Regime-Switching And Levy Jump Dynamics In Option-Adjusted Spreads

Shaw, Charles

Birkbeck

28 December 2018

Regime-Switching And Lévy Jump Dynamics In Option-Adjusted Spreads

Charles Shaw*

May 27, 2019

Abstract

A regime-switching Lévy framework, where all parameter values depend on the value of a continuous time Markov chain as per Chevallier and Goutte (2017), is employed to study US Corporate Option-Adjusted Spreads (OASs). For modelling purposes we assume a Normal Inverse Gaussian distribution, allowing heavier tails and skewness. After the Expectation-Maximization algorithm is applied to this general class of regime switching models, we compare the obtained results with time series models without jumps, including one with regime switching and one without. We find that a regime-switching Lévy model clearly defines two regimes for A-, AA-, and AAA-rated OASs. We find further evidence of regime-switching effects, with data showing relatively pronounced jump intensity around the time of major crisis periods, thereby confirming the presence and importance of volatility regimes. Results indicate that ignoring the complex and dynamic dependence structure in favour of certain model assumptions may lead to a significant underestimation of risk.

*Author email: cshaw11@mail.bbk.ac.uk

1 Introduction

This paper deploys a regime switching Lévy model to examine regime changes in Option-Adjusted Spreads (OASs), which are the calculated spreads between a computed OAS index of all bonds in a given rating category and a spot Treasury curve. An OAS index is constructed using each constituent bond's OAS, weighted by market capitalization. OASs represent a measure of credit risk in option-embedded bonds such[1].

These are of interest for many reasons. One such reason is that corporate bond yields signal the cost of financing for private firms. Higher spreads are indicative that the cost of capital is higher and, therefore, that the profitability of investment opportunities is lower. And since investment in physical capital is a key driver of economic growth, understanding the structure of the cost of financing for private firms helps to identify the barriers to productive investment.

To the best of our knowledge, this study is the first time that Markov-switching Lévy models have been employed to analyze the structure of option-adjusted spreads. For modelling purposes we assume a Normal Inverse Gaussian distribution, allowing heavier tails and skewness. This is followed by a comparison with other time-series models. The maximum likelihood estimates are determined by the EM algorithm. In order to measure the quality of the regime classification, we deploy two measures:

1. The regime classification measure (RCM) introduced by Ang and Bekaert (2002) in [2].
2. The smoothed probability indicator.

The intuition behind regime-switching models is that the parameters of an autoregression rely upon a stochastic and unobservable regime variable which represents the probability of being in a particular state. So, in other words, the evolution of regimes can be inferred – we have to form an inference since it cannot be observed directly – from the data once a law has been specified for the states.

The regular flow of economic activity may occasionally suffer shocks substantive enough to result in different observed dynamics. Sampled time series data may typically show not only periods of low and high volatility but also periods of slower and faster mean growth. In such cases GARCH-type models often do not perform well empirically and may even be inappropriate. But following the seminal work of Hamilton [33, 34], we have a more useful framework: that of stochastic regime switching. When converted into a continuous-time, this model implies that the underlying asset price can switch between two states, exhibiting continuous changes in each state. Such shifts are governed by a Markov (point) process.

In a seminal contribution, [25] has demonstrated the practicality of a segmented trends model, i.e. that a time series may be segmented into a sequence of stochastic time trends. Regime switching models have since been used widely in the literature across various domains of application, including but not limited to analysis of energy prices [35, 48], exchange rates [11], stock returns [34, 23], systemic risk [13], asset allocation [28, 62], international equity markets [1, 50], business cycles [31, 49], economic growth [38], term structure [24], and monetary policy [58, 59, 12]. Surveys are provided by [32] and [3].

The use of jump-diffusion models in financial applications can be traced to [47], and later [10]. These early models relied on the two pivotal ideas. First, that the Poisson jump-driven part of the model explains large market shifts in response to unexpected information. Second, that the diffusive (i.e. the Wiener process-driven) part of the model explains normal asset price variations. Such models were able to specify a finite number of jumps in a finite time interval.

[1]Eg callable and putable bonds, etc

Building on this literature, some of the more recent models have suggested models with infinitely many jumps in finite time intervals.

These models include the variance gamma model of Madan and Seneta [45], and the CGMY model of Carr, Geman, Madan and Yor [16], amongst others. [16] study the variance gamma (VG) and normal inverse Gaussian (NIG) as two examples of time-changed Lévy processes. The VG and NIG are obtained by replacing the time of a Brownian motion with the inverse Gaussian and gamma process, respectively. More concretely, the VG and NIG processes are subordinators belonging to the class of Lévy processes. The distinct advantage of the latter set of models was their ability to capture both large and infrequent jumps, as well as small and frequent ones.

Although the literature on switching regime Lévy processes remains sparse, we are witnessing a rise in application of Lévy models to study different asset classes and markets. Some of the better known examples include [20] who show empirical evidence that commodity prices demonstrate jumps. More recently, empirical work by [15] underpins the importance of jumps across a range of asset classes. The presence of jump-diffusion processes (including Lévy processes) is empirically supported by [37], who studied the S&P 500 index. A survey is provided by [29].

2 Data

The data of interest consists of the Bank of America Merrill Lynch investment-grade ("U.S. Corporate Master") and high-yield ("U.S. High Yield Master") corporate bond indices. The BAML dataset relies on the industry standard for valuations, aggregating data from TRACE as well as other sources. For an more detailed description of this data set see Schaefer and Strebulaev [57]. Figure 1 graphically shows the data series, sourced from Federal Reserve Economic Data (FRED). The date range is 03/01/2000-24/12/2018, n=5014 obs. The individual series are as follows:

