PORTFOLIO DIVERSIFICATION REVISITED

CHARLES SHAW

Abstract. We relax several assumptions in Alexeev and Tapon (2012) to account for non-normally distributed, skewed, multi-regime, and leptokurtic asset return distributions. By calibrating a Markov-modulated Lévy process model to equity market data, we demonstrate that our approach not only effectively captures the empirical moments, including higher-order moments, but also accommodates regime-switching dynamics and path variation. We further argue that much of the literature on portfolio diversification relies on oversimplified assumptions that conflict with observable regularities in financial time series. Ignoring these complexities, particularly in terms of non-linear and multi-regime behaviour, may lead to a material underestimation of risk.

Keywords: Markov-modulated Lévy process; regime-switching models; portfolio diversification

JEL Classification: *G01;G11;G15*

1. Introduction

An important issue in portfolio management is the development of strategies to mitigate risk. It is widely acknowledged that diversification is an effective method for reducing unsystematic (idiosyncratic) risk, provided that no insider information is available. Furthermore, idiosyncratic risk is well-established as a significant contributor to overall volatility. However, eliminating risk entirely is impossible, even with a broad range of investments. Specifically, systemic risk factors will always contribute to total risk, even in the presence of international diversification.

This remains a significant area of research among both investment professionals and academic circles. In simple terms, the objective is to invest an initial amount in financial assets to maximize the expected utility of terminal wealth. Markowitz [48] was among the first to address this issue by developing the mean-variance approach to portfolio optimisation, introducing quantitative methods for selecting optimal portfolios. Black, Scholes, and Merton [50, 20] extended these results, providing explicit formulations for portfolio asset selection and allocation in continuous time. While the techniques developed by Merton and colleagues led to substantial theoretical advancements, practical concerns remain. One key challenge stems from the implicit assumption that the value of a risky asset follows a geometric Brownian motion. Certain

Date: This version: September 9, 2024.

email: charles@fixedpoint.io.

observable patterns in financial time series, such as asymmetry, heavy tails in return distributions, and fluctuating conditional volatility, are not fully consistent with this hypothesis.

Another limitation of the Black-Scholes-Merton framework is its assumption of static coefficients. This becomes a critical consideration for longer-term investment horizons, where macroeconomic conditions may shift multiple times, fundamentally altering investment opportunity sets. In this context, Markov-modulated, or *regime-switching*, models seem well-suited to represent such phenomena. For example, changes in economic conditions can reasonably be quantified using a Markov chain, which adjusts the model's parameters. These models can describe macroeconomic shifts, periods when market behaviour shifts dramatically during crises, or the various stages of business cycles.

When optimizing portfolios, numerous factors must be considered in determining the number of assets required for optimal diversification. Systematic risk measurement can vary depending on the size of the investment universe, investor characteristics, asset attributes (which may change over time), the chosen diversification model, data frequency, market conditions, and the investment horizon. Calculating the optimal number of assets for a fully diversified portfolio — within a specific market, time frame, or set of preferences — poses challenges when approached from the traditional Mean-Variance Optimization framework introduced by Markowitz. However, recent research suggests that the size of a well-diversified portfolio is now larger than ever, that this number is smaller in developing economies compared to established financial markets, and that it decreases as stock correlations with the market increase [3, 4, 23].

The objective of this paper is to critically examine these and related topics. By employing a Markov-modulated L'evy process (where a L'evy process can be described as a suitable combination of two processes: diffusions and jumps), we relax several assumptions found in studies like Alexeev and Tapon [3] to account for non-normally distributed, skewed, multi-regime, and leptokurtic asset return distributions. We fit a model to equity market data, demonstrating that these calibrated models align well with empirical moments. We contend that much of the literature on portfolio diversification relies on assumptions that conflict with observable patterns, potentially leading to significant underestimations of risk when these patterns are ignored.

The structure of this paper is as follows: Section 2 reviews relevant literature, Section 3 presents a calibration exercise involving a Markov-modulated Lévy process applied to Nasdaq data, challenging several assumptions in Alexeev and Tapon [3] to accommodate non-normally distributed, skewed, multi-regime, and leptokurtic asset return distributions. Section 5 discusses the results, and Section 6 offers a conclusion.

2. Literature Review

Classic portfolio allocation theory is grounded in the mean-variance (MV) framework, which provides a method to analyse the trade-off between risk and return to achieve diversification benefits. Despite its limitations, the MV framework remains foundational in many asset allocation decisions, as numerous asset managers, consultants, and investment advisers utilise classical MV optimisation as a standard quantitative approach to portfolio construction. This framework was introduced by Markowitz [48], who, in 1952, developed Modern Portfolio Theory, laying the groundwork for subsequent developments in risk and return theories.

In addition to Markowitz, pioneering researchers such as Evans and Archer [32], Fielitz [33], Solnik [60], Statman [61], and Campbell et al. [23] have made significant contributions to the field of portfolio risk diversification. Their key findings can be summarised as follows:

- A portfolio's risk can be reduced through diversification, which mitigates both systematic and unsystematic risk.
- A portfolio's overall risk increases when the number of stocks it holds approaches that of the entire market.
- Unsystematic risk can be minimised up to the point of achieving optimal portfolio diversification, at which point the total portfolio risk equals systematic risk.

Let us first examine some of the key issues surrounding the measurement of risk diversification. Various methods exist for measuring portfolio risk, each with its own set of advantages and disadvantages, and these have been the subject of ongoing debate. Prominent among these methods is standard deviation, which is widely used in many studies as a commonly accepted metric for risk. This is evident from earlier research ([32, 60, 61, 16]) as well as more recent investigations ([22], [18]).

One major issue with using standard deviation as a risk metric is that it may produce erroneous and misleading results due to its sensitivity to extreme values and outliers. Specifically, it is well-known that standard deviation can lead to inaccurate estimates of extreme occurrences if returns are not normally distributed. Another limitation of the standard deviation approach is that it treats positive and negative deviations from the average return equally.

Given these challenges, a considerable amount of research has explored alternative risk measures, such as expected shortfall (ES) and terminal wealth standard deviation (TWSD), particularly in assessing the impact of financial crises on the optimal number of stocks in a portfolio ([22, 18]). Additionally, the mean absolute deviation (MAD) approach [33], which uses absolute deviation rather than variance, and the unsystematic risk ratio (URR) [59], which provides a measure of diversification relative to its variation, have been proposed as alternative measures of portfolio risk. It is important to note that the portfolio structures generated by different risk measures can vary significantly.

The heterogeneity of viewpoints on what constitutes successful diversification significantly complicates the task of comparing research across the literature on optimal portfolio diversification. For instance, Alexeev and Dungey [4] claim that investing in a diversified portfolio of seven or ten evenly-weighted equities may reduce risk by as much as 85 percent or 90 percent, respectively. However, Tang [64] argues that a portfolio of just 20 equities is needed to eliminate 95 percent of unsystematic risk, while a portfolio of 80 equities may reduce an additional 4 percent. Alekneviciene et al. [2] suggest that a differently-weighted portfolio could remove 97 percent of unsystematic risk with 25 stocks, whereas Stotz and Lu [62] found that in China, 67 percent of unsystematic risk could be mitigated with just 10 stocks. According to Kryzanowski and Singh

[45], investing in 20–25 stocks may minimise 90 percent of a portfolio's risk. They also find that risk-averse investors prefer portfolios with a 99 percent risk reduction, while more aggressive investors, who seek higher returns at the expense of more risk, may be satisfied with a 90 percent reduction.

Alexeev and Tapon [3] further argue that the required number of stocks in a well-diversified portfolio depends on the average correlations between stocks and the market, as well as the prevailing market conditions — whether distressed or stable — when analysing portfolio dynamics over multiple years. Raju and Agarwalla [56] suggest that the appropriate number of stocks is influenced by the investor's risk tolerance, desired level of confidence, and the portfolio's weighting structure. These findings align with intuition, as investors' actions typically vary according to their economic circumstances. Often, their decisions are dynamic and influenced by economic, cultural, and social factors, which may not always be entirely rational [46]. Consequently, their values, preferences, assumptions, and perceptions shift alongside the economy [8]. Moreover, an investor's location also affects both their optimal asset allocation and investment success [34].

It is also possible that the frequency of the data utilised may affect the optimal number of stocks in a portfolio. This could result in an exaggerated number of stocks, as demonstrated by Alexeev and Dungey [4]. They further note that this disparity is amplified during financial market crises. Early research based on (semi)annual and quarterly data suggests that the optimal number of equities for diversification ranges from 8 to 16 [32, 33, 67]. However, Alexeev and Dungey [4] argue that high-frequency data significantly enhances risk assessment and decision-making. With higher frequency data, the number of stocks required to achieve the desired risk reduction decreases.

