities close to 1 year, but we find values

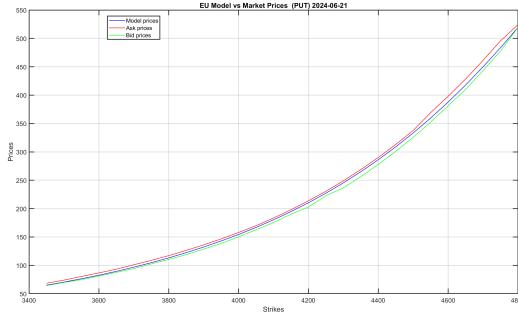


Figure 26: Put prices for the US market

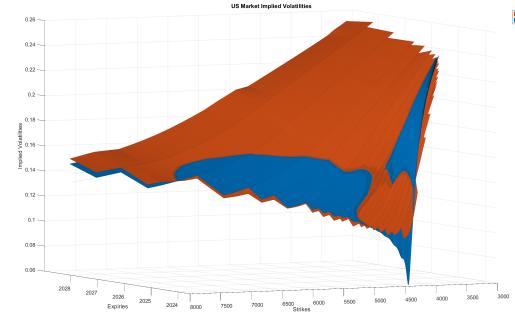


Figure 27: Implied volatility surface for the EU market

As we can notice from the relative percentage errors, this model outperforms US market with respect to the previous ones. It is able to capture in a more accurate way the dynamics of the American index, i.e. it is notable in the volatility surface how the model is able to replicate the smiles also for longer maturities.

Also prices are more accurate for the S&P 500 index rather than the EURO-STOXX 50. Overall, this methodology performs rather well, but at the same time adds a level of complexity to the model.

8.3. Volume Weighted RMSE

Since the dataset also includes volume data, it's possible to utilize these values to construct weights for the RMSE computation. In this scenario, unlike before, the weights are directly applied to the RMSE computation. Specifically, the concept involves using the volumes provided for each maturity and computing a weighted RMSE for each maturity, which is then summed up to obtain the total RMSE. To accomplish this task, we make use of the *rmse* MATLAB function. The objective is to emphasize maturities and strikes with high volumes, which is a metric closely related to liquidity. Consequently, we aim to calibrate our model based on the most liquid options among those already selected.

Calibrating the NIG model with consideration of all maturities yields the following results:

Parameter	σ_{EU}	κ_{EU}	θ_{EU}	σ_{US}	κ_{US}	θ_{US}
Value	0.1191	0.7452	-0.1740	0.1639	5.3924	-0.0890

Table 11: Parameter Values for EU and US Markets (volume weights)

The calibrated parameters are quite similar to those computed with no additional weights, and there doesn't seem to be any substantial improvement in the parameters. Even when neglecting the last American maturity, the calibrated parameters do not show significant improvements; rather, the results become even closer to the

ones obtained without any additional weights.

Once again, we compute the relative percentage error on prices and implied volatilities, obtaining consistent results with those obtained previously but without any further enhancements.

These considerations lead us to refrain from further investigating this idea, as it doesn't yield any improvements and, at the same time, makes the model less parsimonious and more complex.

8.4. Filtered Root Mean Square Error

After exploring different weights for the computation of RMSE, we shifted our focus to a new approach when computing the RMSE. Specifically, we follow the methodology presented by **B&B**, where the RMSE is computed not only by weighting it but also by considering the pricing error only outside the bid-ask spread. Specifically, we consider a positive error (with respect to the mid-price) only if the model price falls outside the bid-ask spread. This choice aims to measure how well the model fits within the trading cost bounds. Additionally, we apply the linear weights on time introduced above to better capture the dynamics of the shorter maturities and penalize the longer maturities that, as we have seen, lead to some inaccuracies. This allows us to include all maturities for both markets in the calibration.

The model performs rather well, reaching a different minimum from the preceding cases, but the parameters are consistent within the provided dataset. The calibration yields the following results:

Parameter	σ_{EU}	κ_{EU}	θ_{EU}	σ_{US}	κ_{US}	θ_{US}
Value	0.1323	0.6420	-0.1623	0.1467	2.0456	-0.1008

Table 12: Parameter Values for EU and US Markets (filtered RMSE)

Once again, the parameters tell us about the behavior of the two markets. Indeed, as we can see, the average volatility of the American index is consistently greater than that of the European one. Moreover, the vol-of-vol of the S&P 500 index is still relatively high, highlighting a less stable volatility surface.

This model, as shown in the following table, outperforms the preceding methodologies. In particular, for the European market, it shows a consistent improvement in both prices and implied volatilities, while the American side is still similar to the previous models. This is probably due to the linear weights, which penalize the long maturities present in the S&P 500 dataset.

Metric	Average Percentage Error	
	EU Market	US Market
Market Price	2.7382%	3.3190%
Implied Volatility	5.1472%	6.2538%

Table 13: Relative Percentage Error for EU and US Markets (filtered RMSE)

The following plots show the goodness of fit of this model. First, we consider, for the same maturity, both a call and a put on two different markets:

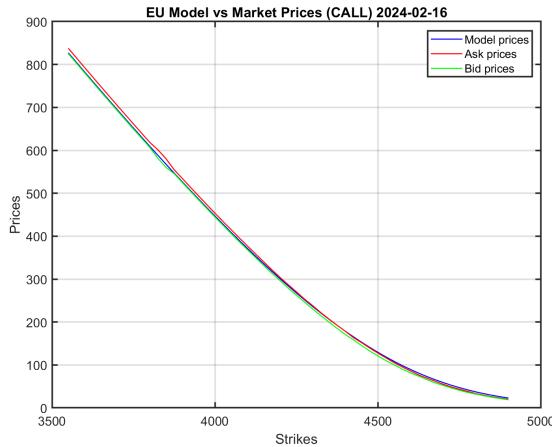


Figure 28: Call prices for the EU market (filtered RMSE)

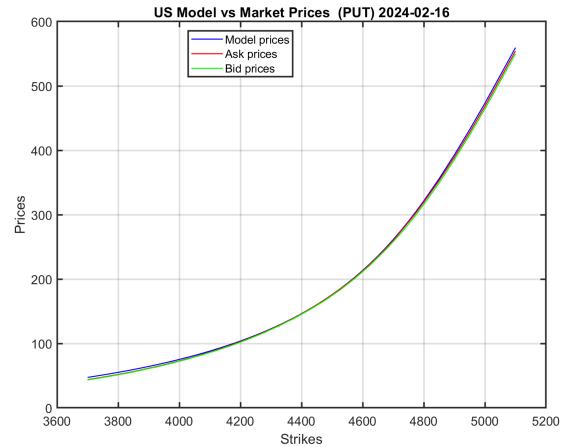


Figure 29: Put prices for the US market (filtered RMSE)

As already highlighted by the relative percentage error, this model captures the price market structure of the two indexes precisely.

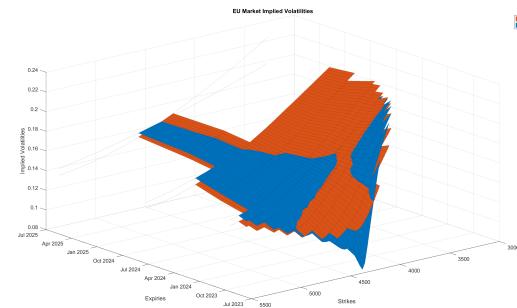


Figure 30: Implied volatility surface for the EU market (filtered RMSE)

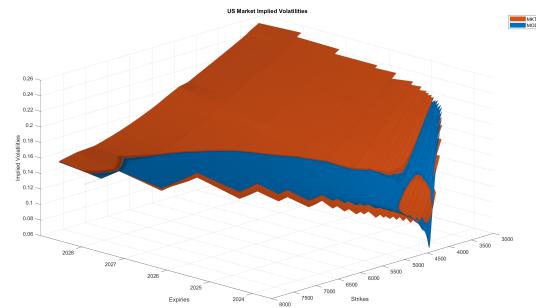


Figure 31: Implied volatility surface for the US market (filtered RMSE)

Regarding the implied volatility surface, in the case of the European index, the model matches the market data accurately, apart from the first maturity where a slight mismatch is present. The American surface is also followed rather well by the model.

From a computational perspective, this model seems to be one of the fastest to converge. However, it presents some downsides; indeed, it is less robust and stable than the model with no additional weights, and it is quite sensitive to the initial conditions of the minimization. The situation improves if we set a larger step (dz) as a hyper-parameter of the FFT. Empirical studies show that the model behaves more consistently when setting $dz = 0.0050$.

8.5. Variance Gamma

For the sake of completeness, we implemented the Variance Gamma subordinator and performed the same calibration as presented above. First, we consider the simpler RMSE computation that yields the following results:

	σ_{EU}	κ_{EU}	θ_{EU}	σ_{US}	κ_{US}	θ_{US}
Not considering the last US maturity						
Unweighted	0.1395	0.8251	-0.1305	0.1598	2.5135	-0.0856
Linear	0.1391	0.4594	-0.1520	0.1466	1.2181	-0.0984
Volumes	0.1369	0.8021	-0.1350	0.1582	2.6319	-0.0861
Triangular	0.1472	1.2294	-0.1115	0.1513	1.9444	-0.0911
Considering the last US maturity						
Unweighted	0.1382	0.8122	-0.1328	0.1654	3.0102	-0.0826
Linear	0.1391	0.4594	-0.1520	0.1469	1.2320	-0.0980
Volumes	0.1340	0.7716	-0.1405	0.1653	3.4244	-0.0822
Triangular	0.1472	1.2294	-0.1115	0.1513	1.9444	-0.0911

Table 14: Calibrated parameters with VG model

As shown by the calibrated parameters, the model appears to be consistent, reaching the same minima even when considering the last American maturity. Once again, the parameters obtained provide insights into the two markets, highlighting the differences between the S&P 500 and the EURO STOXX 50 indexes described in the previous sections.

Numerically, the VG method outperforms its NIG counterpart, demonstrating greater robustness and resilience to changes in hyper-parameters and initial conditions. Further investigation can be conducted on two of the methodologies that have obtained optimal results under the NIG framework.