1 ICE BofAML US Corporate A Option-Adjusted Spread

2 ICE BofAML US Corporate AA Option-Adjusted Spread

3 ICE BofAML US Corporate AAA Option-Adjusted Spread

4 ICE BofAML US High Yield B Option-Adjusted Spread

5 ICE BofAML US High Yield BB Option-Adjusted Spread

6 ICE BofAML US Corporate BBB Option-Adjusted Spread

7 ICE BofAML US High Yield CCC or Below Option-Adjusted Spread

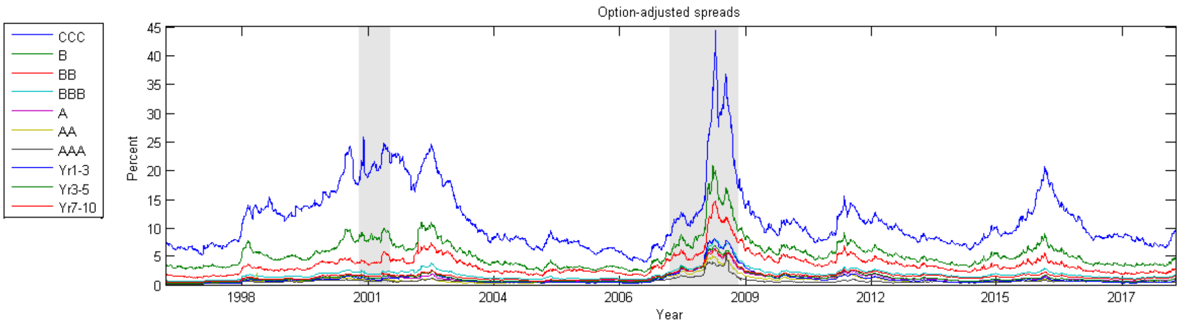
8 ICE BofAML US Corporate 1-3 Year Option-Adjusted Spread

9 ICE BofAML US Corporate 3-5 Year Option-Adjusted Spread

10 ICE BofAML US Corporate 7-10 Year Option-Adjusted Spread

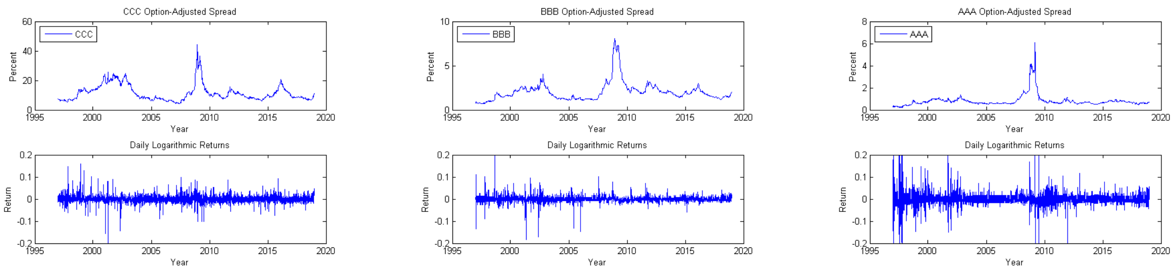
Figure 2 shows time series plots for CCC-, BBB-, and AAA-rated credit in our sample. Common 'stylized facts' of financial time series are clearly apparent. For example, volatility clustering can be observed across all the returns data , particularly at the start of 2000s and during the North Atlantic financial crisis from Q2 2007 to Q2 2009. This is consistent with expectations since it is known that around the middle of 2007 is when dislocation in the subprime mortgage markets first became apparent. The Federal Reserve's first policy response to the crisis, namely provision of liquidity, was in August (Board of the Governors of the Federal Reserve System, 2007, [14]).

Figure 1: ICE BofAML Option-Adjusted Spreads



Source: FRED. Shaded areas indicate U.S. recessions (NBER).

Figure 2: Closing price and log returns: CCC-, BBB-, and AAA-rated credit.



We conduct the Augmented Dickey-Fuller on our closing price data. This tests the null hypothesis that our data follows a unit root process. We fail to reject the null hypothesis of a unit root against the autoregressive alternative for all ten series in our sample. Number of lags were selected using the BIC selection criterion.

When testing for normality we use the Lilliefors goodness-of fit test, which is a version of the Kolmogorov-Smirnov test corrected for the potential presence of parameter uncertainty. P-values for the null hypothesis of normality for the returns of our series return $p < 0.001\%$ in all ten cases. We do not make use of omnibus tests (a la Jarque-Bera) as they are suspected of being underpowered. Demerits of of 'omnibus' tests for normality of the JB type, are discussed in [19, 46, 61] and [63].

Trying to model financial time series is fraught with problems, since the observations can be influenced by events that are largely unpredictable. Such events – which may include natural disasters, statements from central banks, policy announcements from governments – have the potential to profoundly affect the market. As a result, the assumption of stationarity may not hold for financial data. The implication of this is that classic time series analysis techniques may be partially or completely inadequate to model financial data. In some cases, solutions to this problem can be found by deploying Markov-switching models, since these models let us, under certain mild assumptions, address the non-stationarity of time series data.

The idea behind such models is that the distribution of the observations is allowed to change over time. By way of exposition, we can write a general Markov Regime Switching (MRS) model in the following way:

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

where θ is the vector of the parameters of the model, S_t is the state of the model at time t, $\psi_t := \{y_k : k = 1, \dots, t\}$ 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 1 show how the Markov-switching approach can be fruitful for time series applications, since realizations of 1 let us approach specific problems that may be difficult to model in a singe state regime. Although the literature on Markov-switching models is broad, it appears that it can be separated into two general groups. The first consists of models that have complicated distributions for the data or a large number of states, but basic transition laws, such as a first order Markov chain. For examples of studies in this vein see [21, 30, 17]. The second group consists of models with simple assumptions and very few states, often two, but with more complicated transition laws. For examples of such studies see, e.g. [44, 22, 51].

3 Markov-switching model augmented by jumps

We now turn our attention to estimation of a Markov-switching model augmented by jumps, under the form of a Lévy process, with a view to applying this methodology to study OAS returns. In order to motivate this section's modelling approach, we first set up the general structure of Lévy processes, then we outline their properties with reference to path variation and the Lévy-Khintchine theorem. Links to infinitely divisible distributions are also provided. The discussion also highlights the properties of Markov chain, such as irreducibility, aperiodicity, and ergodicity. The section concludes with a discussion on a framework for estimating a jump-robust model tempered by a Markov chain, which can be used to study the relations of dependence within OAS returns and related time series. Estimation in such a framework can be performed using the EM-algorithm.

3.1 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 two. A well-known pure diffusion process used in finance is the oft-used 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 prices 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 generalized to define a common framework of one-dimensional stochastic (Lévy) processes.

3.2 A regime-switching Lévy model

This subsection motivates the introduction of the regime-switching Lévy approach to modelling of our time series. This part of the discussion follows the theoretical methodology recently proposed by [18], namely by combining a Lévy jump-diffusion model with a Markov-switching framework. However, we deploy it to study the structure of OAS returns. This is motivated by the apparent presence of discontinuities in OAS data series, which is a prompt to incorporate stochastic jumps into the modelling process. 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 number of key definitions and notations.

Definition 3.1 (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 3.2 (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: $W_t - W_s \sim N(0, t-s)$ $\forall$ $0 \leq s < t < \infty$

Property (*ii*) implies the Markov property i.e. conditional probability distribution of future states of the process depend 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 Poisson Process, another stochastic process as follows.