Data from various frequency intervals shows a minor variation in unsystematic risk during stable periods, but a substantial difference during times of high volatility. Risk estimates based on lower frequency data tend to be overstated, particularly during financial crises. Higher frequency data provides more accurate risk measurements, which suggests that holding large portfolios is not always necessary, as implied by lower frequency risk measures, especially in times of financial distress. Price fluctuations of financial instruments are also influenced by fundamental factors such as interest rates, economic growth, and currencies. Investors may uncover diversification benefits by constructing portfolios based on these criteria [46]. Moreover, corporate bonds, in addition to equities, may offer significant risk reduction, especially during periods of financial turbulence.

Building on this reasoning, an aggregate stock index composed of the S&P 500's most prominent companies may reduce transaction costs while offering sufficient diversification, according to Aboura and Chevallier [1]. To prevent international financial contagion, investors and mutual fund managers must diversify their portfolios and utilise hedging strategies [35]. Funds with less liquid equities in their portfolios tend to be more diversified, and as noted by Pastor et al. [52], such funds are typically larger, cheaper, and trade more frequently. This represents a significant advancement in active portfolio management. Socially responsible investment (SRI) funds, according to Leite et al. [47], do not significantly affect idiosyncratic risk due to screening intensity, which is a crucial aspect of an investment strategy. Additionally, portfolio managers may opt to adjust their holdings during periods of political instability, favouring companies with more accurate reporting practices [25, 26].

There are several challenges in modelling financial time series, as observations can be influenced by unforeseen events. Events such as natural catastrophes, central bank statements, and government policy announcements may have a significant impact on the market. Consequently, the assumption of stationarity in financial data is often violated. As a result, traditional approaches to time series analysis may prove unsatisfactory. Markov-switching models are thus of interest, as they allow for the mitigation of issues related to suspected non-stationarity, under certain mild assumptions.

A key stylised feature of financial time series is regime switching, wherein economic conditions undergo periodic regime changes. The regular flow of economic activity may occasionally experience shocks significant enough to result in different observed dynamics. As a result, sampled time series data may exhibit not only periods of low and high volatility but also periods of slower and faster mean growth. The economy may oscillate between two states: (1) a stable, low-volatility state characterised by economic expansion, and (2) a panic-driven, high-volatility state defined by economic contraction. These dynamics are often described by deterministic or stochastic ordinary differential equations. Regime shifts indicate transitions between different dynamics, potentially involving changes in state space, the objective function, or both.

Regime changes can occur for various reasons: (i) exogenous changes in dynamics (e.g., due to sudden environmental disasters or social/political reform); (ii) unintentional internal changes in dynamics (e.g., due to human activity-related disasters or firm bankruptcy); (iii) intentional (controlled) shifts to new dynamics (e.g., technological innovations, mergers of firms); (iv) changes in preferences/objectives (e.g., environmental concerns). A single model may incorporate a combination of these triggers. Evidence of such regimes has been widely documented, and switching models have been applied across various domains, including but not limited to exchange rates, asset allocation, and equity markets. Surveys are provided by [5] and [38].

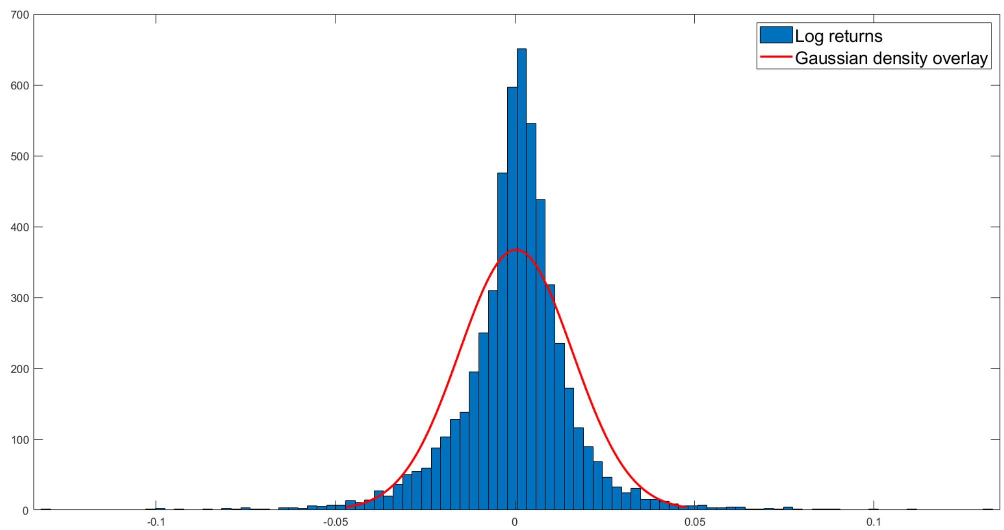
3. Data and model

To facilitate the discussion, we utilise daily closing price data for the Nasdaq 100 index, sourced from Oxford-Man's Realized Library[1]. The data spans from 04 January 2000 to 12 November 2021, resulting in 5,483 daily observations. Figure 1 presents the plot of log returns for the Nasdaq 100 (blue) with a normal distribution overlay (red). The graph clearly demonstrates that Gaussian density underestimates random variation. This can be attributed to the fact that financial returns typically exhibit distributions with certain stylised features, such as excess kurtosis (i.e., fat or semi-heavy tails), skewness (tail asymmetries), and regime shifts.

Practical modelling of the aforementioned stylised features can be achieved using the General Hyperbolic Distribution. The specific case of the General Hyperbolic Distribution that we focus on is the Normal-Inverse Gaussian (NIG) distribution, a continuous probability distribution defined as a normal variance-mean mixture with an inverse Gaussian mixing density [10, 15]. The NIG density is characterised by a four-dimensional parameter vector $[\alpha, \beta, \mu, \delta]$, which determines its form. Due to its extensive parametrisation, the NIG density is well-suited for modelling a wide range of unimodal, positively kurtotic data.

[1]https://realized.oxford-man.ox.ac.uk

Fig. 1. Log returns for Nasdaq 100 (blue) with normal distribution overlay (red)



The α parameter controls the steepness or peakedness of the density, increasing monotonically with larger values of α. A high α value indicates light tails, while a low α value corresponds to heavier tails. The β parameter governs the skewness of the distribution. When $\beta < 0$, the density is skewed to the left; when $\beta > 0$, it is skewed to the right; and when $\beta = 0$, the density is symmetric around the centrality parameter μ. The δ parameter is a scale parameter.

The class of NIG distributions is highly flexible, accommodating fat-tailed and skewed distributions, with the normal distribution, $N(\mu, \sigma^2)$, arising as a special case when $\beta = 0$, $\delta = \sigma^2/\alpha$, and $\alpha \to \infty$.

Consider the following generic Markov Regime Switching (MRS) model:

$$\begin{cases} y_t = f(S_t, \theta, \psi_{t-1}) \\ S_t = g\left(\tilde{S}_{t-1}, \psi_{t-1}\right) \\ S_t \in \Lambda \end{cases} \tag{3.1}$$

where θ is the vector of the parameters of the model, S_t is the state of the model at time t, $\psi_t := \{y_k : k = 1, \ldots, t\}$ is the set of all observations up to t, $\tilde{S}_t := \{S_1, ..., S_t\}$ is the set of all observed states up to t, $\Lambda = \{1, ..., M\}$ is the set of all possible states, and g is the function that regulates transitions between states. Function f indicates how observations at time t depend on $S_t, \theta,$ and ψ_{t-1} and finally, $t \in \{0, 1, ..., T\}$, where $T \in \mathbb{N}$, $T < +\infty$, is the terminal time.

Equations 3.1 enable us to address specific issues that may be difficult to capture in a single-state regime, which proves useful for time series applications. Although the literature on Markov-switching models is extensive, two general categories can be identified. The first category includes models with basic transition laws, such as a first-order Markov chain, but complex distributions for the data or a large number of states. For examples of research in this area, see [28, 37, 24]. The second category comprises models with more complex transition laws, but with simpler assumptions and a limited number of states, often restricted to two. See, for instance, [44, 29, 53].

We now focus on the estimation of a Markov-switching model augmented by jumps, represented as a Lévy process, to model financial returns. To motivate this section's modelling approach, we first outline the general structure of Lévy processes and then detail their properties, such as path variation and the Lévy-Khintchine theorem. The discussion highlights key properties of Markov chains, including irreducibility, aperiodicity, and ergodicity. This section also presents a framework for estimating a jump-robust model tempered by a Markov chain, which can be used to study the dependence relationships within financial returns.

3.1. **NIG-type distribution.** Following [27], we assume that the Lévy process L follows the Normal Inverse Gaussian (NIG) distribution, defined as a variance-mean mixture of a normal distribution with the inverse Gaussian as the mixing distribution (also see Barndorff-Nielsen et al. [10, 14, 11, 12] for a more extensive discussion).

The NIG distribution is a relatively novel process introduced by Barndorff-Nielsen [14] as a model for the log returns of stock prices. It is a subclass of the broader class of hyperbolic Lévy processes. After its introduction, the NIG distribution was shown to provide an excellent fit to stock market log returns [13]. Other studies have demonstrated the superior empirical fit of this distribution for various asset classes [54, 41, 31, 36]. [63] find that the Normal Inverse Gaussian distribution provides a better overall fit for financial data than other subclasses of Generalised Hyperbolic distributions and significantly outperforms Lévy-stable laws. Rachev et al. [55] have successfully applied the NIG distribution to address well-known 'puzzles,' such as (i) the predictability of asset returns, (ii) the equity premium, and (iii) the volatility puzzle.