Unweighted VG

In the first scenario, we compare VG and NIG in the case of no additional weights and neglecting the last maturity for the American index. In this case, both methods are rather robust and are not particularly influenced by the initial condition. Both methods converge relatively fast, yielding the following results:

Parameter	σ_{EU}	κ_{EU}	θ_{EU}	σ_{US}	κ_{US}	θ_{US}
NIG	0.1246	0.8267	-0.1620	0.1559	3.8448	-0.0940
VG	0.1395	0.8251	-0.1305	0.1598	2.5135	-0.0856

Table 15: Calibrated parameters for EU and US Markets (no weights)

Later, we compare the performance via metrics on the prices and the implied volatilities, and the NIG shows a net better result. Indeed, the VG presents a relative percentage error that is consistently higher than the NIG, reaching a difference in the error of more than 2% in the implied volatilities for the S&P 500 index. However, when plotting the implied volatilities surface, we can notice that the Variance Gamma matches the market data even better than the Normal Inverse Gaussian, but it presents a notable discrepancy for the shorter maturities.

Hence, we can conclude that the VG can be a better choice when dealing with derivative pricing on longer maturities, while the NIG should be preferred for the shorter ones.

Filtered RMSE VG

In the second scenario, we explore the VG method with the filtered RMSE technique described above and compare it with the NIG model. The model converges rapidly to the following calibrated parameters:

Parameter	σ_{EU}	κ_{EU}	θ_{EU}	σ_{US}	κ_{US}	θ_{US}
NIG	0.1323	0.6420	-0.1623	0.1467	2.0456	-0.1008
VG	0.1398	0.4961	-0.1483	0.1479	1.3100	-0.0966

Table 16: Calibrated parameters for EU and US Markets (filtered RMSE)

Again, the parameters describe the market data as expected, although the two models yield non-perfect matching parameters. Also, in this case, the error of the VG model is relatively higher. Comparing the two VG methodologies, this second one approximates the market structure with less precision. This is also notable by looking at the plots of the implied volatility surface, which show a more inconsistent behavior of the model.

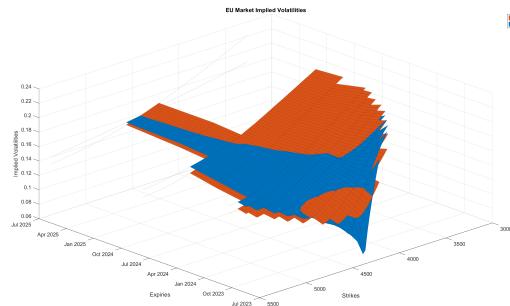


Figure 32: Implied volatilities surface for the EU market with VG model and filtered RMSE

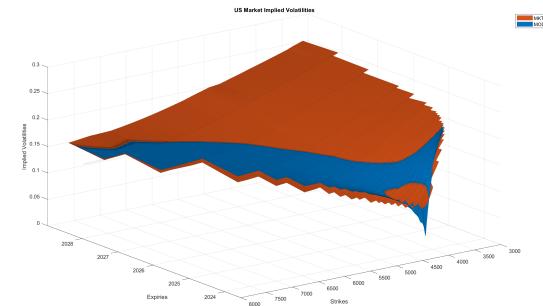


Figure 33: Implied volatilities surface for the US market with VG model and filtered RMSE

8.5.1 A comparison between two subordinators: NIG vs VG

Both the Normal Inverse Gaussian (NIG) and Variance Gamma (VG) distributions offer significant advantages for financial modeling, particularly in capturing the heavy tails and skewness observed in asset returns. These are the main reasons why NIG and VG are widely used in pricing financial derivatives and have been investigated in various studies in terms of their fit to empirical stock returns.

They both present a closed analytical form of the characteristic function that facilitates pricing complex derivatives.

The disadvantages of using NIG and VG models are the computational costs associated with using their distributions and the need for frequent re-calibration to keep the model updated, especially in high-frequency trading environments.

In conclusion, while both the Normal Inverse Gaussian (NIG) and Variance Gamma (VG) distributions offer valuable tools for financial modeling, their suitability depends on the characteristics of the market and how well they can capture its dynamics. The choice between NIG and VG should be made based on factors such as the distribution of returns, the specific needs of the financial application, and the computational resources available. Ultimately, neither model is universally superior, and the decision should be based on a careful evaluation of their strengths and weaknesses in the context of the given market environment.

8.6. Alternative Calibrations

As a final attempt, we test our model on different data combinations. Specifically, we calibrate the parameters on a single maturity and on a bucket of maturities.

We expect the model to perform rather well, although calibrating on a restricted number of maturities may lead to inaccuracies when considering the entire dataset.

8.6.1 Single Maturity Calibration

The model is tested on a single maturity. Specifically, we consider the closest date to the expiry (2024-06-21) of the derivative that will be priced in the following sections.

The minimization is performed using NIG and VG, yielding the following results:

Parameter	σ_{EU}	κ_{EU}	θ_{EU}	σ_{US}	κ_{US}	θ_{US}
NIG	0.1274	1.1805	-0.1463	0.1339	2.1308	-0.1145
VG	0.1499	1.3419	-0.1075	0.1521	1.9681	-0.0901

Table 17: Calibrated parameters for EU and US Markets calibrating only on 2024-06-21

Calibrating on a single maturity gives similar parameters to the full calibration. This methodology is obviously much cheaper from a computational perspective, but it lacks in capturing the entire structure. Indeed, when computing the metrics on prices and implied volatilities, we can observe that the model is truly consistent and precise around the selected maturity but deviates by some basis points as the distance from the calibration date increases. This can be seen in the following graphs:

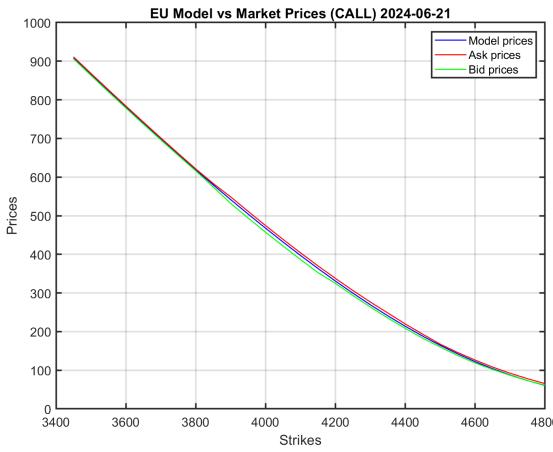


Figure 34: Call prices EU, NIG model calibrated only on 2024-06-21

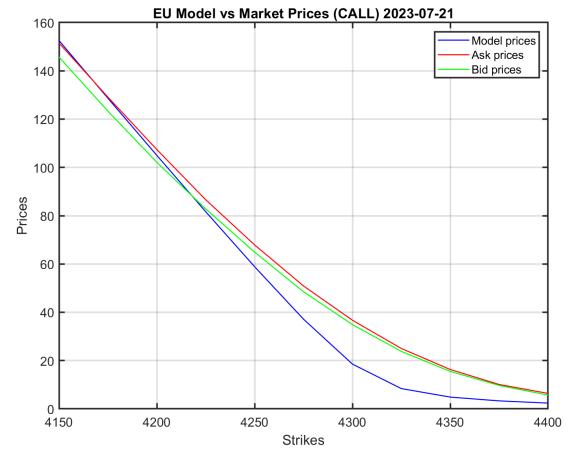


Figure 35: Call prices EU, NIG model calibrated only on 2024-06-21

As we can see, the model fits the call prices for the European index perfectly at the maturity selected for calibration, while considering the first maturity of the same market, we can elucidate a mispricing behavior. Similarly, this happens when checking the implied volatility surfaces for the two markets. However, the overall result is not too far from the ones obtained via full calibration.

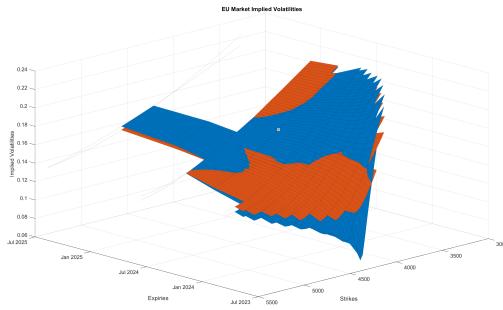


Figure 36: Implied volatilities surface for the EU market with NIG model considering only 2024-06-21

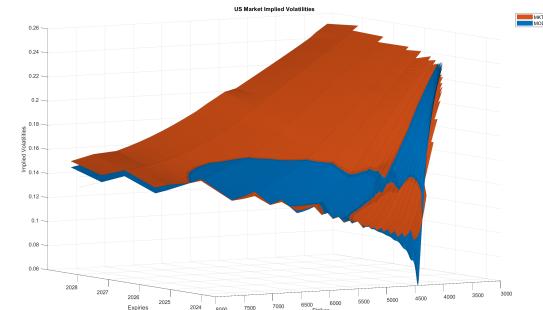


Figure 37: Implied volatilities surface for the US market with NIG model considering only 2024-06-21

8.6.2 Maturity Window Calibration

Later, we extend the idea presented above by taking into consideration a bucket of data around 2024-06-21 (same data of the previous point). We test the results calibrating on 5 dates via NIG and VG models. The calibration yields the following results:

Parameter	σ_{EU}	κ_{EU}	θ_{EU}	σ_{US}	κ_{US}	θ_{US}
NIG	0.1275	1.1374	-0.1462	0.1363	2.2288	-0.1116
VG	0.1450	1.1846	-0.1129	0.1512	1.9791	-0.0911

Table 18: Calibrated parameters for EU and US Markets calibrating on window maturity

This approach can be a satisfactory trade off, indeed this model is able to capture on average rather well the whole framework. Both the NIG and the VG obtain similar performance on our metrics.