Definition 3.3 (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 3.4 (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: $\lim_{t\to s} L_t = L_s$

[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.

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

Definition 3.5. *A real valued random variable Θ has an infinitely divisible distribution if for each $n = 1, 2, \dots,$ there exists a i.i.d. sequence of random variables $\Theta_1, \dots, \Theta_n$ 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 $n = 1, 2, \dots$ 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 3.6 (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,*

$$\Phi(u;t) = \int_{\mathbb{R}} e^{iux} \mu(dx) = e^{-t\Psi(u)} \quad for \quad u \in \mathbb{R}, \tag{2}$$

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

$$\Psi(\lambda) = i\gamma u + \frac{\sigma^2 u^2}{2} + \int_{\mathbb{R}} \left(1 - e^{(iux)} + iux\mathbf{1}_{|x|<1}\right) \Pi(dx) \tag{3}$$

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

From Theorem 3.6 we can say that there exists a probability space where $L = L^{(1)} + L^{(2)} + L^{(3)}$; $L^{(1)}$ 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 1 (a.s.). This is the the Lévy-Itô decomposition, which can be stated as follows

$$L_t = \eta t + \sigma W_t + \int_0^t \int_{|x|\geq 1} x\mu^L(ds, dx) + \int_0^t \int_{|x|<1} x(\eta^L - \Pi^L)(ds, dx). \tag{4}$$

Definition 3.7 (Markov-Switching)**.** *Let $(Z_t)_{t\in[0,T]}$ be a continuous time Markov chain on finite space $S := \{1, \dots, K\}$. 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*

$$\Pi_{ij}^Z \begin{cases} \geq 0, & if\ i \neq j \\ -\sum_{j\neq i} \Pi_{ij}^Z, & otherwise \end{cases} \tag{5}$$

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

Definition 3.8 (Regime-switching Lévy model)**.** *For all $t \in [0, T]$, let Z_t be a continuous time Markov chain on finite space $S := \{1, \dots, K\}$ defined as per Definition 3.7. A regime-switching model is a stochastic process (X_t) which is solution of the stochastic differential equation given by*

$$dX_t = k(Z_t)(\theta(Z_t) - X_t)dt + \sigma(Z_t)dY_t \tag{6}$$

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 $\sigma(Z_t)$: $k(Z_t) := \{k(1), \ldots, k(K)\} \in \mathbb{R}^{K^*}$, $\theta(S) := \{\theta(1), \ldots, \theta(K)\}$, $\sigma(S) := \{\sigma(1), \ldots, \sigma(K)\} \in \mathbb{R}^{K^*}$. 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 stochasticity: 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.

3.3 NIG-type distribution

Following [18], let us 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 [5, 7, 8, 9]).

The NIG type distribution is a relatively novel process introduced by Barndorff-Neilsen [7] as a model for log returns of stock prices. It is a sub-class of the more general class of hyperbolic Lévy processes. After its introduction it was demonstrated that the NIG distribution provides an excellent fit to log returns of stock market data [6]. Other studies have also shown this distribution's superior empirical fit to other asset classes [54, 39, 26, 27]. Using IBEX35 data, [60] find that the Normal Inverse Gaussian distribution provides an overall fit for the data better than any of the other subclasses of Generalized Hyperbolic distributions and much better than the Lévy-stable laws. More recently, Rachev et al [55] have deployed the NIG distribution, together with other statistical machinery, to study and (in their words) resolve such well-known 'puzzles' as (i) Predictability of asset returns (ii) The Equity Premium, and (iii) The Volatility Puzzle.

This type of types of heavy-tailed process is expected to draws increasing interest, particularly since the NIG distribution fulfils the fat-tails condition, is analytically tractable, yet is closed under convolution [36].

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

$$f_{NIG}(x; \alpha, \beta, \delta, \mu) = \frac{\alpha}{\pi} e^{\delta \sqrt{\alpha^2 - \beta^2 + \beta(x-\mu)}} \frac{K_1(\alpha\delta\sqrt{1 + (x-\mu)^2/\delta^2})}{\sqrt{1 + (x-\mu)^2/\delta^2}}, \tag{7}$$

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, 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

$$K_v(z) = \frac{1}{2} \int_0^\infty y^{v-1} e^{\left(-\frac{1}{2}z\left(y + \frac{1}{y}\right)\right)} dy \tag{8}$$

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

$$v^2 y'' + xy' - (x^2 + v^2)y = 0. \tag{9}$$

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

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

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

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

An expectation-maximization algorithm for the normal-inverse Gaussian distribution was proposed by Karlis [40] and more recently generalised to include Lévy processes by Chevallier and Goutte [18]. Again, the following exposition follows [18], in content, in model construction, and in notation.

4 Estimation procedure

We now turn our attention to the estimation procedure of the Lévy regime-switching model in question. The EM algorithm used to estimate the regime-switching Lévy model, namely the SDE given in Equation (6), is provided by [18]. The algorithm fits a regime-switching Lévy model where the stochastic process Y is a Lévy process that follows a Normal Inverse Gaussian distribution. The set of parameters that require estimation is

$$\hat{\Theta} := (\hat{k}_i, \hat{\theta}_i, \hat{\sigma}_i, \hat{\alpha}_i, \hat{\beta}_i, \hat{\delta}_i, \hat{\mu}_i, \hat{\Pi}) \quad i \in S.$$

There are four parameters of the density of the Lévy process L, three parameters of the dynamics of X, and the transition matrix of the Markov chain Z. After model is discretized, the global set of parameters are estimated in a two-step procedure.

4.1 Discretization procedure

Let us consider Wiener process W for stochastic process Y. Let Γ be the increasing sequence of time from which the data values are taken:

$$\Gamma = \{t_j; 0 = t_0 \leq t_1 \leq \dots t_{M-1} \leq t_M = T\}, \ with \ \Delta_t = t_j - t_{j-1} = 1.$$

In this specification, $M + 1$ denotes to the size of historical data. The discretized version of the SDE given in 6 is

$$X_{t+1} = k(Z_t)\theta(Z_t) + (1 - k(Z_t))X_t + \sigma(Z_t)\epsilon_{t+1}. \tag{10}$$

Since Y is a Wiener process, $\epsilon_{t+1} \sim N(0, 1)$. Let $\mathcal{F}^X_{t_k}$ be the vector of historical values of the process X until time $t_k \in \Gamma$. Then $\mathcal{F}^X_{t_k}$ is a vector of $k+1$ values of the discretized model. Thus $\mathcal{F}^X_{t_k} = (X_{t_0}, X_{t_1}, \dots, X_{t_k})$.