This type of heavy-tailed processes is of interest, particularly since the NIG distribution fulfills the fat-tails condition, is analytically tractable, yet is closed under convolution [40].

The density function of a $NIG(\alpha,\beta,\delta,\mu)$ is given by

$$f_{NIG}(x;\alpha,\beta,\delta,\mu) = \frac{\alpha}{\pi} e^{\delta\sqrt{\alpha^2-\beta^2+\beta(x-\mu)}} \frac{K_1(\alpha\delta\sqrt{1+(x-\mu)^2/\delta^2})}{\sqrt{1+(x-\mu)^2/\delta^2}}, \tag{3.2}$$

where $\delta > 0$, $\alpha \geq 0$. The parameters in the Normal Inverse Gaussian distribution can be interpreted as follows: α is the tail heaviness of steepness, β is the skewness, δ is the scale, and μ is the location. The NIG distribution is the only member of the family of general hyperbolic distributions to be closed under convolution. K_v is the Hankel function with index v. This can be represented by

$$K_v(z) = \frac{1}{2}\int_0^\infty y^{v-1} e^{\left(-\frac{1}{2}z\left(y+\frac{1}{y}\right)\right)} dy \tag{3.3}$$

For a given real v, the function K_v satisfies the differential equation given by

$$v^2 y'' + xy' - (x^2+v^2)y = 0. \tag{3.4}$$

The log cumulative function of a Normal Inverse Gaussian variable is given by

$$\phi^{NIG}(z) = \mu z + \delta \left(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + z)^2}\right) \text{ for all } |\beta + z| < \alpha. \tag{3.5}$$

The first two moments are $\mathbb{E}[X] = \mu + \frac{\delta\beta}{\gamma}$, and $Var[X] = \frac{\delta\alpha^2}{\gamma^3}$, where $\gamma = \sqrt{\alpha^2 - \beta^2}$. The Lévy measure of a $NIG(\alpha, \beta, \delta, \mu)$ law is

$$F_{NIG}(dx) = e^{\beta x} \frac{\delta\alpha}{\pi|x|} K_1(\alpha|x|)dx. \tag{3.6}$$

We then follow the two-stage estimation strategy proposed by [27]. Figures 2 and 3 point to the presence of identifiable regimes.

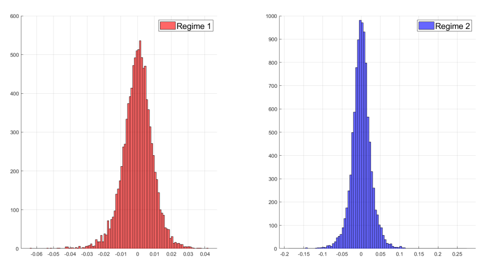


Fig. 2. Nasdaq regimes, separated.

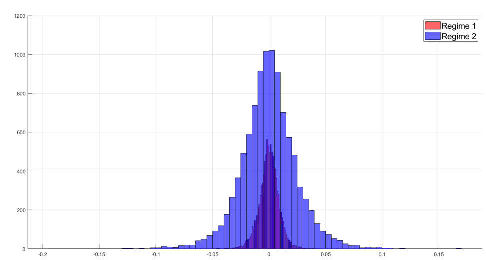


Fig. 3. Regimes, superimposed.

4. Results

Figure 4 shows closing prices for the NASDAQ (top), together with the volatility of log-returns (middle), and the evolution of regimes in our MRS model. Tables 1 and 2 report the NIG distribution parameters of the Lévy jump process fitted to each regime.

Table 1. NIG distribution parameters of the Lévy jump process fitted to each regime.

NIG	α	β	δ	μ
State 1	150.0919	-16.2944	0.011949	0.001276
State 2	41.8416	0.295358	0.026838	-0.00015

In Table 1, the parameter α represents the jump intensity. A higher value of α corresponds to lower jump intensity in a given regime. We observe that $\alpha \approx 150$ during regime 1, indicating a low jump intensity, while $\alpha \approx 42$ during regime 2 points to a high jump intensity. The parameter δ represents the scale, providing a measure of the dispersion of returns. In regime 2, $\delta = 0.027$, compared to $\delta = 0.12$ in regime 1, indicating that the returns are approximately 2.5 times more dispersed in the second regime than in the first.

Table 2 presents the diffusion parameters of the Lévy jump process fitted to each regime. The mean-reverting parameter, κ, is close to unity in both regimes. The volatility parameter, σ, is more than seven times higher during regime 2 than during regime 1. Therefore, the substantially higher jump intensity

Fig. 4. NASDAQ close prices, log returns, and regimes.

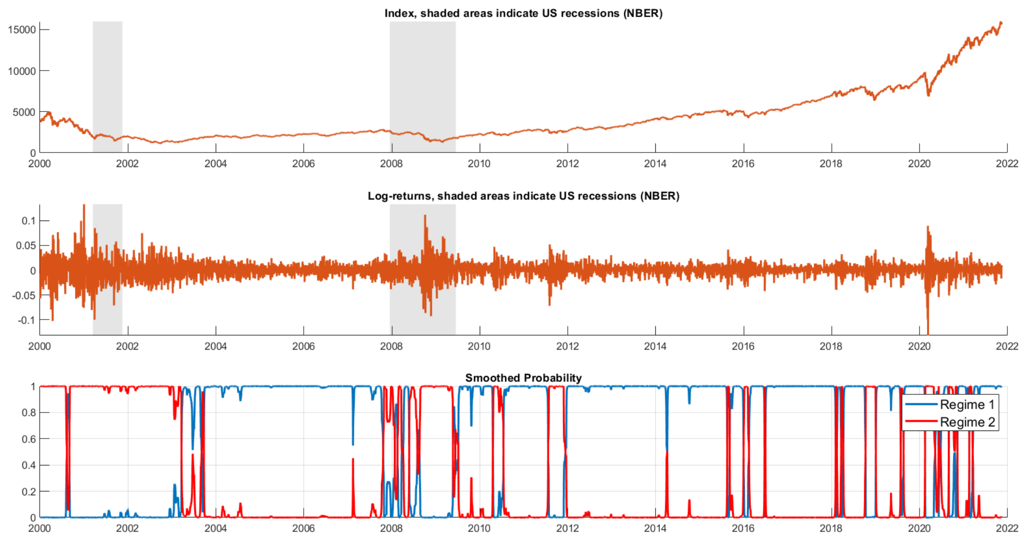


Table 2. NIG diffusion parameters.

NIG	κ	θ	σ	P^Z_{ii}
State 1	1.011003	0.000994	0.000082	0.716268
State 2	1.082196	-0.00154	0.000643	0.283732

observed in regime 2 translates into increased volatility. The smoothed probabilities in the lower panel of Figure 4 illustrate the probability of remaining in the current regime, which is high for regime 1. This is quantified in the last column of Table 2: the probability of staying in regime 1 (in regime 2) is 71.6% (28.3%). If the chain exhibits general persistence (e.g., a high likelihood of remaining in a particular regime), this could have significant implications for computing Value-at-Risk and dynamic portfolio allocation.

The graphs and numerical results primarily demonstrate that the stochastic process governing each regime does not exhibit the same jump characteristics across the sample period. Certain periods in the index have a high frequency of jumps (captured under regime 2), while others do not. These findings highlight the merits of using the regime-switching Lévy model to simulate price dynamics. In the following section, we examine the quality of regime classification and model fit.

4.0.1. *Regime Classification Measure of Ang and Bekaert (2002).* A robust model should sharply classify the regimes, with smoothed probabilities close to either ≈ 0 or ≈ 1. To evaluate this directly, Regime Classification Measures (RCMs) were proposed by Ang and Bekaert [7] as a method for determining whether the number of regimes, K, is appropriate. The RCM statistic ranges from 0 (indicating perfect regime classification) to 100 (indicating a failure to detect any regime classification) and can be formalised as follows:

$$RCM(K) = 100 \times \Big(1 - \frac{K}{K-1}\frac{1}{T}\sum_{k=1}^{N}\sum_{Z_{t_k}}^{N}\Big(P(Z_{t_k} = i|\mathcal{F}^X_{t_M}; \hat{\Theta}^{(n)}_1) - \frac{1}{K}\Big)^2\Big), \tag{4.7}$$

where $P(Z_{t_k} = i|\mathcal{F}^X_{t_M}; \hat{\Theta}^{(n)}_1)$ corresponds to the smoothed probability and $\hat{\Theta}^{(n)}_1$ is the vector of estimated parameters. $RCM \in [0, 100]$ and lower values are preferred to higher ones. In this sense, a 'perfect' model will be associated with an RCM of almost 0, a good model will have an RCM close to ≈ 0, while a model that cannot distinguish between regimes at all will have an RCM close to 100. A good model implies that the smoothed probability is less than 0.1 or greater than 0.9. This means that the data at time $t \in [0, T]$ is in one of the regimes at the 10% error level. The RCM was extended for multiple states by Baele [9].