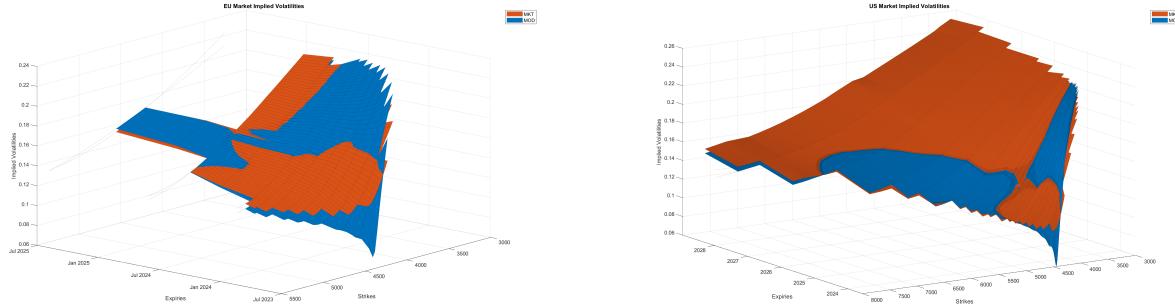


Figure 38: Implied volatilities surface for the EU market with NIG model considering window maturity

Figure 39: Implied volatilities surface for the US market with NIG model considering a window of maturity

8.7. Final Consideration on the Calibration

In summary, the calibration provides coherent insights into the two markets. Indeed, the parameters obtained in the methodologies discussed above yield similar results. It is noteworthy that the EURO STOXX 50 seems to be less volatile and more consistent over the maturities. This is in accordance, from a financial interpretation perspective, with the rates term structure described in the previous section. In general, the European equity market seems to be more stable, while the S&P 500 shows a more variable behaviour, probably due to the larger maturity dataset.

Regarding the numerical procedure, we encountered several issues related to the robustness of the minimization. The presence of many local minima led the built-in MATLAB function *fmincon* to not always converge to the global minimum. Indeed, even when exploring different algorithms, the model was still quite sensitive to the initial guess and other hyper-parameters. As a solution, we proposed several approaches, and some of them outperformed the trivial model.

For instance, simply neglecting the last maturity of the American market (due to the problems discussed previously), we obtained a more robust model that is able to nicely capture the whole structure.

Alternatively, the model that outperformed the others and takes into account the last US maturity as well is the *Filtered RMSE*. Even though it adds some level of complexity to the model and imposes additional weights in the RMSE computation.

As a last remark, we highlight the performance of the window cut on maturities, which overall was able to capture the whole market and at the same time match with fine precision the set of strikes and maturities around the expiry date of the derivative we are asked to price.

Moreover, the model selection between the NIG or the VG is subordinated to the market we are dealing with. Indeed, the choice of the right subordinator can lead to a more accurate estimation of market data. As an example, we recall the study of A.Göncü and H.Yang (2016), where a refined research on the goodness of fit of the subordinator introduced above led to the conclusion that: *"For high-frequency Chinese index returns, the NIG model is more robust and provides a better fit to the empirical distributions of returns at different time scales."*

9. Calibration of the Idiosyncratic and Systematic Factors

The second step of calibration involves tuning the parameters of the idiosyncratic process $Y(t)$ and the systematic component $Z(t)$ to match the correlation between the two markets and leverage the convolution condition.

A multidimensional modeling approach allows for any univariate Lévy processes for $Y(t)$ and $Z(t)$, with the resulting distribution of the margin $X(t)$ attainable through the characteristic function. Alternatively, it is possible to ensure that $X(t) = Y(t) + aZ(t)$ follows a specified distribution by imposing convolution conditions on the processes $Y(t)$ and $Z(t)$. For computational simplicity, we select $X(t)$, $Y(t)$, and $Z(t)$ from the same family of Lévy processes, either NIG or VG.

As a consequence, we need to manage the following parameters:

- $a_j \in \mathbb{R}$, $j = \text{US}, \text{EU}$ represents the systematic part coefficient. The sign of this parameter leads to a corresponding positive or negative correlation between the marginal $X(t)$ and the systematic part $Z(t)$.
- $\beta_j \in \mathbb{R}$, $j = Z, \text{US}, \text{EU}$ is the skewness of the distribution. This parameter measures the asymmetry of the process.
- $\gamma_j \geq 0$, $j = Z, \text{US}, \text{EU}$ represents the average volatility.
- $\nu_j \geq 0$, $j = Z, \text{US}, \text{EU}$ indicates the volatility rate (volatility of the volatility) of the j -th process.

The implementation of the algorithm for the computation of ν for the systematic and the two idiosyncratic factors leverages the MATLAB built-in function *fmincon*. We explored different algorithms, such as those presented in the previous section (e.g., genetic algorithm, interior-point, etc.). In this case, we opted for the sequential quadratic programming algorithm since it captures the minimum of the problem more accurately.

The main idea of the minimization problem is based on the theoretical derivation in the third section. We set the objective function as the squared difference between the historical correlation and the linear correlation formula involving the three parameters.

$$\text{Objective function} := \left(\sqrt{\frac{\nu_{\text{US}} \cdot \nu_{\text{EU}}}{(\nu_{\text{US}} + \nu_Z)(\nu_{\text{EU}} + \nu_Z)}} - \rho_{\text{mkt}} \right)^2$$

The historical computation, considering yearly log-returns for the two markets, yields $\rho = 0.8010$. Moreover, we set some linear and non-linear constraints. In particular, we require that the systematic parameter ν_Z is larger than the maximum of the calibrated κ values of the two markets, ensuring that all the variance rate parameters for both idiosyncratic and systematic factors are positive. The non-linear constraints ensure that the convolution conditions are respected and that the two relations for computing the linear correlation lead to the same value.

$$\begin{aligned} \sqrt{\frac{\nu_{\text{US}} \cdot \nu_{\text{EU}}}{(\nu_{\text{US}} + \nu_Z)(\nu_{\text{EU}} + \nu_Z)}} - \frac{\sqrt{\kappa_{\text{EU}} \cdot \kappa_{\text{US}}}}{\nu_Z} &= 0 \\ \frac{\nu_{\text{US}} \cdot \nu_Z}{\nu_{\text{US}} + \nu_Z} - \kappa_{\text{US}} &= 0 \\ \frac{\nu_{\text{EU}} \cdot \nu_Z}{\nu_{\text{EU}} + \nu_Z} - \kappa_{\text{EU}} &= 0 \end{aligned}$$

As a final step, we tuned the initial guess, as this model lacks some robustness. Even with more robust algorithms, the minimization did not show significant improvements. Hence, we opted to implement a calibration that satisfies the introduced constraints while refining the options of the MATLAB function to better explore the surface we are minimizing.

The calibrated parameters with the initial guess $X_0 = 0.5 \cdot \text{ones}(1, 3)$ are reported in the following table:

Subordinator		ν_{US}	ν_{EU}	ν_Z
NIG	Unweighted	8.0652	0.9315	7.3476
	Filtered RMSE	4.8679	0.7848	3.5282
	Time Window	5.4296	1.6268	3.7806
VG	Unweighted	5.7632	1.0125	4.4576
	Filtered RMSE	3.6190	0.6542	2.0532
	Time Window	4.9744	1.8522	3.2867

Table 19: Calibration of ν_{US} , ν_{EU} , and ν_Z for different scenarios

Even though the different approaches lead to different values, we can notice that all scenarios follow the same pattern. Namely, the ν related to the American idiosyncratic factor is consistently higher, aligning with the other parameters already implemented, showing a higher variance rate for the S&P 500 index. The parameters seem rather large, but considering that we are calibrating the two markets together for different maturities, we did not expect low levels as in Ballotta and Bonfiglioli's study.

Given the ν parameters, we can discuss how our model captures the correlation between the two indices. We report the results in the following table:

	Unweighted	Filtered RMSE	Time Window
NIG	0.2426	0.3248	0.4211
VG	0.3231	0.3927	0.4659

Table 20: Calibrated correlation for NIG and VG models

As we can see, neither of these models reaches the historical correlation ($\rho = 0.8010$). The "time window" model is the one that best captures this relationship, achieving more than 50%. Moreover, we can notice that the Variance Gamma (VG) model consistently provides a closer fit compared to the NIG model. The level of correlation expressed is rather low, which might be a limitation of our calibration.

After calibrating the ν parameters, we calibrated the other parameters by leveraging the convolution condition. Specifically, we set up a system of equations and solved it using the built-in MATLAB function *solve*:

$$\begin{cases} a_{US} \cdot \beta_Z - \left(\frac{\kappa_{US} \cdot \theta_{US}}{\nu_Z} \right) = 0 \\ a_{EU} \cdot \beta_Z - \left(\frac{\kappa_{EU} \cdot \theta_{EU}}{\nu_Z} \right) = 0 \\ \kappa_{US} \cdot \sigma_{US}^2 - \nu_Z \cdot a_{US}^2 \cdot \gamma_Z^2 = 0 \\ \kappa_{EU} \cdot \sigma_{EU}^2 - \nu_Z \cdot a_{EU}^2 \cdot \gamma_Z^2 = 0 \end{cases}$$

For the sake of simplicity, we report only the results for the NIG model (see Appendix C for the VG results).

	Parameter	a	gamma	nu	Beta
Unweighted	Y_EU	0.7835	0.1174	0.9315	-0.1438
	Y_US	2.1143	0.1077	8.0652	-0.0448
	Z	-	0.0534	7.3476	-0.0233
Filtered RMSE	Y_EU	0.4267	0.1197	0.7848	-0.1327
	Y_US	0.8445	0.0951	4.8679	-0.0424
	Z	-	0.1323	3.5282	-0.0692
Time Window	Y_EU	0.4368	0.1066	1.6268	-0.1022
	Y_US	0.6535	0.0873	5.4296	-0.0458
	Z	-	0.1601	3.7806	-0.1007

Table 21: Calibrated Parameters for Idiosyncratic and Systematic Factors under Different Scenarios, NIG subordinator

As shown in the table, across the different methodologies, signs of the parameters are consistent, and the values seem acceptable with respect to their financial meaning. Indeed, the coefficient a is positive, as expected, since both indexes are positively impacted by the systematic factor. In fact, the two indexes tend to move together.