5 Estimation procedure

Next, we proceed with estimating our model in two stages.

- **Stage 1**: Estimation of the regime-switching model 6 in the Wiener case. Here we estimate the parameters of the discretized model 10. We use the EM-algorithm. But in order to deploy the EM-algorithm, the parameter space estimate $\hat{\Theta}$ is first divided into $\hat{\Theta}_1 := (\hat{k}_i, \hat{\theta}_i, \hat{\sigma}_i, \hat{\Pi}_i)$ for $i \in S$.

- **Stage 2**: Estimation of the parameters of the Lévy process fitted to each regime. Using the regime classification obtained in Step 1, we estimate the next subset of parameters $\hat{\Theta}_2 := (\hat{\alpha}_i, \hat{\beta}_i, \hat{\delta}_i, \hat{\mu}_i)$ for $i \in S$. This relates to the Normal Inverse Gaussian distribution parameters of the Lévy jump process fitted for each regime.

Further details of the 2-stage estimation, iincluding the EM algorythm are given in the Appendix.

6 Empirical findings

6.1 Regime Classification Measure of Ang and Bekaert (2002)

A great model is one that is able to sharply classifies the regimes, whilst smoothed probabilities should be either $\approx$ zero or $\approx$ one. To address this, Regime Classification Measures (RCMs) have been proposed by Ang and Bekaert [2] as a way to determine if the number of regimes K is appropriate. The RCM statistic spans from 0 (perfect regime classification) to 100 (failure to detect any regime classification). The RCM was extended for multiple states by Baele [4].

$$RCM(K) = 100 \times \Big(1 - \frac{K}{K-1}\frac{1}{T}\sum_{k=1}^{N}\sum_{Z_{t_k}}^{N}\Big(P(Z_{t_k} = i|\mathcal{F}_{t_M}^{X};\hat{\Theta}_1^{(n)}) - \frac{1}{K}\Big)^2\Big), \tag{11}$$

where $P(Z_{t_k} = i|\mathcal{F}_{t_M}^{X};\hat{\Theta}_1^{(n)})$ corresponds to the smoothed probability and $\hat{\Theta}_1^{(n)}$ 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 a RCM of almost 0, a good model will have a RCM of close to ≈ 0, while a model that cannot distinguish between regimes at all will have a RCM close to 100. A good model is one that 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.

6.2 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 $k \in 1, \ldots, N$ 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 1 by $P^{p\%}$.

6.3 Capturing the quality of a model's regime qualification performance

In Table 1 we observe that the RCM statistic for A-rated indices is less than 10, indicating a good MRS model fit. We take this as evidence that the our model is able to model the data reasonably well with two regimes. This model fit is also comparable, albeit with weaker results, with B- or C-rated bond data. We can also observe that the smoothed probability indicator is equal to 0.96 for AAA OASs, to 0.92 for AA bonds, and 0.91 for A bonds. We take this as evidence that in this case both regimes are clearly defined. These regimes can be thought of as different means in the growth rate. On the other hand, the indicator for B- and C- rated OASs is not as near to the upper bound of 1. Performance for probability indicators for Corporate 7-10 Year, 3-5 Year, and 1-3 Year sits somewhere in between.

As a comparison, we run the model on various stock market indices data and found results to be significantly worse. Using data available from Oxford-Man Institute of Quantitative Finance[3], we ran our model on 31 stock indices. Same time horizon was used where possible[4]. For example, running this model on Nikkei data yielded RCM of 33.27 and probability indicator p=0.66. Out of 31 stock index series, only the FTSE and the Amsterdam Exchange index performed well with RCM < 10 and $p > 0.9$. This seems to be in contrast to some select claims made in the

[3]Data is available via `https://realized.oxford-man.ox.ac.uk/data/assets`

[4]Not all time series have the same starting point, for example data for the OMX Copenhagen 20 commences in late 2005. However, all data series extend until the end of 2018, as per other time series in study

Table 1: Regime Classification Measure, Ang and Bekaert (2002) and smoothed probability indicator.

US Corporate OASs	RCM	p^{10}
AAA	4.2696	0.9569
AA	7.2448	0.9222
A	8.2493	0.9107
BBB	14.3738	0.8468
BB	15.9911	0.8347
B	14.0505	0.8494
CCC and below	13.6773	0.8552
7-10 Year	13.9954	0.8536
3-5 Year	10.3878	0.8923
1-3 Year	8.8628	0.9019

literature. After going back to the literature and conducting light replications of some studies that highlighted the relative performance of similar frameworks (MRS-Lévy, Variance Gamma, NIG), it appears that such models are sensitive to the selection of the span of the selected time series.

That notwithstanding, given the set of models considered in our analysis, and using a reasonably long time span (n=5014 obs), we can conclude that a Lévy regime-switching model clearly defines two regimes for the A-, AA-, and AAA-rated US Corporate OASs in our sample.

If we look at Figures 3, 4, 5, and 6 then we can observe that we can see that our regime switching model captures the effect of the North Atlantic Financial Crisis of 2007-2008. In particular, Figures 3 shows that the AAA option-adjusted spread switches to a high regime of variance during the period 2008-2010[5]. Thus, during this period of financial instability our model is in a regime of high variance which maps to higher levels of volatility in the US corporate bond market. This result is in line with expectations, since we expect that volatility should be higher in times of crisis than in other economic periods. This reaffirms our point that the use of two different Markov chains enables us to highlight different levels of mean for a same level of variance or different levels of variance for a same level of mean. Other indicators of risk, such as the VIX, increased sharply but briefly at the same time, see Figure 7.

7 Comparison against benchmark models

We now compare the performance of our model against benchmark models, including one with a single regime specification.

1. Regime-switching Lévy model. This is our headline model specified by Equation 5. The process $Y = L$ is a Lévy process such that $L_1 \sim NIG(\alpha, \beta, \mu)$.

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

2. Regime switching Gaussian. This is the same as Equation 5 but the process $Y = W$ is a Brownian motion.

$$dX_t = k(Z_t)(\theta(Z_t) - X_t)dt + \sigma(Z_t)dY_t dW_t \tag{13}$$

[5]The other two episodes in higher variance correspond to the time of the 2001 recession in US and the Russian default of 98. However, it should be noted that for the purposes of our model estimation our time series starts in 03/01/2000.