4.0.2. *Smoothed probability indicator.* The quality of classification may also be observed when the smoothed probability is less than p or greater than $1-p$ with $p \in [0, 1]$. Thus the data at time $k \in 1, \dots, N$ has a probability higher than $(100-2p)\%$ in one of the regimes for the $2p\%$ error. This percentage is the smoothed probability indicator with $p\%$ error, denoted in Table 3 by $P^{\%}$.

Table 3. Regime Classification Measure (RCM) and Smoothed Probability Indicator

Nasdaq	RCM	p^{10}
	8.63	91.39

4.0.3. *Measuring the quality of our model's regime classification performance.* A natural question to ask is how to measure the quality of our model's regime classification performance. In Table 3, we observe that the RCM statistic for our model is less than 10, indicating a good fit for the Markov Regime-Switching (MRS) model. This serves as evidence that our model, with two regimes, provides a satisfactory fit to the data. In other words, given the set of models considered in our analysis, and using a sufficiently long period (n=5,483 observations), we conclude that the Lévy regime-switching model effectively identifies two regimes for the Nasdaq index returns within our sample.

4.0.4. *Value-at-Risk.* It is instructive to compare the Value-at-Risk (VaR) estimates produced by the standard Gaussian model with those of the Normal Inverse Gaussian (NIG) model. Over longer time horizons, the estimates from both models converge. If both the NIG and Gaussian models are appropriately fitted to the same dataset, the resulting VaR estimates for extended periods will be nearly identical. However, for shorter time horizons, the divergence between the two models' VaR estimates becomes pronounced, highlighting the limitations of using a simple, universal risk metric. As discussed in [21], VaR is an inconsistent measure, which complicates its application in identifying an optimal portfolio to minimise risk.

4.1. **Construction of Random Portfolios.** Following Alexeev and Tapon [3], we use a simulation approach to construct random portfolios based on actual daily equity returns. Our sample spans from 04 January 2000 to 12 November 2021. To depart from the classical Gaussian limitations of [3], we generate random values from the Markov-modulated Normal-Inverse Gaussian (NIG) distribution. The model parameters are specified in Tables 1 and 2.

It is important to recognise the relationship between security returns and their associated uncertainty. However, diversification is achievable due to the low or negative correlations between assets, which complicates the analysis. While examining the variances and correlations is straightforward for small portfolios, such as five stocks, this approach becomes impractical as the number of assets increases. For instance, a portfolio of 200 stocks would result in 200 variances and 19,900 correlations or covariances, making direct examination ineffective. A practical solution to this complexity is to apply *dimension reduction*. Specifically, we utilise principal component analysis (PCA), a well-established statistical technique, to simplify the analysis by focusing on a subset of the variances and correlations or covariances that are assumed to be uncorrelated.

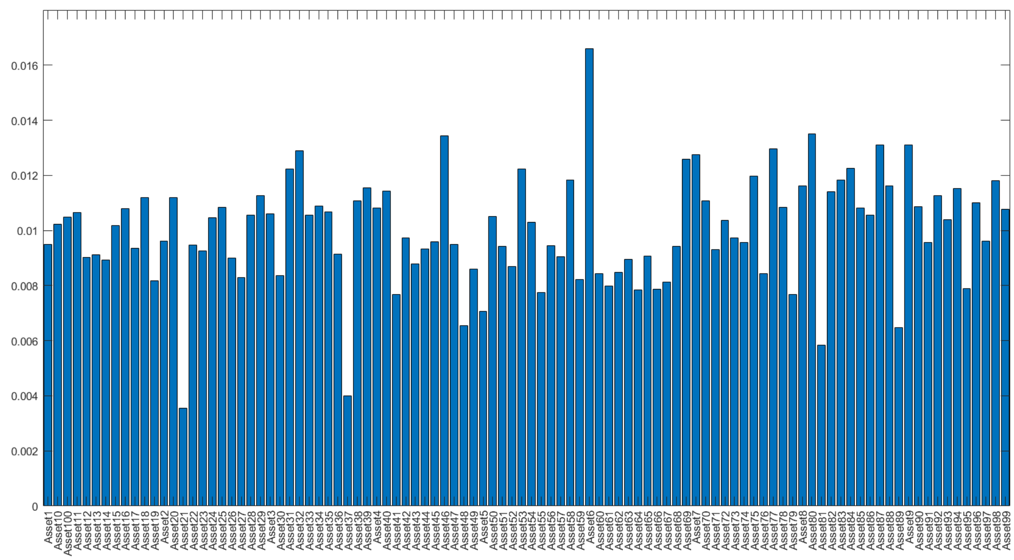
For our illustrative example, we adopt the methodology suggested in [49] (p. 131). A factor model is used to optimise asset allocation within a mean-variance framework. Multifactor models are commonly employed for risk modelling, portfolio management, and performance attribution. These models reduce the dimensionality of the investment universe while explaining a significant portion of market variability. Factors may be statistical, macroeconomic, or fundamental in nature. In this case, we generate statistical factors from asset returns and optimise our allocation according to these factors. Figure 5 shows the asset weights derived from our factor model.

Our factor model optimisation reveals that 94 components are required to explain at least 95% of the total variance, while 87 components explain 90%. These values were derived using a Principal Component Analysis (PCA) approach, applied to the returns of 100 assets over 5,483 daily observations. PCA is a widely-used statistical technique for dimensionality reduction, where the original asset return space is transformed into a set of orthogonal factors (principal components), each of which captures a portion of the overall variance in returns. In this context, each component represents a linear combination of the original asset returns, weighted in a way that maximises the variance explained by that component.

In this analysis, the first principal components typically explain the majority of the variance, and as more components are added, the incremental variance explained diminishes. This is consistent with the concept of diminishing marginal utility in portfolio diversification, where adding additional assets results in increasingly smaller risk reduction benefits. The cumulative variance explained by successive components is shown in Figure 5, where the portfolio weights corresponding to each component have been optimised within a mean-variance framework.

The decision to focus on the 95% and 90% thresholds for explained variance is a common heuristic in financial modelling and portfolio optimisation. These thresholds represent a trade-off between capturing enough of the variance in returns (to ensure effective diversification) and maintaining a manageable portfolio size. The corresponding portfolio sizes (94 and 87 components) provide a measure of the dimensionality

Fig. 5. Weights of our constructed portfolio (5483 observations, 100 assets) based on parameters recovered from market data based on distributional and functional form assumptions.



required to sufficiently capture the relevant market dynamics, including the higher-order statistical properties of the return distributions.

Unlike simpler models that assume normally distributed returns, our approach uses a Markov-modulated Lévy process to account for the non-normal, skewed, multi-regime, and leptokurtic nature of asset returns. Such distributions exhibit heavy tails and regime shifts, which introduce additional risks not captured by traditional Gaussian-based models. Incorporating these complexities into the factor model increases the number of components needed to explain the total variance because these factors attempt to capture both linear and non-linear relationships between assets over time.

5. Discussion

5.1. **Dealing with Regimes and Higher Moments.** Several factors must be considered when determining how many assets are necessary to achieve optimal diversification. These include the investment universe (size, asset classes, and features of the asset classes), the investor's characteristics, changes in asset features over time, the model adopted to measure diversification (e.g., equally weighted versus an "optimal" allocation), and the frequency of the data used. We argue that it is equally important to account for well-documented nuances in financial time series data, particularly the presence of non-normal distributions, skewness, multi-regime behaviour, leptokurtic characteristics, and non-linear or discontinuous price processes.

Additionally, the cost of overseas investment must be considered when evaluating the benefits of international diversification. Investors face several challenges when constructing portfolios outside their home market, including currency risks, political instability, limitations on money transfers, and differing regulatory environments across countries. In contrast, domestic investments are free from such complications, making them easier to manage from a diversification, security analysis, and asset allocation perspective.

There is a range of opinions in the academic literature on what constitutes successful diversification. Recent research—see, for example, [3], [56], [4], and the literature cited therein—suggests that the average size of a well-diversified portfolio has increased compared to earlier studies, which may be attributed in part to reduced transactional costs. This research also indicates that the ideal number of equities in a well-diversified portfolio has risen significantly when longer time horizons are considered. According to [66], U.S. unsystematic risk has grown markedly in recent decades relative to overall stock market volatility, underscoring the need for larger portfolios to reduce diversifiable risk as much as possible.

However, not all investors may achieve the same level of diversification by following uniform portfolio size guidelines. In making these decisions, it is important to consider several factors, including the frequency of data collection, the level of risk accounted for, the confidence of investors, and the potential diversification benefits of the chosen investment opportunity set, among others.