Moreover, the negative β is consistent with what we found in the previous calibration, since θ was also negative, indicating a smile that is slightly asymmetric in favor of OTM puts.

In conclusion, these three methodologies seem to perform similarly, with a slight preference for the "Time Window" model. However, this method neglects part of the market data and is tuned to better capture just the maturities we need to price our derivative. The other two models, on the other hand, are less stable but are calibrated on the whole market, hence can be more flexible in pricing derivatives with different maturities.

10. Black model

In this section we consider an alternative model for the forward dynamics of the two assets. Specifically, the two assets forward exponents are defined as:

$$X_i(t) = -\frac{1}{2}\sigma_i^2 t + \sigma_i W_i(t) \quad \text{with} \quad \text{corr}(W_i(t), W_j(t)) = \rho$$

The model we take into consideration is the Black model since it is based on the Brownian Motions $W_i(t)$, $W_j(t)$. Black model is widely used in financial engineering for pricing derivatives and modeling asset prices due to its relatively simple and intuitive formulation, which makes it easier to understand and implement. The model describes the forward price of the underlying asset as following a log-normal distribution, which aligns with the empirical observations of asset price behaviour.

$$\begin{cases} dF(s, t) = F(s, t)\sigma dW_s \\ F(t_0, t) = F_0 \end{cases}$$

Black model has just one parameter σ that needs calibration. In this particular case, because of the bivariate nature of the market, we have to calibrate the model on the two implied volatilities surfaces separately and select a ρ that matches the previously observed historical correlation.

10.1. Calibration

Calibration of the Black model requires computing the prices at each maturity as functions of the implied volatilities for both indexes. The computed prices are then compared with the market prices to minimize the sum of the root mean squared errors RMSE over all maturities.

MATLAB provides a built-in function called "blkprice" that calculates the model prices for call and put options given the known market parameters. Although we have prices for every strike, both for calls and puts, we filter the prices to consider only the out-of-the-money call and put options due to the reasons already explained in previous sections.

The calibrated volatilities are the σ_{Black}^{US} and σ_{Black}^{EU} that are the ones which minimize the objective function. To build the objective function we consider the RMSE over each maturity and then sum the RMSE vector.

$$rmse_i = \sqrt{\frac{\sum_{j=1}^{N_i} (price_{OTM,j}^{mkt} - price_{OTM,j}^{mod})^2}{N_i}}$$

where N_i is the number of strikes for the i^{th} maturity.

$$\text{objective function} = \sum_{i=1}^{\#\text{maturities}} rmse_i$$

To perform the calibration we leverage the built-in function *fmincon*, which in this case resulted robust and stable for any initial condition.

Calibrated Parameters and Metrics

By applying the algorithm described above, the following parameters are obtained:

Parameter	σ_{EU}	σ_{US}
Black	0.1569	0.1641

Table 22: Calibrated parameters for EU and US Markets via Black Model

This calibration produces reasonable results, finding market volatilities in keeping with what has been found in previous calibrations via Lévy. As before, the volatilities of the two markets are of the same order of magnitude,

with that of the US market confirmed to be slightly higher.

This calibration procedure is convenient in terms of computational cost because it is based on a simple model. Although Black's model is not as accurate as the one based on Lévy's multivariate processes, in that it is less precise than the latter in identifying any market shocks and the true characteristics of volatility, it is an excellent way to have a benchmark in terms of the order of magnitude of the parameters.

The reduced accuracy of Black's model compared to the competing Lévy model can be observed by calculating the Average Percentage Error:

Metric	Average Percentage Error	
	EU Market	US Market
Market Price	8.8059%	13.168%

Table 23: Relative Percentage Error for EU and US Markets (filtered RMSE)

On average, pricing the options with the calibrated parameters via Black produces a higher error than with Lévy. The lower accuracy of option prices found with the calibrated parameters compared to actual bid-ask prices can also be appreciated graphically:

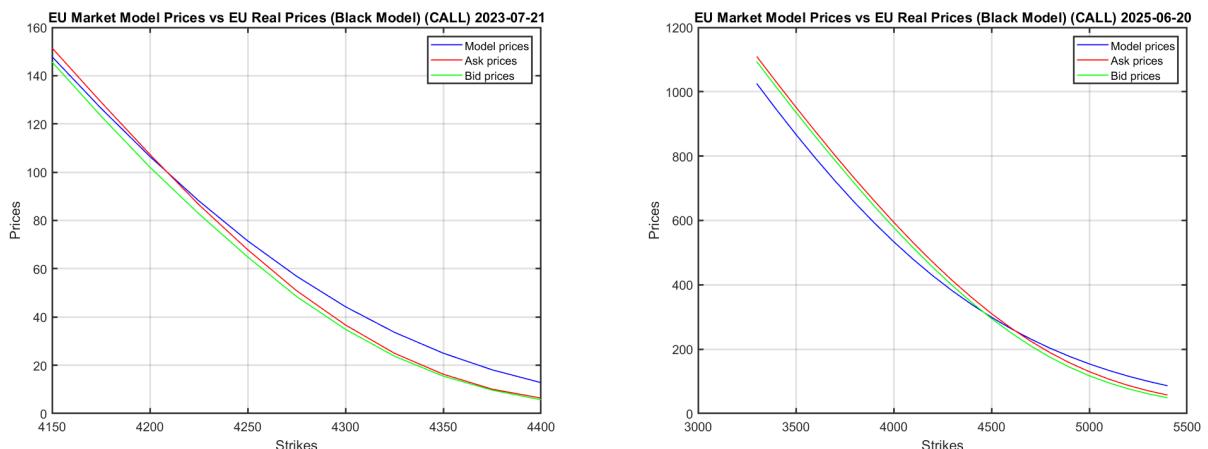


Figure 40: First expiry's call prices for EU Market via Black model calibrated

Figure 41: Last expiry's call prices for EU Market via Black model calibrated

Correlation in Black model

The two methods also have an important difference in terms of maximum achievable correlation. In fact, Black's method, which is less rigorous in describing the market, has weaker assumptions about the value that ρ can achieve.

As seen in section 4, Lévy's model describes assets in terms of their idiosyncratic component and systemic component, and therefore their correlation can be easily rewritten as a function of these components

$$\rho_{12}^X = \frac{\sqrt{k_1 k_2}}{\nu_Z} * sign(a_1 a_2).$$

The maximum value attainable by ρ is then capped by the model's constraints. In contrast, Black's model expresses the two assets as geometric BMs: their description is less detailed and their correlation is that between two BMs. Consequently, as far as follows from the theory of stochastic calculus on BMs, the maximum correlation that the two assets can present is exactly equal to the maximum correlation that two BMs can exhibit, i.e., $|\rho|=1$.

11. Derivative Pricing

In this section, we are pricing a one-year maturity derivative based on the S&P 500 and EURO-STOXX 50 indexes having the following payoff:

$$(S_{US}(t) - S_{US}(t_0))^+ \cdot \mathbb{1}_{S_{EU}(t) < 0.95 \cdot S_{EU}(t_0)}$$

In order to compute its price we exploit both Lévy and Brownian models as we show in the following subsections. Moreover, we derive a semi-closed formula, where a numerical method is needed to compute the final integral.

11.1. Lévy model

Monte Carlo approach can be applied to the Lévy model to price derivatives. We implemented two algorithms to achieve this result. The first step coincides for both method. It is required to compute the discounts $B_i(0, T)$ with $T = 1$ year in order to discount at the end of the method to get the derivative fair price and in order to compute the forward values of the assets since $F_i(0, T) = S_i(0)/B_i(0, T)$. We find two values for the discounts, one per market since we derive it from the discounts of each market, we interpolate on the corresponding zero rates and we get back to the discount at the desired maturity. It is time to focus on the two algorithms separately.

Lévy 1

The first method exploits the relation presented earlier in this report $X_i(t) = Y_i(t) + a_i Z(t)$ to get the marginal NIG processes. From equations (10) in **B&B** we obtain the fundamental idiosyncratic parameters to implement the method which consists of simulating the idiosyncratic and systematic parts of the marginal NIG or VG processes separately and combining them together afterwards.

$$\begin{aligned}\beta_{US} &= \theta_{US} - a_{US}\beta_Z \\ \beta_{EU} &= \theta_{EU} - a_{EU}\beta_Z \\ \gamma_{US} &= \sqrt{\sigma_{US}^2 - a_{US}^2\gamma_Z^2} \\ \gamma_{EU} &= \sqrt{\sigma_{EU}^2 - a_{EU}^2\gamma_Z^2}\end{aligned}$$

Moreover, we compute the drift compensator for the two markets as

$$\begin{aligned}\text{drift compensator}_{US}^{NIG} &= -\frac{1}{k_{US}}(1 - \sqrt{1 - 2k_{US}\theta_{US} - k_{US}\sigma_{US}^2}) \\ \text{drift compensator}_{EU}^{NIG} &= -\frac{1}{k_{EU}}(1 - \sqrt{1 - 2k_{EU}\theta_{EU} - k_{EU}\sigma_{EU}^2}) \\ \text{drift compensator}_{US}^{VG} &= -\frac{1}{k_{US}}\log(1 - \theta_{US}\kappa_{US} - \frac{1}{2}\kappa_{US}\sigma_{US}^2) \\ \text{drift compensator}_{EU}^{VG} &= -\frac{1}{k_{EU}}\log(1 - \theta_{EU}\kappa_{EU} - \frac{1}{2}\kappa_{EU}\sigma_{EU}^2)\end{aligned}$$

We simulate $Y_{US}(t)$, $Y_{EU}(t)$ and $Z(t)$.

$$\begin{aligned}Y_{US}(t) &= -\gamma_{US}^2(\frac{1}{2} + \beta_{US})G_{US}t + \gamma_{US}\sqrt{tG_{US}}g_{US} \\ Y_{EU}(t) &= -\gamma_{EU}^2(\frac{1}{2} + \beta_{EU})G_{EU}t + \gamma_{EU}\sqrt{tG_{EU}}g_{EU} \\ Z(t) &= -\gamma_Z^2(\frac{1}{2} + \beta_Z)G_Zt + \gamma_Z\sqrt{tG_Z}g_Z\end{aligned}$$

where G_i are Inverse Gaussian random variables with unitary mean and variance $\frac{\nu_i}{\Delta t}$ in the case of the NIG model, while G_i are extracted from the Variance Gamma in the VG model. Moreover, g are independent standard normal.