3. Vasicek. This model is specified as per Equation 5 but here the process $Y = W$ is a Brownian motion without regime switches. The corresponding stochastic differential equation does not depend on the Markov chain Z.

$$dX_t = k(Z_t)(\theta(Z_t) - X_t)dt + \sigma dW_t \tag{14}$$

We now examine the goodness-of-fit of these competing models by calculating log-likelihood values with the EM-algorithm. We also report the information critera (AIC and BIC) for each model. The lowest model selection criteria, and therefore the best model, is highlighted in italics. The results in Table 2 show that a regime switching Gaussian model without jumps (i.e. a regime-switching Gaussian diffusion model where only the drift component is regime-dependent) is preferred.

Table 2: Benchmark model selection

OASs	RS Lévy		RS Gaussian		Vasicek	
	BIC	*AIC*	*BIC*	*AIC*	*BIC*	*AIC*
AAA	73.84	21.68	*-31028.78*	*-31080.94*	-517.28	-536.84
AA	71.14	18.98	*-30867.58*	*-30919.74*	-22114.85	-22134.41
A	69.30	17.14	*-30391.76*	*-30443.92*	-22004.36	-22023.92
BBB	69.45	17.29	*-27650.98*	*-27703.15*	-21706.83	-21726.39
BB	72.82	20.66	*-14197.10*	*-14249.26*	-10368.04	-10387.61
B	73.42	21.26	*-11145.15*	*-11197.31*	-7112.83	-7132.39
CCC	74.83	22.67	*-7425.13*	*-31080.94*	-517.28	-536.84
7-10 Year	69.48	17.32	*-28151.61*	*-28203.77*	-22772.30	-22791.87
3-5 Year	69.38	17.22	*-28853.66*	*-28905.82*	-22416.79	-22436.35
1-3 Year	71.31	19.15	*-27742.38*	*-27794.54*	-18938.98	-18958.54

As is customary, we will now briefly discuss possible directions for future work. Looking forward, our results suggest that there is still much to be achieved by virtue of departures from the modeling assumptions used in a traditional time series models a la ARCH-GARCH. One obvious direction is to extend the model presented herein to a more generalized framework whilst making use of recent research on MRS models. For example, until recently deriving a likelihood ratio test statistic for testing the number of regimes in MRS models remained an open statistical problem. However, Kasahara and Shimotsu [41][6] have recently presented an asymptotic distribution of the likelihood ratio test statistic for testing the number of regimes in MRS models. This framework has the potential to be usefully employed to augment the model employed in this paper and extend it to a more general setting.

8 Conclusion

To conclude, a regime-switching Lévy framework, where all parameter values depend on the value of a continuous time Markov chain as per Chevallier and Goutte (2017), was employed to study Option-Adjusted Spreads (OASs). For modelling purposes we assumed a Normal Inverse Gaussian distribution, allowing heavier tails and skewness.

We motivated this paper's modelling approach by setting up the general structure of Lévy processes before outlining their properties with reference to path variation and the Lévy-Khintchine theorem. Estimation was done using the EM-algorithm.

[6] As yet unpublished working paper, arXiv:1801.06862

We found that a regime-switching Lévy model clearly defines two regimes for A-, AA-, and AAA-rated OASs. We found further evidence of regime-switching effects, with data showing relatively pronounced jump intensity around the time of major crisis periods, thereby confirming the presence and importance of volatility regimes.

The discussion highlighted the properties of Markov chain, such as irreducibility, aperiodicity, and ergodicity. We discussion the potential merits and demerits of estimating a jump-robust model tempered by a Markov chain. When comparing the regime-switching Lévy model to other benchmark time series models, we concluded that a regime-switching Gaussian diffusion model where only the drift component is regime-dependent is preferred.

Results indicate that ignoring the complex and dynamic dependence structure in favour of certain model assumptions may lead to a significant underestimation of risk.

9 Appendix

9.1 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. Then calculation in the previous iteration of the algorithm gives the following vector set $\hat{\Theta}_1^{(n)} := \left(\hat{k}_i^{(n)}, \hat{\theta}_i^{(n)}, \hat{\sigma}_i^{(n)}, \hat{\Pi}_i^{(n)}\right)$

9.2 EM-algorithm

9.2.1 Expectation step (E step)

We aim to estimate both filtered probability and smoothed probability. Optimality is achieved when a model is able to 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 $k = \{1, 2, \ldots, M\}$, estimate the following

$$\begin{aligned} P\Big(Z_{t_k} = i|\mathcal{F}_{t_k}^X; \hat{\Theta}_1^{(n)}\Big) &= \frac{P\Big(Z_{t_k}, X_{t_k}|\mathcal{F}_{t_{k-1}}^X; \hat{\Theta}_1^{(n)}\Big)}{f\big(X_{t_k}|\mathcal{F}_{t_{k-1}}^X; \hat{\Theta}_1^{(n)}\big)} \\ &= \frac{P\Big(Z_{t_k} = i|\mathcal{F}_{t_{k-1}}^X; \hat{\Theta}_1^{(n)}\Big) f\Big(X_{t_k}|Z_{t_k} = i; \mathcal{F}_{t_{k-1}}; \hat{\Theta}_1^{(n)}\Big)}{\sum_{j \in S} P\Big(Z_{t_k} = j|\mathcal{F}_{t_{k-1}}^X; \hat{\Theta}_1^{(n)}\Big) f\Big(X_{t_k}|Z_{t_k} = j; \mathcal{F}_{t_{k-1}}; \hat{\Theta}_1^{(n)}\Big)} \end{aligned} \tag{15}$$

such that

$$
\begin{aligned}
P\Big(Z_{t_k} = i|\mathcal{F}^X_{t_{k-1}};\hat{\Theta}^{(n)}_1\Big) &= \sum_{j\in S} P\Big(Z_{t_k} = i, Z_{t_{k-1}} = j|\mathcal{F}^X_{t_{k-1}};\hat{\Theta}^{(n)}_1\Big) \\
&= \sum_{j\in S} P\Big(Z_{t_k} = i, Z_{t_{k-1}} = j|\hat{\Theta}^{(n)}_1\Big) P\Big(Z_{t_{k-1}} = j|\mathcal{F}^X_{t_{k-1}};\hat{\Theta}^{(n)}_1\Big) \\
&= \sum_{j\in S} \Pi^{(n)}_{ij} P\Big(Z_{t_{k-1}} = j|\mathcal{F}^X_{t_{k-1}};\hat{\Theta}^{(n)}_1\Big),
\end{aligned} \tag{16}
$$