5.2. **Simple Heuristic Diversification Rules.** Our results also highlight the so-called 1/n investment puzzle, where some market participants allocate their contributions equally across all assets. Since this approach contradicts the core principles of modern portfolio theory, it is often regarded as a naive method. There is evidence [17] suggesting that many participants in defined contribution plans use basic heuristic diversification methods to distribute their contributions across different asset classes. The 1/n rule, commonly referred to as an "equally-weighted portfolio," is one such heuristic. However, it has been criticised for not lying on the efficient frontier, as it does not represent an optimal portfolio. Some academics argue that pension systems should be less flexible to prevent individuals from making suboptimal investment decisions.

Nevertheless, [65] presents arguments demonstrating that this behaviour may be less naive than it initially seems, showing that the 1/n rule offers certain advantages in terms of robustness. While we do not advocate for any heuristic diversification rule in this paper, we note that further work may be required to improve upon this simple diversification criterion. In this context, [39] observes that "naive formation rules such as the equal weight rule can outperform the Markowitz rule." Similarly, [51] notes that due to estimation risk, "an equally weighted portfolio may often be substantially closer to true mean-variance optimality than an optimised portfolio."

5.3. **Attempts at a Unified Theory of Portfolio Management.** Before concluding our discussion, we should briefly address the feasibility of applying the Mean-Variance Optimisation (MVO) framework to assets beyond equities, such as fixed income. Any attempt at diversification naturally involves diversifying across asset classes. However, when constructing bond portfolios, the established MVO approach poses certain challenges, making the development of a unified theory for bond portfolios non-trivial.

To elaborate, stocks and bonds differ in several key ways, the most significant being the fixed maturity date at which bonds expire and disappear from the market, whereas stock characteristics can fluctuate in response to business news or management decisions. In an unconstrained market, the maturity period of bonds can take on an infinite number of values, leading to an infinite number of distinct bonds. Consequently, the price of a stock is determined solely by its associated risks (e.g., market risk, idiosyncratic risk), while the price of a bond is influenced by both its risks and the time remaining until maturity.

The stochastic differential equations (SDEs) used to model stock prices are typically autonomous, meaning that the coefficients are time-independent functions of the prices, as seen in geometric Brownian motion or mean-reverting processes. However, any model for bond prices must account for the fact that volatility approaches zero as the time to maturity decreases—at least, this is how it is mathematically expressed.

A simple portfolio of stocks and bonds is made more analytically challenging by the fact that prices for each asset class follow different principles, even in the most basic scenario. Furthermore, due to the dependency on maturity, many strategies permissible in the stock market are not directly applicable in the bond market. For example, a simple buy-and-hold strategy results in bonds converting into cash upon maturity. In the context of fixed-income derivatives, it is natural to model interest rates with an uncountably large set of traded instruments.

While some notable attempts toward a unified theory of portfolio management incorporating both stocks and bonds have been made, the literature remains relatively sparse, and the problem of multi-asset portfolio optimisation is still an open research question. In this vein, Ekeland and Taffin [30] made strides toward introducing a bond portfolio management theory, building on foundations similar to stock portfolio management by modelling the discounted price curve as infinite-dimensional dynamics in a Banach space of continuous functions, driven by a cylindrical Wiener process. The interested reader is referred to [57], who extend these results using the HJM framework, and [19], who apply them to commodity futures markets. Extending these results presents a promising avenue for future research.

6. Conclusion

This paper discusses some of the issues which arise in the portfolio allocation literature. However, we wish to focus the reader's attention on challenges of a different kind, which seem to be absent from the discussions in the optimum portfolio diversification literature. Specifically, there is merit in exploring certain stylized features of financial time series in an attempt to capture dynamics of non-normally distributed, skewed, multi-regime, and leptokurtic asset return distributions.

We promoted the estimation of a Markov-switching model augmented by jumps, under the form of a Lévy process. After setting up the general structure of Lévy processes, we outlined their properties concerning path variation and the Lévy-Khintchine theorem. The analysis highlighted the properties of Markov chains, such as irreducibility, aperiodicity, and ergodicity.

By deploying a Markov-modulated Lévy process, we relaxed several assumptions in Alexeev and Tapon [3] to account for account for non-normally distributed, skewed, multi-regime, and leptokurtic asset return distributions. We calibrated a regime-switching Lévy model to equity market data to demonstrate that such a model is a) analytically tractable and computationally effective, b) intuitive, and c) does a good job of matching the empirical moments, including those of higher order. Finally, we argue that a part of the related literature on portfolio diversification relies on assumptions that are in tension with certain observable regularities and which, if ignored, may lead to material underestimation of risk.

Code availability: Matlab code is available at `https://github.com/shawcharles/Portfolio-Diversification-Revisit`

Appendix A. Lévy processes

Lévy processes can be thought of as a combination of two distinct processes, namely diffusions and jumps. The attractive properties of such a combination can demonstrated by sketching the connections between the two. A well-known pure diffusion process used in finance is the Wiener process, a continuous-time Markovian stochastic process with a.s. continuous sample paths. A well-known pure jump process is the Poisson process, which is a non-decreasing process that, unlike Wiener, does not have continuous paths. Whilst the Poisson process has paths of bounded variation over finite time horizons, the paths of a Wiener process exhibit unbounded variation over finite time horizons.

When combined, these become interesting and, crucially, tractable tools for modelling financial time series due to their ability to match the empirically observed behaviour of financial markets more accurately than when armed with simple Wiener process-based models. These tools are useful, for example, in modelling jumps, spikes, and other such discontinuous variations in the price signal that are frequently observed in asset price processes. Such jump dynamics may be due to short-term liquidity challenges, microstructure frictions, or news shocks. Despite their apparent differences, these two processes have much in common. Both processes are initiated from the origin, both have right-continuous paths with left limits[2], and both have independent and stationary increments. Hence, these common features can be generalised to define a common framework of one-dimensional stochastic (Lévy) processes.

A.1. **A regime-switching Lévy model.**

This subsection motivates the introduction of the regime-switching Lévy approach to the modelling of our time series. This part of the discussion follows the methodology proposed by [27], namely by combining a Lévy jump-diffusion model with a Markov-switching framework. The regime-switching Lévy model offers the possibility of identification of such stochastic jumps, together with disentangling different market regimes and capturing the regime-switching dynamics. We begin by introducing a set of key definitions and notations.

Definition A.1 (Stochastic Process)**.** *A stochastic process X on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a collection of random variables $(X_t)_{0\leq t<\infty}$.*

If $X_t \in \mathcal{F}_t$, the process X is adapted to the filtration $\mathcal{F}$, or equivalently, $\mathcal{F}_t$-measurable.

Definition A.2 (Brownian Motion)**.** *Standard Brownian motion $W = (W_t)_{0\leq t<\infty}$ has the following three properties:*
(i) $W_0 = 0$
(ii) W has independent increments: $W_t - W_s$ is independent of $\mathcal{F}_s$, $0 \leq s < t < \infty$
(iii) $W_t - W_s$ is a Gaussian random variable: $W_t - W_s \sim N(0, t-s)\ \forall\ 0 \leq s < t < \infty$

[2]We adopt the convention that all Lévy processes have sample paths that are cadlag or RCLL i.e. right-continuous with left limits at every t.

Property (ii) implies the Markov property i.e. conditional probability distribution of future states of the process depends only on the present state. Property (iii) indicates that knowing the distribution of W_t for $t \leq \tau$ provides no predictive information about the state of the process when $t > \tau$. We can also define a Poisson Process, another stochastic mechanism as follows.

Definition A.3 (Poisson Process)**.** *A Poisson process $N = (N_t)_{0\leq t<\infty}$ satisfies the following three properties:*
(i) $N_0 = 0$
(ii) N has independent increments: $N_t - N_s$ is independent of $\mathcal{F}_s$, $0 \leq s < t < \infty$
(iii) N has stationary increments: $P(N_t - N_s \leq x) = P(N_{t-s} \leq x)$, $0 \leq s < t < \infty$

SDEs formulated with only the Poisson process or Brownian motion may not be very useful in investing or risk management. Arguably one needs more realistic models to describe the complex dynamics of an evolving system. However, their common properties may be combined, thus establishing a more general process.

Definition A.4 (Lévy Process)**.** *Let L be a stochastic process. Then L_t is a Lévy process if the following conditions are satisfied:*
(i) $L_0 = 0$
(ii) L has independent increments: $L_t - L_s$ is independent of $\mathcal{F}_s$, $0 \leq s < t < \infty$
(iii) L has stationary increments: $\mathbb{P}(L_t - L_s \leq x) = \mathbb{P}(L_{t-s} \leq x)$, $0 \leq s < t < \infty$
(iii) L_t is continuous in probability: $\lim_{t\to s} L_t = L_s$

Condition (iii) follows from (i) and (ii). For proof see [42].

Definition A.5. *A real-valued random variable Θ has an infinitely divisible distribution if for each $n = 1, 2, \ldots$, there exists a i.i.d. sequence of random variables $\Theta_1, \ldots, \Theta_n$ such that*

$$\Theta \stackrel{d}{=} \Theta_{1,n} + \ldots + \Theta_{n,n}$$

This says that the law μ of a real-valued random variable is infinitely divisible if for each $n = 1, 2, \ldots$ there exists another law μ_n of a real-valued random variable such that $\mu = \mu_n^{*n}$, the n-fold convolution of μ_n.