Then, we can compute the marginal processes by the previously introduced linear combination of auxiliary processes such that:

$$\begin{aligned} X_{US} &= Y_{US}(t) + a_{US}Z(t) \\ X_{EU} &= Y_{EU}(t) + a_{EU}Z(t) \end{aligned}$$

Then, we obtain the dynamics of the assets as follows:

$$\begin{aligned} S_{US} &= F_{US}(0) \cdot e^{\text{drift compensator}_{US} \cdot T + X_{US}(t)} \\ S_{EU} &= F_{EU}(0) \cdot e^{\text{drift compensator}_{EU} \cdot T + X_{EU}(t)} \end{aligned}$$

The final step is to average the discounted payoffs obtained from the simulated paths.

$$Price(t_0) = B(0, T) \mathbb{E}[(S_{US}(t) - S_{US}(t_0))^+ \cdot \mathbb{1}\{S_{EU}(t) < 0.95 \cdot S_{EU}(t_0)\}]$$

Lèvy 2

The second algorithm simulates two Normal Inverse Gaussian (NIG), or two Variance Gamma (VG), processes $X_{US}(t)$, $X_{EU}(t)$, using the calibrated parameters directly. As a result, the difference between the two methods relies only on the implementation of the simulations to obtain the NIG (or VG) processes. We focus on $f_t \equiv \log \frac{F_t}{F_0}$.

The dynamics up to time t can be written as

$$f_t = \sqrt{t-t_0} \sigma_i \sqrt{G_i} g - \left(\frac{1}{2} + \theta_i\right)(t-t_0) \sigma_i^2 G_i - \frac{t-t_0}{k_i} (1 - \sqrt{1 - 2k_i \theta_i - k_i \sigma_i^2}), \quad i = \text{US, EU}$$

where $G_i > 0$ is an Inverse Gaussian with unitary mean and variance $\frac{k_i}{\Delta t}$ (or a Variance Gamma) and g is a standard normal.

In order to preserve the correlation between the two indexes through the simulations, g is set to be a matrix containing one column per market filled with correlated simulated standard random variables. The correlation structure ensures that the simulated returns for the US and EU markets reflect the empirical correlation between them.

Using the simulated paths of $X_{US}(t)$ and $X_{EU}(t)$, we compute the derivative payoff in each simulation through the following relationships:

$$\begin{aligned} S_{US}(t) &= F_{US}(0) \cdot e^{f_t^{US}} \\ S_{EU}(t) &= F_{EU}(0) \cdot e^{f_t^{EU}} \end{aligned}$$

Hence, we compute the average of the simulated payoffs and discount it with the usual procedure.

11.2. Black model

The derivative price can be computed via Monte Carlo simulations using the Black model as well. Specifically, we simulate the assets as Geometric Brownian Motions, as their dynamics in the Black model suggest. To simulate the values at time T , we use Z_{US} and Z_{EU} which are correlated Gaussian random variables. Each path represents a possible future scenario of the index prices.

$$\begin{aligned} S_{US}(T) &= F_{US}(0) \cdot \exp \left\{ \left(-\frac{1}{2} \sigma_{US}^2 \right) \cdot T + \sigma_{US} \sqrt{T} \cdot Z_{US} \right\} \\ S_{EU}(T) &= F_{EU}(0) \cdot \exp \left\{ \left(-\frac{1}{2} \sigma_{EU}^2 \right) \cdot T + \sigma_{EU} \sqrt{T} \cdot Z_{EU} \right\} \\ \begin{pmatrix} Z_{US} \\ Z_{EU} \end{pmatrix} &\sim \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right) \end{aligned}$$

For each simulated path, we compute the payoff of the derivative at maturity. This results in a vector of possible payoffs, one per simulation. Then, the fair price of the derivative is estimated by taking the discounted average across the simulations of the payoffs.

11.3. Semi-closed formula

The semi-closed form solution for a derivative price relies on the concept of the expected discounted payoff. This means that we begin our procedure from the payoff the derivative delivers at its maturity date t , we weight each scenario by its probability and discount it back to the present time t_0 using the stochastic discount $D(t_0, t)$. Finally, we average this discounted payoff across all scenarios to arrive at the derivative's fair price.

The semi-closed formula is computed within the Black framework. Specifically, we utilize the Black & Scholes model to express the spot dynamics of the two indexes.

$$\text{Spot dynamics: } S_i(t) = S_i(0)e^{(r_i - \frac{1}{2}\sigma_i^2)t + \sigma_i W_i(t)}, \quad \text{where } i = US, EU$$

We assume that $W_{US}(t)$ and $W_{EU}(t)$ are correlated Brownian Motions with correlation ρ . We can divide the process to obtain the final result into three main steps:

- **Step 1: Integral form.**

In order to express the expectation as an integral we require the joint density probability function. Utilizing the Black & Scholes model, we can think of our underlying indexes in terms of the correlated Brownian Motions. At time t the two correlated BMs become correlated Gaussian random variables. From probability theory, we can express the joint probability density function as the product of the conditional distributions of one variable given the other and the marginal distribution of the second one.

$$f_{X,Y}(x, y) = f_{Y|X}(y|x)f_X(x) = f_{X|Y}(x|y)f_Y(y)$$

Hence, we can formulate the following integral:

$$B(0, T) \int \{(S_{US}(T) - S_{US}(0))^+ \cdot \mathbb{1}\{W_{EU}(T) < A\} f_{W_{US}|W_{EU}} dw_{US} | w_{EU}\} f_{W_{EU}}(w_{EU}) dw_{EU}$$

Given that the indicator function depends on the conditioning variable, we can take it outside the inner integral, and treat it as the integration limits of the outer integral whose integration variable is W_{EU} .

$$B(0, T) \int_{-\infty}^A \mathbb{E}_0[(S_{US}(T) - S_{US}(0))^+ | W_{EU}(T) = w_{EU}] f_{W_{EU}}(w_{EU}) dw_{EU}$$

- **Step 2: Computation of the expected value in the integral.**

Since conditional expectation as well as expectation is a linear operator, we are allowed to split it up into two conditional expectations. We focus on them separately.

$$\underbrace{\mathbb{E}_0[S_{US}(T) \cdot \mathbb{1}\{S_{US}(T) \geq S_{US}(0)\} | W_{EU}(T) = w_{EU}]}_{(I)} - \underbrace{\mathbb{E}_0[S_{US}(0) \cdot \mathbb{1}\{S_{US}(T) \geq S_{US}(0)\} | W_{EU}(T) = w_{EU}]}_{(II)}$$

Let's begin by considering the second one. We are able to compute it thanks to probabilistic theory, in fact we observe that the indicator function can be written as

$$\mathbb{1}\{W_{US}(T) \geq -\frac{(r - \frac{\sigma_{US}^2}{2})T}{\sigma_{US}}\}$$

and the density function of $W_{US}(t)$ is known, as it is a correlated Gaussian random variable. Indeed, after some computations, we get the value of the second expectation which is

$$\mathcal{N}(-B)S_{US}(0)$$

For the first expectation the core idea remains the same, requiring just a few more steps. The result is as follows:

$$S_{US}(0) \exp\left((r_{US} - \frac{1}{2}\sigma_{US}^2)T + \frac{T(1 - \rho^2)\sigma_{US}^2}{2} + \sigma_{US}\rho\xi\right) \mathcal{N}(D)$$

- **Step 3: Final result.**

In the end we have to combine the previous results in order to complete the task. The result consists of an integral that needs to be computed by a numerical technique. The theoretical result is:

$$B(0, T)S_{US}(0) \int_{-\infty}^A \{ \exp\left(r_{US}T - \frac{1}{2}\sigma_{US}^2\rho^2T + \sigma_{US}\rho\xi\right) \mathcal{N}(D) - \mathcal{N}(B) \} \frac{1}{\sqrt{2\pi T}} \exp\left(-\frac{\xi^2}{2T}\right) d\xi$$

If the reader is interested in a step-by-step deduction of the semi-closed formula, they can find it in Appendix B. Appendix B, along with the theoretical discussion above, should provide a complete description of the procedure to obtain the desired formula.

11.4. Comments and Results

In this section we analyze the pricing results obtained via the several models implemented previously and we compare the obtained prices of the exotic derivative.

Semi-closed formula and Black Model

In the present subsection, we show the results derived from the semi-closed formula and the Black Model. We discuss them together since they are strictly connected. Indeed, the semi-closed formula is derived from the assumptions of the Black Model.

The semi-closed formula contains an integral that can be solved by the built-in MATLAB function *integral* that implements a quadrature method to approximate the result.

Method	Price(€)	Confidence Interval(€)
Black Model	14.71	[14.65, 14.77]
Semi-Closed Formula	14.69	N/A

Table 24: Prices and Confidence Intervals Black and Semi-Closed formula

Although the underlying spot price is 4424.46€ the derivative price is - for the closed formula - only 14.69€, the 0.33% of the spot price. This is due to the high correlation between the two indexes and the nature of the payoff. In fact, the only combination that allows the owner of the option to get the money is that the American index at maturity is beyond the US spot price and the European one below its 95% of its spot price. As mentioned before, we expect that this situation does not happen frequently because of the high correlation between the S&P 500 and the EURO-STOXX 50.

A notable result is the fact that the semi-closed formula price falls into the 99% Confidence Interval of the Black model price.

Lévy

Now we analyze the prices obtained via the Lévy approach, considering the parameters calibrated above.