where $f\Big(X_{t_k}|Z_{t_k} = i;\mathcal{F}_{t_{k-1}};\hat{\Theta}^{(n)}_1\Big)$ is the density of the process X at time t_k, conditional that the process is in regime $i \in S$. Using model 10 we can observe that, given $\mathcal{F}^X_{t_{k-1}}$, the process X_{t_k} has a conditional Gaussian distribution $\sim N\Big(k^{(n)}_i\theta^{(n)}_i + (1-k^{(n)}_i)X_{t_{k-1}}, \sigma^{2(n)}_i\Big)$. The density of this distribution is given by

$$
f\Big(X_{t_k}|Z_{t_k} = i;\mathcal{F}_{t_{k-1}};\hat{\Theta}^{(n)}_1\Big) = \frac{1}{\sqrt{2\pi\sigma^{(n)}_i}} \exp\Big[\frac{X_{t_k} - \big(1-k^{(n)}_i\big)X_{t_{k-1}} - \theta^{(n)}_i k^{(n)}_i\big)^2}{2\big(\sigma^{(n)}_i\big)^2}\Big] \tag{17}
$$

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 $\mathcal{F}^X_T$. For all $i \in S$ and $k = \{M-1, M-2, \ldots, 1\}$ we obtain

$$
P\Big(Z_{t_k} = i|\mathcal{F}^X_{t_M};\hat{\Theta}^{(n)}_1\Big) = \sum_{j\in S}\Big(\frac{P\big(Z_{t_k} = i|\mathcal{F}^X_{t_k};\hat{\Theta}^{(n)}_1\big)P\big(Z_{t_{k+1}} = j|\mathcal{F}_{t_M};\hat{\Theta}^{(n)}_1|\Pi^{(n)}_{ij}\big)}{P\big(Z_{t_{k+1}} = j|\mathcal{F}^X_{t_k};\hat{\Theta}^{(n)}_1\big)}\Big) \tag{18}
$$

9.2.2 Maximization step (M step)

We are able to obtain explicit formula of the maximum likelihood estimator of the initial subset of parameters $\hat{\Theta}_1$. The maximum likelihood estimates $\hat{\Theta}^{(n+1)}_1$ for all parameters, for all $i \in S$, can be obtained by

$$
\begin{aligned}
\theta^{(n+1)}_i &= \frac{\sum_{k=2}^M \big[P\big(Z_{t_k} = i|\mathcal{F}_{t_M};\hat{\Theta}^{(n)}_1\big)\big(X_{t_k} - (1-k^{(n+1)}_i)X_{t_{k-1}}\big]}{k^{n+1}_i \sum_{k=2}^M \big[P\big(Z_{t_k} = i|\mathcal{F}_{t_M};\hat{\Theta}^{(n)}_1\big)\big]} \\
k^{(n+1)}_i &= \frac{\sum_{k=2}^M \big[P\big(Z_{t_k} = i|\mathcal{F}_{t_M};\hat{\Theta}^{(n)}_1\big)X_{t_{k-1}}B_1\big]}{\sum_{k=2}^M \big[P\big(Z_{t_k} = i|\mathcal{F}_{t_M};\hat{\Theta}^{(n)}_1\big)X_{t_{k-1}}B_2\big]} \\
\sigma^{(n+1)}_i &= \frac{\sum_{k=2}^M \big[P\big(Z_{t_k} = i|\mathcal{F}_{t_M};\hat{\Theta}^{(n)}_1\big)\big(X_{t_k} - k^{(n+1)}_i\theta^{(n+1)}_i(1-k^{(n+1)}_i)X_{t_{k-1}}\big)^2\big]}{\sum_{k=2}^M \big[P\big(Z_{t_k} = i|\mathcal{F}_{t_M};\hat{\Theta}^{(n)}_1\big)\big]}
\end{aligned} \tag{19}
$$

where

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

We then obtain the transition probabilities:

$$
\Pi^{(n+1)}_{ij} = \frac{\sum_{k=2}^M \Big[P\big(Z_{t_k} = j|\mathcal{F}_{t_M};\hat{\Theta}^{(n)}_1\big)\frac{\Pi^n_{ij}P\big(Z_{t_{k-1}}=i|\mathcal{F}_{t_{k-1}};\hat{\Theta}^{(n)}_1\big)}{P\big(Z_{t_k}=j|\mathcal{F}_{t_{k-1}};\hat{\Theta}^{(n)}_1\big)}\Big]}{\sum_{k=2}^M \big[P\big(Z_{t_{k-1}} = i|\mathcal{F}_{t_{k-1}};\hat{\Theta}^{(n)}_1\big)\big]} \tag{20}
$$

3. 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 $n+1 \leq N$ 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 [56].

9.3 Stage 2: Lévy distribution fitted to each regime

We have estimated the regime-switching model 6 using the EM algorithm. Now, we estimate the set of parameters $\hat{\Theta_2}$ by fitting a NIG distribution for each regime.

$$X(\text{Regime } 1) - L_1(\alpha^1, \beta^1, \delta^1, \mu^1) \tag{21}$$

$$X(\text{Regime } 2) - L_2(\alpha^2, \beta^2, \delta^2, \mu^2) \tag{22}$$

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 [18], initialization of the algorithm is performed by the method of moments.

References

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

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

[3] Ang, A., Timmermann, A. (2012) Regime Changes and Financial Markets, *Annual Review of Financial Economics*, 4, 313–337.

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

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

[6] 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.

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

[8] 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.

[9] 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.

[10] Bates, D. S. (1991). The crash of 87: Was it expected? The evidence from options markets. *J. Finance* 46 1009-1044.

[11] Bekaert, G., Hodrick R., (1993) On Biases in the Measurement of Foreign Exchange Risk Premium. *Journal of International Money and Finance* 12, 115-138.

[12] Bianchi, F. (2013), Regime Switches, Agents' Beliefs, and Post-World War II U.S. Macroeconomic Dynamics, *Review of Economic Studies*, 80, 463–490.

[13] Billio, M., Pelizzon. L., (2000) Value-at-Risk: a Multivariate Switching Regime Approach. *Journal of Empirical Finance*, 7, 531-554.

[14] Board of the Governors of the Federal Reserve System. 2007. FRB: Press Release FOMC Statement: The Federal Reserve Is Providing Liquidity to Facilitate the Orderly Functioning of Financial Markets. August 10, 2007.

[15] Brooks, C., Prokopczuk, M., (2013) The Dynamics of Commodity Prices. *Quantitative Finance* 13 (4): 527-542.