The full extent to which we may characterize infinitely divisible distributions is carried out via their characteristic function (or Fourier transform of their law) and the Lévy-Khintchine formula.

Theorem A.6 (Lévy-Khintchine formula)**.** *Suppose that $\mu \in \mathbb{R}$, $\sigma \geq 0$, and Π is a measure concentrated on $\mathbb{R}/\{0\}$ such that $\int_{\mathbb{R}} \min(1, x^2)\Pi(dx) < \infty$. A probability law μ of a real-valued random variable L has characteristic exponent $\Psi(u) := -\frac{1}{t}\log \mathbb{E}[e^{iuL_t}]$ given by,*

$$\Phi(u;t) = \int_{\mathbb{R}} e^{iux}\mu(dx) = e^{-t\Psi(u)} \quad for \quad u \in \mathbb{R}, \tag{A.8}$$

if (and only if) there exists a triple (γ, σ, Π), where $\gamma \in \mathbb{R}, \sigma \geq 0$ and Π is a measure supported on $\mathbb{R} \setminus \{0\}$ satisfying $\int_{\mathbb{R}}(1 \wedge x^2)\Pi(dx) < \infty$, such that

$$\Psi(\lambda) = i\gamma u + \frac{\sigma^2 u^2}{2} + \int_{\mathbb{R}} \left(1 - e^{(iux)} + iux\mathbf{1}_{|x|<1}\right) \Pi(dx) \tag{A.9}$$

for all $u \in \mathbb{R}$.

From Theorem A.6 we can say that there exists a probability space where $L = L^{(1)} + L^{(2)} + L^{(3)}$. We can build intuition as follows: $L^{(1)}$ is standard Brownian motion with drift, $L^{(2)}$ is a compound Poisson process, and $L^{(3)}$ is a square-integrable martingale with countable number of jumps of magnitude less than unity (a.s.). This is the Lévy-Itô decomposition, which can be stated as follows

$$L_t = \eta t + \sigma W_t + \int_0^t \int_{|x|\geq 1} x\mu^L(ds, dx) + \int_0^t \int_{|x|<1} x(\eta^L - \Pi^L)(ds, dx). \tag{A.10}$$

Definition A.7 (Markov-Switching). *Let $(Z_t)_{t\in[0,T]}$ be a continuous time Markov chain on finite space $S := \{1, \ldots, K\}$. Let $\mathcal{F}_t^Z := \{\sigma(Z_s); 0 \leq s \leq t\}$ be the natural filtration generated by the continuous time Markov chain Z. The generator matrix of Z, denoted by Π^Z, is given by*

$$\Pi_{ij}^Z \begin{cases} \geq 0, & \text{if } i \neq j \\ -\sum_{j\neq i} \Pi_{ij}^Z, & \text{otherwise} \end{cases} \tag{A.11}$$

We can now define the Regime-switching Lévy model as follows.

Definition A.8 (Regime-switching Lévy model). *For all $t \in [0, T]$, let Z_t be a continuous time Markov chain on finite space $S := \{1, \ldots, K\}$ defined as per Definition A.7. A regime-switching model is a stochastic process (X_t) which is a solution of the stochastic differential equation given by*

$$dX_t = k(Z_t)(\theta(Z_t) - X_t)dt + \sigma(Z_t)dY_t \tag{A.12}$$

where $k(Z_t)$, $\theta(Z_t)$, and $\sigma(Z_t)$ are functions of the Markov chain Z. They are scalars which take values in $k(Z_t)$, $\theta(Z_t)$, and:

$$\begin{aligned} &\sigma(Z_t) : k(Z_t) := \{k(1), \ldots, k(K)\} \in \mathbb{R}^{K^*}, \\ &\theta(S) := \{\theta(1), \ldots, \theta(K)\}, \\ &\sigma(S) := \{\sigma(1), \ldots, \sigma(K)\} \in \mathbb{R}^{K^*}, \end{aligned}$$

where Y is a Wiener or a Lévy process. Here, k denotes the mean reverting rate, θ denotes the long-run mean, and σ denotes the volatility of X.

The above model exhibits two sources of randomness: the Markov chain Z, and the stochastic process Y which appears in the dynamics of X. In other words, there is stochasticity due to the Markov chain Z, $\mathcal{F}^Z$, and stochasticity due to the market information which is the initial continuous filtration $\mathcal{F}$ generated by the stochastic process Y.

Appendix B. Technical supplement

The following arguments follow directly from [27].

B.1. **Stage 1: The regime-switching model.** We aim to estimate the set of parameters $\Theta = \hat{\Theta}_1 := (\hat{k}_i, \hat{\theta}_i, \hat{\sigma}_i, \hat{\Pi}_i)$ for $i \in S$.

(1) Start with initial vector $\hat{\Theta}_1^{(0)} := \left(\hat{k}_i^{(0)}, \hat{\theta}_i^{(0)}, \hat{\sigma}_i^{(0)}, \hat{\Pi}_i^{(0)}\right)$ for $i \in S$. Let $N \in \mathbb{N}$ be the maximum number of iterations. Fix a positive constant ϵ as a convergence constant for the estimated log-likelihood function.
(2) Assume that we are at the $n+1 \leq N$ steps. The calculation in the previous iteration of the algorithm gives the following vector set $\hat{\Theta}_1^{(n)} := \left(\hat{k}_i^{(n)}, \hat{\theta}_i^{(n)}, \hat{\sigma}_i^{(n)}, \hat{\Pi}_i^{(n)}\right)$

B.2. **EM-algorithm.**

B.2.1. *Expectation step (E step).* We aim to estimate both filtered probability and smoothed probability. Optimality is achieved when a model can identify regimes sharply, such that smoothed probabilities approach either zero or one. Filtered probability is given by the probability such that the Markov chain Z is in regime $i \in S$ at time t with respect to $\mathcal{F}_T^X$:

For all $i \in S$ and $k = \{1, 2, \dots, M\}$, estimate the following

$$
\begin{aligned}
P\Big(Z_{t_k} = i | \mathcal{F}_{t_k}^X; \hat{\Theta}_1^{(n)}\Big) &= \frac{P\Big(Z_{t_k}, X_{t_k} | \mathcal{F}_{t_{k-1}}^X; \hat{\Theta}_1^{(n)}\Big)}{f\big(X_{t_k} | \mathcal{F}_{t_{k-1}}^X; \hat{\Theta}_1^{(n)}\big)} \\
&= \frac{P\Big(Z_{t_k} = i | \mathcal{F}_{t_{k-1}}^X; \hat{\Theta}_1^{(n)}\Big) f\Big(X_{t_k} | Z_{t_k} = i; \mathcal{F}_{t_{k-1}}; \hat{\Theta}_1^{(n)}\Big)}{\sum_{j \in S} P\Big(Z_{t_k} = j | \mathcal{F}_{t_{k-1}}^X; \hat{\Theta}_1^{(n)}\Big) f\Big(X_{t_k} | Z_{t_k} = j; \mathcal{F}_{t_{k-1}}; \hat{\Theta}_1^{(n)}\Big)}
\end{aligned} \tag{B.13}
$$

such that

$$
\begin{aligned}
P\Big(Z_{t_k} = i | \mathcal{F}_{t_{k-1}}^X; \hat{\Theta}_1^{(n)}\Big) &= \sum_{j \in S} P\Big(Z_{t_k} = i, Z_{t_{k-1}} = j | \mathcal{F}_{t_{k-1}}^X; \hat{\Theta}_1^{(n)}\Big) \\
&= \sum_{j \in S} P\Big(Z_{t_k} = i, Z_{t_{k-1}} = j | \hat{\Theta}_1^{(n)}\Big) P\Big(Z_{t_{k-1}} = j | \mathcal{F}_{t_{k-1}}^X; \hat{\Theta}_1^{(n)}\Big) \\
&= \sum_{j \in S} \Pi_{ij}^{(n)} P\Big(Z_{t_{k-1}} = j | \mathcal{F}_{t_{k-1}}^X; \hat{\Theta}_1^{(n)}\Big),
\end{aligned} \tag{B.14}
$$

where $f\Big(X_{t_k} | Z_{t_k} = i; \mathcal{F}_{t_{k-1}}; \hat{\Theta}_1^{(n)}\Big)$ is the density of the process X at time t_k, conditional that the process is in regime $i \in S$. Using previous arguments we can observe that, given $\mathcal{F}_{t_{k-1}}^X$, the process X_{t_k} has a conditional Gaussian distribution $\sim N\Big(k_i^{(n)}\theta_i^{(n)} + (1-k_i^{(n)})X_{t_{k-1}}, \sigma_i^{2(n)}\Big)$. The density of this distribution is given by

$$f\Big(X_{t_k}|Z_{t_k}=i;\mathcal{F}_{t_{k-1}};\hat{\Theta}_1^{(n)}\Big) = \frac{1}{\sqrt{2\pi\sigma_i^{(n)}}}\exp\Big[\frac{X_{t_k}-\big(1-k_i^{(n)})X_{t_{k-1}}-\theta_i^{(n)}k_i^{(n)}\big)^2}{2\big(\sigma_i^{(n)}\big)^2}\Big] \tag{B.15}$$