First, we focus on the Unweighted model, which neglects the last maturity of the S&P 500 index. By taking into account the NIG subordinator we achieve the following results:

Method	Price €	Confidence Interval €
Lévy 1	16.34	[16.24, 16.44]
Lévy 2	15.53	[15.44, 15.63]

Table 25: Prices and Confidence Intervals for Lévy Models 'Unweighted' (NIG)

As expected, the price of the derivative is slightly higher than the one obtained via the semi-closed formula and Black simulation. Indeed, as highlighted in the previous sections, the Lévy model for this particular choice does not capture the entire correlation between the markets (historical yearly correlation $\rho = 0.8010$). Hence, since the two indicators function of the payoff requires opposite movements of the two indices, having a lower correlation makes the trigger of the two functions easier, resulting in a payoff that is more often greater than zero. This translates into a higher price of the derivative.

As proof of this statement, we can explore different initial conditions for the ν calibration, which yield slightly different values of the calibrated parameters and consequently differences in the calibrated correlation. Indeed, if

we consider the following initial guesses:

$$X_1 = 0.2 \cdot \text{ones}(1, 3)$$

$$X_2 = [4 \quad 0.1 \quad 4]$$

we obtain respectively a calibrated correlation of 0.22506 and 0.34033, which are again respectively lower and greater than the one obtained with the parameters discussed above, $\rho = 0.2426$. The prices as a consequence result to be higher in the first scenario and lower in the second, following the same reasoning as before.

Method	Price €	Confidence Interval €
Lévy Model X_1	16.73	[16.63, 16.84]
Lévy Model X_2	14.08	[13.99, 14.17]

Table 26: Prices and Confidence Intervals for Lévy Models 'Unweighted'(different initial guess)

Moreover, we computed the prices using the Variance Gamma (VG) model. We attempted to implement an algorithm to simulate a random variable distributed as VG, however, we are not completely sure that our methodology is working properly, and it could require longer computational time but is still feasible. Indeed, the prices of this method deviate slightly from those obtained with other models. The following table reports the prices via VG:

Method	Price €	Confidence Interval €
Lévy 1	31.91	[31.77, 32.05]
Lévy 2	16.11	[15.97, 16.26]

Table 27: Prices and Confidence Intervals for Lévy Models 'Unweighted'(VG)

Secondly, we focus on the Filtered RMSE. This method takes into account all available maturities and yields the following results for the NIG subordinator.

Method	Price €	Confidence Interval €
Lévy 1	17.38	[17.28, 17.48]
Lévy 2	16.79	[16.69, 16.88]

Table 28: Prices and Confidence Intervals for Lévy Models 'Filtered RMSE'(NIG)

Since the ρ explained by the model is lower than the historical one, following the previous reasoning, we expect to obtain larger prices than those derived via the semi-closed formula.

In the following table, we provide the derivative prices using the VG model within the Filtered RMSE framework.

Method	Price €	Confidence Interval €
Lévy 1	24.27	[24.16, 24.39]
Lévy 2	11.05	[10.93, 11.17]

Table 29: Prices and Confidence Intervals for Lévy Models 'Filtered RMSE'(VG)

Also in this case, results may be a little inconsistent due to the same reasons as the previous VG point. Lastly, we present the results obtained with both Normal Inverse Gaussian and Variance Gamma subordinators in the time-window framework. This model is fitted to the expiries close to the 1-year maturity of the exotic derivative we are pricing. Hence, we expect the presented prices to be accurate and capture the dynamics of the underlying assets rather well.

Method	Price €	Confidence Interval €
Lévy 1	17.61	[17.51, 17.71]
Lévy 2	18.14	[18.05, 18.24]

Table 30: Prices and Confidence Intervals for Lévy Models 'Time Window' (NIG)

As pointed out previously, the prices are higher than those obtained via the semi-closed formula and Black's model, since the correlation captured in this case is just above 50%.

Method	Price €	Confidence Interval €
Lévy 1	14.75	[14.62, 14.88]
Lévy 2	14.83	[14.68, 14.97]

Table 31: Prices and Confidence Intervals for Lévy Models 'Time Window' (VG)

12. Final consideration on the implemented models

To conclude this report, we summarize the results obtained and compare the implemented models. The Black model sets a benchmark for all models in finance, being one of the most popular approaches. This is thanks to its simplicity and parsimony, which help interpret the financial meaning behind the computation. Indeed, calibration via the Black model returns a unique parameter to describe the volatility of the market. However, we know this is an oversimplification of the real world. Traded volatility follows a surface and changes over time and strike. Moreover, this simple approach based on Geometric Brownian Motion (GBM) does not admit jumps and shocks in equity dynamics, which instead occur frequently in real markets. Additionally, equities tend to have fatter tails than those predicted by GBM.

Hence, other models have been implemented to better explain and predict these behaviors. The Lévy model, in particular, is widely used as it is parsimonious but at the same time is able to capture market shocks by providing a better description of volatility.

While the Black model uses only σ^{Blk} to describe the entire volatility structure, the Lévy model introduces a different parameterization via σ^{Levy} , κ , and θ . These parameters represent, respectively, the average volatility, the variance rate, and the skewness. This triplet guarantees a better approximation of the volatility surface. However, this model also has some drawbacks. Firstly, to calibrate its parameters, some additional inputs are required such as the subordinator, which, as we have seen, could be a crucial point considering different markets. Moreover, we have to implement a numerical method to solve the integral that arises from the Lewis formula. There are several methods such as quadrature and fast Fourier transform, but these numerical approximations also require setting additional hyperparameters.

Furthermore, when considering multivariate processes, the correlation between assets that Lévy is able to capture is limited by the convolution condition. Hence, we are not always able to reproduce the link between assets, while the simple Black model, being driven by simple Brownian motions, is able to achieve any level of correlation.

This was analyzed in our work; while the Black model was able to capture the entire historical correlation between the two indices, the Lévy model was limited by its own calibrated parameters. This has repercussions when pricing a derivative.

The prices computed for the derivative are slightly different between these models. On one side, Lévy is able to capture shocks in the market, the fatter tail of the equity dynamics and the shape of the traded volatility surface, leading to a better understanding of the risks, which can be reflected in the price. Moreover, another influence is brought by the calibrated correlation. As explained in the previous section, a lower correlation may lead the indicator functions of the payoff to be positively triggered together, bringing the price of the derivative up. However, there is a trade-off between the correlation and the risks that Lévy takes into account when pricing the derivative. Hence, in conclusion, we expect Lévy to be more precise in estimating exotic derivative prices that are highly impacted by the stochasticity of volatility.

As a last remark, we also highlight the difference in numerical computation. Indeed, the calibration of Lévy can require more expensive computational time, which can rapidly increase when considering many assets and maturities. On the other hand, the calibration of the Black model is cheap and robust, guaranteeing a reliable benchmark.

13. Appendix A: L. Ballotta & E. Bonfiglioli (2014)

Proposition 13.1. (*Prop. 1*) Let $Z(t)$, $Y_j(t)$, $j = 1, \dots, n$, be independent Lévy processes on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with characteristic function $\Phi_z(u; t)$ and $\Phi_{Y_j}(u; t)$, for $i = 1, \dots, n$, respectively. Then, for $a_j \in \mathbb{R}$, $j = 1, \dots, n$,

$$\mathbf{X}(t) = (X_1(t), \dots, X_n(t))^T = (Y_1(t) + a_1 Z(t), \dots, Y_n(t) + a_n Z(t))^T$$

is a Lévy process on \mathbb{R}^n . The resulting characteristic function is:

$$\phi_{\mathbf{X}}(\mathbf{u}; t) = \phi_Z \left(\sum_{j=1}^n a_j u_j; t \right) \prod_{j=1}^n \phi_{Y_j}(u_j; t), \quad \mathbf{u} \in \mathbb{R}^n$$

Proposition 13.2. (*Section 2.2 - Def.*) A subordinated Brownian Motion $X = (X(t) : t \geq 0)$ is a Lévy process obtained by observing a (arithmetic) BM on a time scale governed by an independent subordinator, i.e. an increasing, positive Lévy process. Hence, $X(t)$ has general form

$$X(t) = \theta G(t) + \sigma W(G(t)), \quad \theta \in \mathbb{R}, \quad \sigma > 0$$

where $W = (W(t) : t \geq 0)$ is a BM and $G = (G(t) : t \geq 0)$ is a subordinator independent of W . The resulting characteristic function is

$$\Phi_X(u, t) = e^{t\varphi_G(u\theta + iu(\sigma^2/2))}, \quad u \in \mathbb{R}$$

where $\varphi_G(\cdot)$ denotes the characteristic exponent of the subordinator.

14. Appendix B: Some definitions & computations

Proposition 14.1. (*Def.*) In mathematics, a degenerate distribution is the probability distribution of a discrete random variable whose support consists of only one value. Examples include a two-headed coin and rolling a die whose sides all show the same number. While this distribution does not appear random in the everyday sense of the word, it does satisfy the definition of random variable.