[16] Carr, P., Geman H., Madan, D. B., Yor, M., (2003) Stochastic Volatility for Lévy Processes. *Mathematical Finance* 13: 345-382.

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

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

[19] D'Agostino, R. B., (1986) Tests for the normal distribution. In R. B. D'Agostino and M. A. Stephens, eds., *Goodness of Fit Techniques*. Marcel Dekker, New York, 367-419.

[20] Deaton, A., Laroque, G., (1992) On the Behaviour of Commodity Prices. *Review of Economic Studies* 59: 1-23.

[21] 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.

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

[23] Driffill, J., Sola, M., (1998), Intrinsic Bubbles and Regime Switching. *Journal of Monetary Economics* 42, 357-373.

[24] Driffill, J., T. Kenc, Sola M., Spagnolo F., (2004) An Empirical Examination of Term Structure Models with Regime Shifts, Discussion paper, *Centre for Economic Policy Research.*

[25] Engle, R., Hamilton, JD., (1990) Long Swings in the Dollar: Are They in the Data and Do Markets Know It? *American Economic Review* 80: 689-713.

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

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

[28] Guidolin, M., and A., Timmermann (2008). Size and Value Anomalies Under Regime Shifts. *Journal of Financial Econometrics*, 6, 1-48.

[29] Hainaut D. (2010) Switching Lévy processes : a toolbox for financial applications. Working paper.

[30] 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.

[31] Hamilton JD., (2005), What's Real About the Business Cycle? *Federal Reserve Bank of St. Louis Review*, 87, 435–452.

[32] Hamilton JD., (2016), Macroeconomic Regimes and Regime Shifts, in *Handbook of Macroeconomics*, eds. Taylor, J. B. and Uhlig, H. Elsevier, vol. 2, chap. 3.

[33] Hamilton JD., 1989 Rational-Expectations Econometric Analysis of Changes in Regime. *Journal of Economic Dynamics and Control* 12: 385-423.

[34] Hamilton JD., Susmel, R. (1994), Autoregressive Conditional Heteroskedasticity and Changes in Regime. *Journal of Econometrics* 64: 307-333.

[35] Huisman, R., Mahieu, R., (2003) Regime Jumps in Electricity Prices. *Energy Economics* 25 (5): 425-434.

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

[37] Kaeck, A. (2013) Asymmetry in the Jump-Size Distribution of the S&P 500: Evidence from Equity and Option Markets. *Journal of Economic Dynamics and Control* 37 (9): 1872-1888.

[38] Kahn, J.A., Rich, R.W., (2007) Tracking the New Economy: Using Growth Theory to Detect Changes in Trend Productivity, *Journal of Monetary Economics*, 54, 1670–1701.

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

[40] Karlis D., (2002) An EM type algorithm for maximum likelihood estimation of the normal-inverse Gaussian distribution. *Statistics & probability letters*, 57(1):43–52.

[41] Kasahara, H., Shimotsu, K., (2018) Testing the number of regimes in Markov regime switching models, working paper, arXiv:1801.06862

[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] Madan, D. B., Seneta, E., (1990) The VG Model for Share Market Returns. *Journal of Business* 36: 364-419.

[46] Mantalos, P., (2011) The three different measures of the sample skewness and kurtosis and the effects to the Jarque-Bera test for normality. *International Journal of Computational Economics and Econometrics*, 2, 47-62.

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

[48] Monfort, A. Feron, O., (2012) Joint Econometric modelling of Spot Electricity Prices, Forwards and Options. *Review of Derivatives Research*, 15, 217-256.

[49] Morley, J., Piger, J., (2012) The Asymmetric Business Cycle, *Review of Economics and Statisitcs*, 94, 208–221.

[50] Okimoto, T., (2008) New Evidence of Asymmetric Dependence Structures in International Equity Markets, *Journal of Financial and Quantitative Analysis*, 43, 787–815.

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

[52] Prause, K. (1999a) The generalized hyperbolic model: financial derivatives and risk measures. In Eleventh European Young Statisticians Meeting, Marly-le-Roi, August 24-28, 1999, pp. 165-169. Bernoulli Society.

[53] Prause, K. (1999b) How to use NIG laws to measure market risk. FDM Preprint 65, University of Freiburg.

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

[57] Schaefer, S. M. and Strebulaev, I. A. (2008) Structural Models of Credit Risk are Useful: Evidence from Hedge Ratios on Corporate Bonds, *J. of Financial Economics* 90(1), 1–19.

[58] Schorfheide, F. (2005), Learning and Monetary Policy Shifts, *Review of Economic Dynamics*, 8, 392–419.

[59] Sims, C., Zha, T. (2006), Were There Regime Switches in U.S. Monetary Policy? *American Economic Review*, 96, 54–81.

[60] 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.

[61] Thadewald, T, and H. Buning, (2004) Jarque-Bera test and its competitors for testing normality - A power comparison. Discussion Paper Economics 2004/9, School of Business and Economics, Free University of Berlin.

[62] Tu, J. (2010) Is Regime Switching in Stock Returns Important in Portfolio Decisions? *Management Science* 56 (7): 1198-1215.

[63] Urzua, C. M., (1996) On the correct use of omnibus tests for normality. *Economics Letters*, 53, 247-251.

Appendix

Figure 3: Regime discontinuities in OASs (top to bottom): AAA, AA, A.