On the other hand, to estimate smoothed probability we need to examine when Markov chain Z is in regime $i \in S$ at time t with respect to all the historical data $\mathcal{F}_T^X$. For all $i \in S$ and $k = \{M-1, M-2, \ldots, 1\}$ we obtain

$$P\Big(Z_{t_k}=i|\mathcal{F}_{t_M}^X;\hat{\Theta}_1^{(n)}\Big) = \sum_{j\in S}\Big(\frac{P\big(Z_{t_k}=i|\mathcal{F}_{t_k}^X;\hat{\Theta}_1^{(n)}\big)P\big(Z_{t_{k+1}}=j|\mathcal{F}_{t_M};\hat{\Theta}_1^{(n)}|\Pi_{ij}^{(n)}\big)}{P\big(Z_{t_{k+1}}=j|\mathcal{F}_{t_k}^X;\hat{\Theta}_1^{(n)}\big)}\Big) \tag{B.16}$$

B.2.2. *Maximization step (M step).* We can obtain an explicit formula of the maximum likelihood estimator of the initial subset of parameters $\hat{\Theta}_1$. The maximum likelihood estimates $\hat{\Theta}_1^{(n+1)}$ for all parameters, for all $i \in S$, can be obtained by

$$\begin{aligned}
\theta_i^{(n+1)} &= \frac{\sum_{k=2}^M\big[P\big(Z_{t_k}=i|\mathcal{F}_{t_M};\hat{\Theta}_1^{(n)}\big)\big(X_{t_k}-(1-k_i^{(n+1)})X_{t_{k-1}}\big]}{k_i^{n+1}\sum_{k=2}^M\big[P\big(Z_{t_k}=i|\mathcal{F}_{t_M};\hat{\Theta}_1^{(n)}\big)\big]}\\
k_i^{(n+1)} &= \frac{\sum_{k=2}^M\big[P\big(Z_{t_k}=i|\mathcal{F}_{t_M};\hat{\Theta}_1^{(n)}\big)X_{t_{k-1}}B_1\big]}{\sum_{k=2}^M\big[P\big(Z_{t_k}=i|\mathcal{F}_{t_M};\hat{\Theta}_1^{(n)}\big)X_{t_{k-1}}B_2\big]}\\
\sigma_i^{(n+1)} &= \frac{\sum_{k=2}^M\big[P\big(Z_{t_k}=i|\mathcal{F}_{t_M};\hat{\Theta}_1^{(n)}\big)\big(X_{t_k}-k_i^{(n+1)}\theta_i^{(n+1)}(1-k_i^{(n+1)})X_{t_{k-1}}\big)^2\big]}{\sum_{k=2}^M\big[P\big(Z_{t_k}=i|\mathcal{F}_{t_M};\hat{\Theta}_1^{(n)}\big)\big]}
\end{aligned} \tag{B.17}$$

where

$$\begin{aligned}
B_1 = X_{t_k}-X_{t_{k-1}} &= \frac{\sum_{k=2}^M\big[P\big(Z_{t_k}=i|\mathcal{F}_{t_M};\hat{\Theta}_1^{(n)}\big)\big(X_{t_k}-X_{t_{k-1}}\big)}{\sum_{k=2}^M\big[P\big(Z_{t_k}=i|\mathcal{F}_{t_M};\hat{\Theta}_1^{(n)}\big)\big]}\\
B_2 &= \frac{\sum_{k=2}^M\big[P\big(Z_{t_k}=i|\mathcal{F}_{t_M};\hat{\Theta}_1^{(n)}\big)X_{t_{k-1}}\big]}{\sum_{k=2}^M\big[P\big(Z_{t_k}=i|\mathcal{F}_{t_M};\hat{\Theta}_1^{(n)}\big)\big]}X_{t_{k-1}}.
\end{aligned}$$

We then obtain the transition probabilities:

$$\Pi_{ij}^{(n+1)} = \frac{\sum_{k=2}^M\Big[P\big(Z_{t_k}=j|\mathcal{F}_{t_M};\hat{\Theta}_1^{(n)}\big)\frac{\Pi_{ij}^n P\big(Z_{t_{k-1}}=i|\mathcal{F}_{t_{k-1}};\hat{\Theta}_1^{(n)}\big)}{P\big(Z_{t_k}=j|\mathcal{F}_{t_{k-1}};\hat{\Theta}_1^{(n)}\big)}\Big]}{\sum_{k=2}^M\big[P\big(Z_{t_{k-1}}=i|\mathcal{F}_{t_{k-1}};\hat{\Theta}_1^{(n)}\big)\big]} \tag{B.18}$$

Let $\hat{\Theta}_1^{(n+1)} := (k_i^{(n+1)}, \theta_i^{(n+1)}, \sigma_i^{(n+1)}, \Pi_i^{(n+1)})$ be the new parameters of the algorithm. These are iterated in step 2 until convergence of the EM algorithm is achieved. The procedure can be stopped if either:

a) the procedure has been performed N times; or
b) the difference between the log-likelihood at step $n+1 \le N$ and the log-likelihood at step n, satisfies the equation $logL(n+1) - logL(n) < \epsilon$.

Proof of consistency of the (quasi) maximum likelihood estimators is provided in [43]; see also [58].

B.3. **Stage 2: Lévy distribution fitted to each regime.** We have estimated the regime-switching model A.12 using the EM algorithm. Now, we estimate the set of parameters $\hat{\Theta_2}$ by fitting a NIG distribution for each regime.

$$X(\text{Regime } 1) - L_1(\alpha^1, \beta^1, \delta^1, \mu^1) \tag{B.19}$$

$$X(\text{Regime } 2) - L_2(\alpha^2, \beta^2, \delta^2, \mu^2) \tag{B.20}$$

where L_1 and L_2 relate to relates to a separate set of Normal Inverse Gaussian distribution parameters of the Lévy jump process. Estimation of the distribution parameters is done by maximum likelihood, where $\Phi^1 = (\alpha^1, \beta^1, \delta^1, \mu^1)$ and $\Phi^1 = (\alpha^2, \beta^2, \delta^2, \mu^2)$. Directly following from [27], initialization of the algorithm is performed by the method of moments.

References

[1] Aboura, S., and J. Chevallier. 2017. A new weighting scheme for equity indexes. International Review of Financial Analysis 54: 159–75.

[2] Alekneviciene, V., E. Alekneviciute, and R. Rinkeviciene. 2012. Portfolio Size and Diversification Effect in Lithuanian Stock Exchange Market. Engineering Economics 23: 338–47.

[3] Alexeev, V., and F. Tapon. 2012. Equity Portfolio Diversification: How Many Stocks are Enough? Evidence from Five Developed Markets. SSRN Electronic Journal.

[4] Alexeev, V., and M. Dungey. 2015. Equity portfolio diversification with high frequency data. Quantitative Finance 15: 1205–15.

[5] Ang, A., & Timmermann, A. (2012). Regime changes and financial markets. Annu. Rev. Financ. Econ., 4(1), 313-337.

[6] Ang, A., Bekaert, G. (2002) International Asset Allocation with Regime Shifts, *Review of Financial Studies*, 15, 1137–1187.

[7] Ang, A., Bekaert, G. (2002) Regime Switching in Interest Rates. *Journal of Business and Economic Statistics* 20: 163-182.

[8] Arjoon, V., and C. S. Bhatnagar. 2017. Dynamic herding analysis in a frontier market. Research in International Business and Finance 42: 496–508.

[9] Baele, L. (2005) Volatility Spillover Effects in European Equity Markets. *Journal of Financial and Quantitative Analysis* 40:373-402

[10] Barndorff-Nielsen, O. E. (1997) Normal Inverse Gaussian Distributions and Stochastic Volatility modelling. *Scand. J. Statist.*, 24:1-13.

[11] Barndorff-Nielsen, O. E. and Blaesild, P., (1981) Hyperbolic Distributions and Ramifications: Contributions to Theory and Application. *Statistical Distributions in Scientific Work*, 4: 19–44.

[12] Barndorff-Nielsen, O. E., Baesild, P., Seshadri, V. (1992) Multivariate Distributions with Generalized Inverse Gaussian Marginals, and Associated Poisson Mixtures. *The Canadian Journal of Statistics*, 20: 109–120.

[13] Barndorff-Nielsen, O. E., Blaesild, P., Jensen, J. L., Sorensen, M. (1985). The fascination of sand. In Atkinson, A. C., Fienberg, S. E., editors, *A Celebration of Statistics*, p. 57-87. Springer, New York.