Proposition 14.2. (*Correlation index derivation*)

$$\begin{aligned}
 \rho_{12}^X &= \frac{a_1 a_2 \text{Var}(Z(1))}{\sqrt{\text{Var}(X_1(1))} \sqrt{\text{Var}(X_2(1))}} = \\
 &= \frac{a_1 a_2 (\gamma_Z^2 + \beta_Z^2 \nu_Z)}{\sqrt{\sigma_1^2 + \theta_1^2 k_1} \sqrt{\sigma_2^2 + \theta_2^2 k_2}} = \\
 \left\{ \frac{\sigma_i^2}{k_i \theta_i^2} = c \right\} &= \frac{a_1 a_2 (\gamma_Z^2 + \beta_Z^2 \nu_Z)}{\sqrt{\sigma_1^2 + \frac{\sigma_1^2}{c}} \sqrt{\sigma_2^2 + \frac{\sigma_2^2}{c}}} = \\
 &= \frac{a_1 a_2 (\gamma_Z^2 + \beta_Z^2 \nu_Z)}{\sigma_1 \sigma_2 (1 + \frac{1}{c})} = \\
 \left\{ \frac{\gamma_Z^2}{\nu_Z \beta_Z^2} = c \right\} &= \frac{a_1 a_2 (\gamma_Z^2 + \beta_Z^2 \nu_Z)}{\sigma_1 \sigma_2 \frac{(\gamma_Z^2 + \beta_Z^2 \nu_Z)}{\gamma_Z^2}} = \\
 \text{Equation [9].2} \quad \left\{ \sigma_i^2 = \frac{\nu_Z a_i^2 \gamma_Z^2}{k_i} \right\} &= \frac{|a_1 a_2| \sqrt{\gamma_Z^2}}{\frac{\nu_Z a_1 a_2 \gamma_Z^2}{\sqrt{k_1 k_2}}} = \\
 &= \frac{\sqrt{k_1 k_2}}{\nu_Z} * \text{sign}(a_1 a_2) = \\
 \left\{ k_i = \frac{\nu_Z \nu_i}{\nu_Z + \nu_i} \right\} &= \text{sign}(a_1 a_2) * \sqrt{\frac{\nu_1 \nu_2}{(\nu_1 + \nu_Z)(\nu_2 + \nu_Z)}}
 \end{aligned}$$

15. Appendix C: Calibrated parameters idiosyncratic and systematic NIG and VG

In the following table, we present the calibrated parameters for the idiosyncratic and systematic factors. In particular, we show the parameters from the NIG subordinators:

NIG Model					
	Parameter	a	gamma	nu	Beta
Unweighted	Y_EU	0.7835	0.1174	0.9315	-0.1438
	Y_US	2.1143	0.1077	8.0652	-0.0448
	Z	-	0.0534	7.3476	-0.0233
Filtered RMSE	Y_EU	0.4267	0.1197	0.7848	-0.1327
	Y_US	0.8445	0.0951	4.8679	-0.0424
	Z	-	0.1323	3.5282	-0.0692
Time Window	Y_EU	0.4368	0.1066	1.6268	-0.1022
	Y_US	0.6535	0.0873	5.4296	-0.0458
	Z	-	0.1601	3.7806	-0.1007

Table 32: Calibrated Parameters for NIG Model under Different Scenarios

In the following table, we present the calibrated parameters for the idiosyncratic and systematic factors. Specifically, we show the parameters from the VG subordinators:

VG Model					
	Parameter	a	gamma	nu	Beta
Unweighted	Y_EU	0.8739	0.1259	1.0125	-0.1063
	Y_US	1.7474	0.1055	5.7632	-0.0374
	Z	-	0.0687	4.4576	-0.0276
Filtered RMSE	Y_EU	0.1827	0.1217	0.6542	-0.1125
	Y_US	0.3141	0.0890	3.6190	-0.0350
	Z	-	0.3762	2.0532	-0.1962
Time Window	Y_EU	-0.0444	0.1160	1.8522	-0.0722
	Y_US	-0.0598	0.0953	4.9744	-0.0362
	Z	-	1.9627	3.2867	0.9177

Table 33: Calibrated Parameters for VG Model under Different Scenarios

16. Appendix D: Proof of the semi-closed formula for the derivative

Payoff: $(S_{US}(\text{ttm}) - S_{US}(t_0))^+ \cdot \mathbb{1}\{S_{EU}(\text{ttm}) < 0.95 \cdot S_{EU}(t_0)\}$

Forward dynamics: $F_i(t) = F_i(0)e^{-\frac{1}{2}\sigma_i^2 t + \sigma_i W_i(t)}$ $i = US, EU$

We use the Black & Scholes model to compute the dynamics of the spot.

We know that:

- $W_i(t) \sim \mathcal{N}(0, t)$,
- $\text{corr}(W_i(t), W_j(t)) = \rho$,
- $t_0 = 0$,
- $\text{ttm} = T = 1 \text{ year}$

Spot dynamics: $S_i(t) = S_i(0)e^{(r_i - \frac{1}{2}\sigma_i^2)t + \sigma_i W_i(t)}$ $i = US, EU$

We take the expected value in t_0 of the Payoff to get the price.

$$\begin{aligned} p &= \mathbb{E}_0[D(0, T)\{\text{Payoff}\}] \\ &= \mathbb{E}_0[D(0, T)\{(S_{US}(T) - S_{US}(0))^+ \cdot \mathbb{1}\{S_{EU}(T) < 0.95 \cdot S_{EU}(0)\}\}] \\ &= B(0, T)\mathbb{E}_0[(S_{US}(T) - S_{US}(0))^+ \cdot \mathbb{1}\{S_{EU}(T) < 0.95 \cdot S_{EU}(0)\}] \\ &= B(0, T)\mathbb{E}_0[(S_{US}(T) - S_{US}(0))^+ \cdot \mathbb{1}\{S_{EU}(0)e^{(r_{EU} - \frac{1}{2}\sigma_{EU}^2)T + \sigma_{EU}W_{EU}(T)} < 0.95 \cdot S_{EU}(0)\}] \\ &= B(0, T)\mathbb{E}_0[(S_{US}(T) - S_{US}(0))^+ \cdot \mathbb{1}\{W_{EU}(T) < \frac{\log(0.95) - (r - \frac{1}{2}\sigma_{EU}^2)T}{\sigma_{EU}^2}\}] \\ &= B(0, T)\mathbb{E}_0[(S_{US}(T) - S_{US}(0))^+ \cdot \mathbb{1}\{W_{EU}(T) < A\}], \text{ where } A = \frac{\log(0.95) - (r - \frac{1}{2}\sigma_{EU}^2)T}{\sigma_{EU}^2} \end{aligned}$$

At this step, we exploit the relation between the joint probability density function and the marginals in the continuous case.

$$f_{X,Y}(x, y) = f_{Y|X}(y|x)f_X(x) = f_{X|Y}(x|y)f_Y(y)$$

where $f_{Y|X}(y|x)$ and $f_{X|Y}(x|y)$ are the conditional distributions of Y given $X = x$ and X given $Y = y$ respectively, and $f_X(x)$ and $f_Y(y)$ are the marginal distributions for X and Y respectively.

$$\begin{aligned} &= B(0, T) \int \int (S_{US}(T) - S_{US}(0))^+ \cdot \mathbb{1}\{W_{EU}(T) < A\} f_{W_{US}, W_{EU}}(w_{US}, w_{EU}) dw_{US} dw_{EU} \\ &= B(0, T) \int \int \{(S_{US}(T) - S_{US}(0))^+ \cdot \mathbb{1}\{W_{EU}(T) < A\} f_{W_{US}|W_{EU}} dw_{US}|w_{EU}\} f_{W_{EU}}(w_{EU}) dw_{EU} \\ &= B(0, T) \int \mathbb{E}_0[(S_{US}(T) - S_{US}(0))^+ \cdot \mathbb{1}\{W_{EU}(T) < A\} | W_{EU}(T) = w_{EU}] f_{W_{EU}}(w_{EU}) dw_{EU} \\ &= B(0, T) \int \mathbb{E}_0[(S_{US}(T) - S_{US}(0))^+ | W_{EU}(T) = w_{EU}] \mathbb{1}\{W_{EU}(T) < A\} f_{W_{EU}}(w_{EU}) dw_{EU} \\ &= B(0, T) \int_{-\infty}^A \mathbb{E}_0[(S_{US}(T) - S_{US}(0))^+ | W_{EU}(T) = w_{EU}] f_{W_{EU}}(w_{EU}) dw_{EU} \end{aligned}$$

It is time to focus on the expected value contained into the integral.

$$\begin{aligned}
& \mathbb{E}_0[(S_{US}(T) - S_{US}(0))^+ | W_{EU}(T) = w_{EU}] \\
&= \mathbb{E}_0[(S_{US}(T) - S_{US}(0)) \cdot \mathbb{1}\{S_{US}(T) \geq S_{US}(0)\} | W_{EU}(T) = w_{EU}] \\
&= \overbrace{\mathbb{E}_0[S_{US}(T) \cdot \mathbb{1}\{S_{US}(T) \geq S_{US}(0)\} | W_{EU}(T) = w_{EU}]}^{(I)} - \overbrace{\mathbb{E}_0[S_{US}(0) \cdot \mathbb{1}\{S_{US}(T) \geq S_{US}(0)\} | W_{EU}(T) = w_{EU}]}^{(II)}
\end{aligned}$$

Let's solve the two expected values separately. We start with (II) since it is necessary for (I), too. Moreover, we introduce a new notation to light up the writing

$$\begin{aligned}
W_{US} &= X \\
w_{US} &= x \\
w_{EU} &= \xi \\
A &= \frac{\log(0.95) - (r - \frac{1}{2}\sigma_{EU}^2)T}{\sigma_{EU}^2} \\
B &= \frac{(r_{US} - \frac{\sigma_{US}^2}{2})T + \rho\xi\sigma_{US}}{\sigma_{US}\sqrt{T(1 - \rho^2)}} \\
C &= \frac{(r_{US} - \frac{1}{2}\sigma_{US}^2)T}{\sigma_{US}} \\
D &= \frac{(r_{US} - \frac{1}{2}\sigma_{US}^2)T + [(T(1 - \rho^2)\sigma_{US} + \rho\xi]\sigma_{US}}{\sqrt{T(1 - \rho^2)}\sigma_{US}} = B - \sqrt{T(1 - \rho^2)}\sigma_{US}
\end{aligned}$$

Solution of (II)

$$\begin{aligned}
(II) : \mathbb{E}_0[S_{US}(0) \cdot \mathbb{1}\{S_{US}(T) \geq S_{US}(0)\} | W_{EU}(T) = \xi] &= \\
&= \mathbb{E}_0[S_{US}(0) \cdot \mathbb{1}\{S_{US}(0)e^{(r_{US} - \frac{1}{2}\sigma_{US}^2)T + \sigma_{US}X(T)} \geq S_{US}(0)\} | W_{EU}(T) = \xi] \\
&= S_{US}(0)\mathbb{E}_0[S_{US}(0) \cdot \mathbb{1}\{S_{US}(0)e^{(r_{US} - \frac{1}{2}\sigma_{US}^2)T + \sigma_{US}X(T)} \geq S_{US}(0)\} | W_{EU}(T) = \xi] \\
&= S_{US}(0)\mathbb{E}_0[\mathbb{1}\{X(T) \geq -\frac{(r - \frac{\sigma_{US}^2}{2})T}{\sigma_{US}}\} | W_{EU}(T) = w_{EU}] \stackrel{(*)}{=}
\end{aligned}$$