[14] Barndorff-Nielsen, O. E., T. Mikosch, Resnick, S. I., (2001) *Lévy Processes: Theory and Applications*. Berlin, Germany: Springer Science & Business Media.

[15] Barndorff-Nielsen, O., & Blaesild, P. (2004). Hyperbolic distributions. Encyclopedia of statistical sciences, 5.

[16] Beck, Kristine L., Sieven B. Perfect, and Pamela P. Peterson. 1996. The Role of Alternative Methodology on the Relation Between Portfolio Size and Diversification. The Financial Review 31: 381–406.

[17] S. Benartzi and R. Thaler. Naive Diversification Strategies in Defined Contribution Savings Plans. The American Economic Review, pages 79–98, March 2001.

[18] Benjelloun, H. 2010. Evans and Archer—Forty years later. Investment Management and Financial Innovations 7: 8.

[19] Benth, F. E., and Lempa, J. (2014). Optimal portfolios in commodity futures markets. Finance and Stochastics, 18(2), 407-430.

[20] Black, F., and M. Scholes (1973): "The Pricing of Options and Corporate Liabilities," Journal of Political Economy, 81, 637–654

[21] Bolviken, Erik, and Fred Espen Benth. "Quantification of risk in Norwegian stocks via the normal inverse Gaussian distribution." Proceedings of the AFIR 2000 Colloquium, Tromso, Norway. Vol. 8798. AFIR, 2000.

[22] Brands, S., and D.R. Gallagher. 2005. Portfolio selection, diversification and fund-of-funds: A note. Accounting & Finance 45: 185–97.

[23] Campbell, J.Y., M. Lettau, B.G. Malkiel, and Y. Xu. 2001. Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk. The Journal of Finance 56: 1–43.

[24] Chauvet M., (1998) An econometric characterization of business cycle dynamics with factor structure and regime switching. *International economic review*, pp 969–996.

[25] Chen, F., O. Hope, Q. Li, and X. Wang. 2018. Flight to Quality in International Markets: Investors' Demand for Financial Reporting Quality during Political Uncertainty Events. Contemporary Accounting Research 35: 117–55.

[26] Chen, K., L. Vitiello, S. Hyde, and S. Poon. 2018. The reality of the stock market jumps diversification. Journal of International Money and Finance 86: 171–88.

[27] Chevallier, J., Goutte, S. (2016) On the estimation of regime-switching Lévy models. *Studies in Nonlinear Dynamics & Econometrics*, 21(1), pp. 3-29.

[28] Di Persio L., and Frigo M., (2016) Gibbs sampling approach to regime-switching analysis of financial time series. *Journal of Computational and Applied Mathematics*, 300:43 – 55.

[29] Diebold, FX, Lee, JH., and GC Weinbach. (1994) Regime switching with time-varying transition probabilities. *Business Cycles: Durations, Dynamics, and Forecasting*, pages 144–165.

[30] Ekeland, I., and E. Taflin. "A theory of bond portfolios." The Annals of Applied Probability 15, no. 2 (2005): 1260-1305.

[31] Eriksson, A., Ghysels, E., Wang W. (2009) The Normal Inverse Gaussian Distribution and the Pricing of Derivatives, *The Journal of Derivatives*, 16, 23-37;

[32] Evans, JL., and SH. Archer. 1968. Diversification and the Reduction of Dispersion: An Empirical Analysis. The Journal of Finance 23: 761–67.

[33] Fielitz, BD. 1974. Indirect versus Direct Diversification. Financial Management 3: 54–62.

[34] Flavin, TJ., and Michael R.Wickens. 2006. Optimal International Asset Allocation with Time-Varying Risk. Scottish Journal of Political Economy 53: 543–64.

[35] Fong, Tom Pak Wing, Angela Kin Wan Sze, and Edmund Ho Cheung Ho. 2018. Determinants of equity mutual fund flows—Evidence from the fund flow dynamics between Hong Kong and global markets. Journal of International Financial Markets, Institutions and Money 57: 231–47.

[36] Fragiadakis, K., Karlis, D., and Meintanis S.G. (2009) Tests of fit for normal inverse Gaussian distributions, *Statistical Methodology*, 6,553-564;

[37] Hamilton JD., (1989) A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the Econometric Society*, pp357–384, 1989.

[38] Hamilton, J. D. (2016). Macroeconomic regimes and regime shifts. Handbook of macroeconomics, 2, 163-201.

[39] Jobson J. D. and B. M. Korkie. Estimation for Markowitz Efficient Portfolios. Journal of the American Statistical Association, 75(371):544–554, 1980.

[40] Jonsson H., Masol V., Schoutens W., (2010) Normal Inverse Gaussian Model, in Cont R. (Edt.)*Encyclopedia of Quantitative Finance*, Vol 3, 1311-1314.

[41] Kalemanova A., Schmid B., Werner R., (2007) The normal inverse Gaussian distribution for synthetic CDO pricing. *Journal of Derivatives*, 14(3):80.

[42] Keller, U., (1997) Realistic Modelling of Financial Derivatives. University of Freiburg/Breisgau.

[43] Kim, CJ., (1994) Dynamic Linear Models with Markov Switching, *Journal of Econometrics*, 60, 1-22.

[44] Kim, CJ., Nelson, CR et al. (1999) *State-space models with regime switching: classical and Gibbs sampling approaches with applications*, volume 2. MIT press Cambridge, MA, 1999.

[45] Kryzanowski, L., and S.Singh. 2010. Should Minimum Portfolio Sizes Be Prescribed for Achieving Sufficiently Well- Diversified Equity Portfolios? SSRN Electronic Journal 7: 1–37.

[46] Kushnirovich, N. 2016. Immigrant investors in financial markets: Modes of financial behavior. Journal of Business Economics and Management 17: 992–1006.

[47] Lee, D.D., J.E. Humphrey, K.L. Benson, and J.Y.K. Ahn. 2010. Socially responsible investment fund performance: The impact of screening intensity. Accounting & Finance 50: 351–70.

[48] Markowitz, H. 1952. Portfolio Selection. The Journal of Finance 7: 77–91.

[49] Meucci, A. (2005). Risk and asset allocation (Vol. 1). New York: Springer.

[50] Merton, R. C. (1976) Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics* 3 125-144.

[51] Michaud R. Efficient Asset Management: A Practical Guide to Stock Portfolio Optimization and Asset Allocation. Harvard Business School Press, 1998.

[52] Pastor, L., R.F. Stambaugh, and L.A. Taylor. 2020. Fund tradeoffs. Journal of Financial Economics 138: 614–34.

[53] Pelagatti M., (2005) Duration dependent markov-switching vector autoregression: properties, bayesian inference, software and application. *Bayesian Inference, Software and Application*

[54] Prause, K. (1999c) The generalized hyperbolic model: Estimation, financial derivatives and risk measures. PhD thesis, University of Freiburg.

[55] Rachev S., Stoyanov S., Mittnik S., Fabozzi F., (2017) Behavioral Finance – Asset Prices Predictability, Equity Premium Puzzle, Volatility Puzzle: The Rational Finance Approach, Papers 1710.03211, arXiv.org.

[56] Raju, R., and S. Kumar Agarwalla. 2021. Equity Portfolio Diversification: How Many Stocks Are Enough? Evidence from India. SSRN Electronic Journal.

[57] Ringer, N., and Tehranchi, M. (2006). Optimal portfolio choice in the bond market. Finance and Stochastics, 10(4), 553-573.

[58] Ruiz, E., (1994), Quasi-maximum likelihood estimation of stochastic volatility models, *Journal of Econometrics*, 63, issue 1, p. 289-306

[59] Sharma, P., and Vipul. 2018. Improving portfolio diversification: Identifying the right baskets for putting your eggs. Managerial and Decision Economics 39: 698–711.

[60] Solnik, BH. 1974. Why Not Diversify Internationally Rather Than Domestically? Financial Analysts Journal 30: 48–54.

[61] Statman, M. 1987. How Many Stocks Make a Diversified Portfolio? Journal of Financial and Quantitative Analysis 22: 353–63.

[62] Stotz, A., and W. Lu. 2014. Eight Stocks are Enough in China. SSRN Electronic Journal.

[63] Suarez-Garcia P., Gomez-Ullate D., (2013) Scaling, stability and distribution of the high-frequency returns of the Ibex35 index, *Physica A*, 392:6, 1409-1417.

[64] Tang, G.Y.N. 2004. How efficient is naive portfolio diversification? an educational note. Omega 32: 155–60.

[65] Windcliff, H., and Boyle, P. P. (2004). The 1/n pension investment puzzle. North American Actuarial Journal, 8(3), 32-45.

[66] Zaimovic, A., A. Omanovic, and A. Arnaut-Berilo. 2021. "How Many Stocks Are Sufficient for Equity Portfolio Diversification? A Review of the Literature" Journal of Risk and Financial Management 14, no. 11: 551.

[67] Zhou, Y. 2014. Number of Stocks in Portfolio and Risk Reduction. Master's thesis, University of California, Los Angeles, CA, USA.