We compute the bivariate conditional distribution from probabilistic theory.

$$\begin{aligned}
\begin{pmatrix} X \\ Y \end{pmatrix} &\sim \mathcal{N} \left(\begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \begin{pmatrix} \sigma_X^2 & cov_{XY} \\ cov_{XY} & \sigma_Y^2 \end{pmatrix} \right) \\
Y|X = x &\sim \mathcal{N}(\mu_Y + cov_{XY}\sigma_X^{-1}(x - \mu_X), \sigma_Y^2 - cov_{XY}^2\sigma_X^{-2}) \\
W_{US}(T) &\sim \mathcal{N}(0, T) \\
W_{EU}(T) &\sim \mathcal{N}(0, T)
\end{aligned}$$

Then, in our case it becomes:

$$\begin{aligned}
X|W_{EU} = \xi &\sim \mathcal{N}(\rho\xi, T - \rho^2T) \\
f_{X|W_{EU} = \xi} &= \frac{1}{\sqrt{2\pi T(1 - \rho^2)}} \exp\left(-\frac{(x - \rho\xi)^2}{2T(1 - \rho^2)}\right)
\end{aligned}$$

$$\begin{aligned}
&\stackrel{(*)}{=} S_{US}(0) \int \frac{1}{\sqrt{2\pi T(1-\rho^2)}} \exp\left(-\frac{(x-\rho\xi)^2}{2T(1-\rho^2)}\right) \mathbb{1}\{x \geq -C\} dx \\
&= S_{US}(0) \int_{-C}^{\infty} \int \frac{1}{\sqrt{2\pi T(1-\rho^2)}} \exp\left(-\frac{(x-\rho\xi)^2}{2T(1-\rho^2)}\right) dx \\
&= \int_{-B}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{\tilde{x}^2}{2}} d\tilde{x} \\
&= S_{US}(0)(1 - \mathcal{N}(-B)) \\
&= S_{US}(0) \mathcal{N}(B)
\end{aligned}$$

$$\begin{aligned}
\tilde{x} &= \frac{(x-\rho\xi)^2}{\sqrt{2T(1-\rho^2)}} \\
d\tilde{x} &= \frac{1}{\sqrt{T(1-\rho^2)}} dx
\end{aligned}$$

Solution of (I)

$$\begin{aligned}
(I) : \mathbb{E}_0[S_{US}(T) \cdot \mathbb{1}\{S_{US}(T) \geq S_{US}(0)\} | W_{EU}(T) = \xi] &= \\
&= \mathbb{E}_0[S_{US}(0) e^{(r_{US} - \frac{1}{2}\sigma_{US}^2)T + \sigma_{US}X(T)} \cdot \mathbb{1}\{S_{US}(0) e^{(r_{US} - \frac{1}{2}\sigma_{US}^2)T + \sigma_{US}X(T)} \geq S_{US}(0)\} | W_{EU}(T) = \xi] \\
&= \mathbb{E}_0[S_{US}(0) e^{(r_{US} - \frac{1}{2}\sigma_{US}^2)T + \sigma_{US}X(T)} \cdot \mathbb{1}\{X(T) \geq -\frac{(r_{US} - \frac{1}{2}\sigma_{US}^2)T}{\sigma_{US}}\} | W_{EU}(T) = \xi] \\
&= \int_{-C}^{\infty} S_{US}(0) \frac{1}{\sqrt{2\pi T(1-\rho^2)}} \exp\left((r_{US} - \frac{1}{2}\sigma_{US}^2)T + \sigma_{US}x - \frac{(x-\rho\xi)^2}{2T(1-\rho^2)}\right) dx \\
&= S_{US}(0) \exp\left(r_{US} - \frac{1}{2}\sigma_{US}^2\right) \int_{-C}^{\infty} \frac{1}{\sqrt{2\pi T(1-\rho^2)}} \exp\left(x\sigma_{US} + \frac{-x^2 + 2x\rho\xi - \rho^2\xi^2}{2T(1-\rho^2)}\right) dx \\
&= S_{US}(0) \exp\left((r_{US} - \frac{1}{2}\sigma_{US}^2)T + \frac{T(1-\rho^2)\sigma_{US}}{2} + \sigma_{US}\rho\xi\right) \int_{-C}^{\infty} \frac{1}{\sqrt{2\pi T(1-\rho^2)}} \exp\left(\frac{(x - (T(1-\rho^2)\sigma_{US} + \rho\xi))^2}{2T(1-\rho^2)}\right) dx \stackrel{(*)}{=}
\end{aligned}$$

We show in this section how to treat the exponent in order to complete the computations.

$$\begin{aligned}
&x\sigma_{US} + \frac{-x^2 + 2x\rho\xi - \rho^2\xi^2}{2T(1-\rho^2)} \\
&= \frac{-x^2 + 2x(T(1-\rho^2)\sigma_{US} + \rho\xi) - \rho^2\xi^2}{2T(1-\rho^2)} \\
&= \frac{-x^2 + 2x(T(1-\rho^2)\sigma_{US} + \rho\xi) - (T(1-\rho^2)\sigma_{US} + \rho\xi)^2 + (T(1-\rho^2)\sigma_{US} + \rho\xi)^2 - \rho^2\xi^2}{2T(1-\rho^2)} \\
&= -\frac{(x - (T(1-\rho^2)\sigma_{US} + \rho\xi))^2}{2T(1-\rho^2)} + \frac{T^2(1-\rho^2)^2\sigma_{US}^2 + \rho^2\xi^2 + 2T(1-\rho^2)\sigma_{US}\rho\xi - \rho^2\xi^2}{2T(1-\rho^2)} \\
&= -\frac{(x - (T(1-\rho^2)\sigma_{US} + \rho\xi))^2}{2T(1-\rho^2)} + \frac{T(1-\rho^2)\sigma_{US}^2 + 2\sigma_{US}\rho\xi}{2}
\end{aligned}$$

$$\begin{aligned}
&\stackrel{(*)}{=} S_{US}(0) \exp \left(\left(r_{US} - \frac{1}{2} \sigma_{US}^2 \right) T + \frac{T(1-\rho^2)\sigma_{US}^2}{2} + \sigma_{US}\rho\xi \right) \int_{-C}^{\infty} \frac{1}{\sqrt{2\pi T(1-\rho^2)}} \exp \left(-\frac{(x-(T(1-\rho^2)\sigma_{US}+\rho\xi))^2}{2T(1-\rho^2)} \right) dx \\
&= S_{US}(0) \exp \left(\left(r_{US} - \frac{1}{2} \sigma_{US}^2 \right) T + \frac{T(1-\rho^2)\sigma_{US}^2}{2} + \sigma_{US}\rho\xi \right) \int_{-D}^{\infty} \frac{1}{\sqrt{2\pi T}} \exp \left(-\frac{\tilde{x}^2}{2T} \right) d\tilde{x} \\
&= S_{US}(0) \exp \left((r_{US} - \frac{1}{2} \sigma_{US}^2)T + \frac{T(1-\rho^2)\sigma_{US}^2}{2} + \sigma_{US}\rho\xi \right) \mathcal{N}(D)
\end{aligned}$$

$$\begin{aligned}
\tilde{x} &= \frac{x - (T(1-\rho^2)\sigma_{US} + \rho\xi)}{\sqrt{2T(1-\rho^2)}} \\
d\tilde{x} &= \frac{1}{\sqrt{T(1-\rho^2)}} dx
\end{aligned}$$

Final result:

$$\begin{aligned}
p &= \mathbb{E}_0[D(0, T)\{\text{Payoff}\}] \\
&= \mathbb{E}_0[D(0, T)\{(S_{US}(T) - S_{US}(0))^+ \cdot \mathbb{1}\{S_{EU}(T) < 0.95 \cdot S_{EU}(0)\}\}] \\
&= B(0, T)S_{US}(0) \int_{-\infty}^A \{ \exp \left(r_{US}T - \frac{1}{2} \sigma_{US}^2 \rho^2 T + \sigma_{US}\rho\xi \right) \mathcal{N}(D) - \mathcal{N}(B) \} \frac{1}{\sqrt{2\pi T}} \exp \left(-\frac{\xi^2}{2T} \right) d\xi
\end{aligned}$$

17. Appendix E: Python Code criticalities

Due to our low level of knowledge of python, we thought to implement a version of the code as simple as possible, excluding all graphs and model checks. In particular, we focused on the basic calibration (Unweighted rmse - NIG marginals) with the goal of being able to first of all produce a result and possibly as close as possible to what we found with MATLAB.

We succeeded with some critical issues: we were unable to obtain consistent results with MATLAB for ν_Z and for Black's sigmas and therefore decided to take the values from the MATLAB script and continue with the code processing.

The pricing results are consistent for Black and his semi-closed formula. Levy's pricing (done without decomposing the assets into the systemic and idiosyncratic components) differs slightly, but this could derive from differences in the languages and their executions.

18. References

- [1] Synthetic forwards and cost of funding in the equity derivative market, by Michele Azzone, Roberto Baviera (Politecnico di Milano), July 2021.
- [2] Multivariate asset models using Lévy processes and applications, by Laura Ballotta, Efrem Bonfiglioli (BCE), The European Journal of Finance, April 2014.
- Multivariate Time Changes for Lévy Asset Models: Characterization and Calibration, by E.Luciano, P.Semeraro
- The US Economy's Resilience Is Now Undeniable, by J.Levin (2024)
- Financial modelling for jump processes, by R.Cont, P.Tankov
- Variance-Gamma and Normal-Inverse Gaussian models: Goodness-of-fit to Chinese high-frequency index returns, by A.Göncü and H.Yang (2016).
- Degenerate distribution definition
- EU Yield Curve
- US Yield Curve
- Arbitrage Free Implied Volatility Surfaces, by M.Roper (University of Sydney), March 2010
- Implied Volatility Functions: Empirical Tests, by B.Dumas, J.Fleming, R.Whaley (National Bureau of Economic Research), March 1996