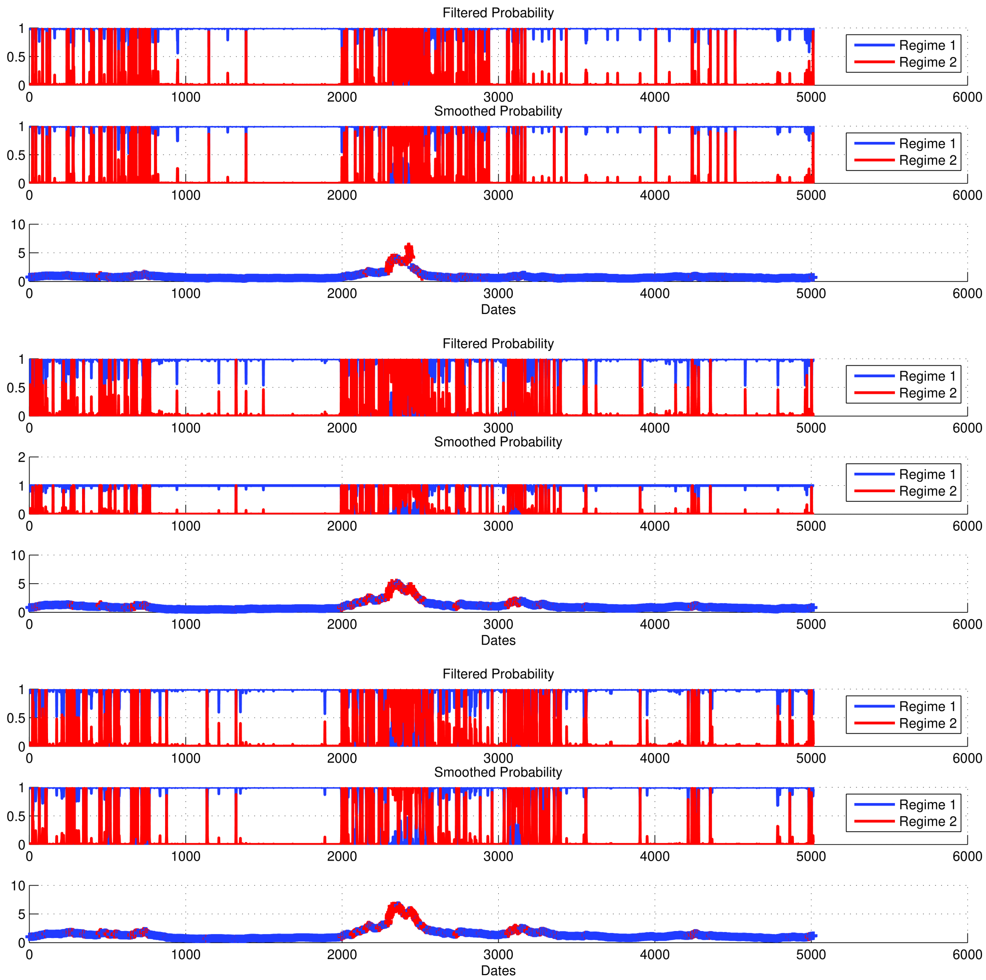


Figure 4: Regime discontinuities in OASs (top to bottom): BBB, BB, B

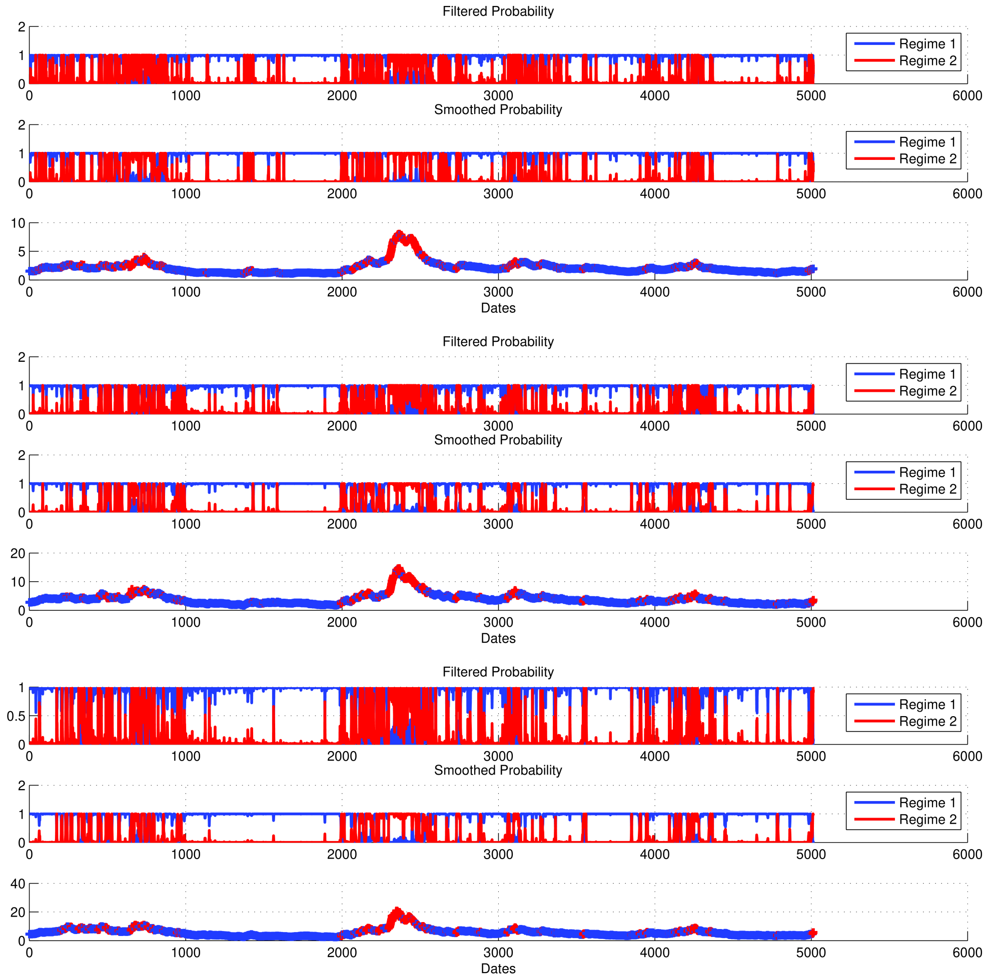


Figure 5: Regime discontinuities in OASs: CCC.

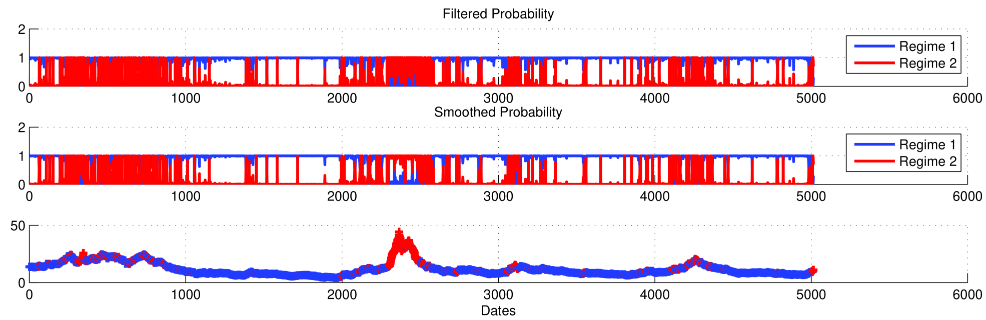


Figure 6: Regime discontinuities in OASs (top to bottom): US Corporate 7-10 Year, 3-5 Year, and 1-3 Year.

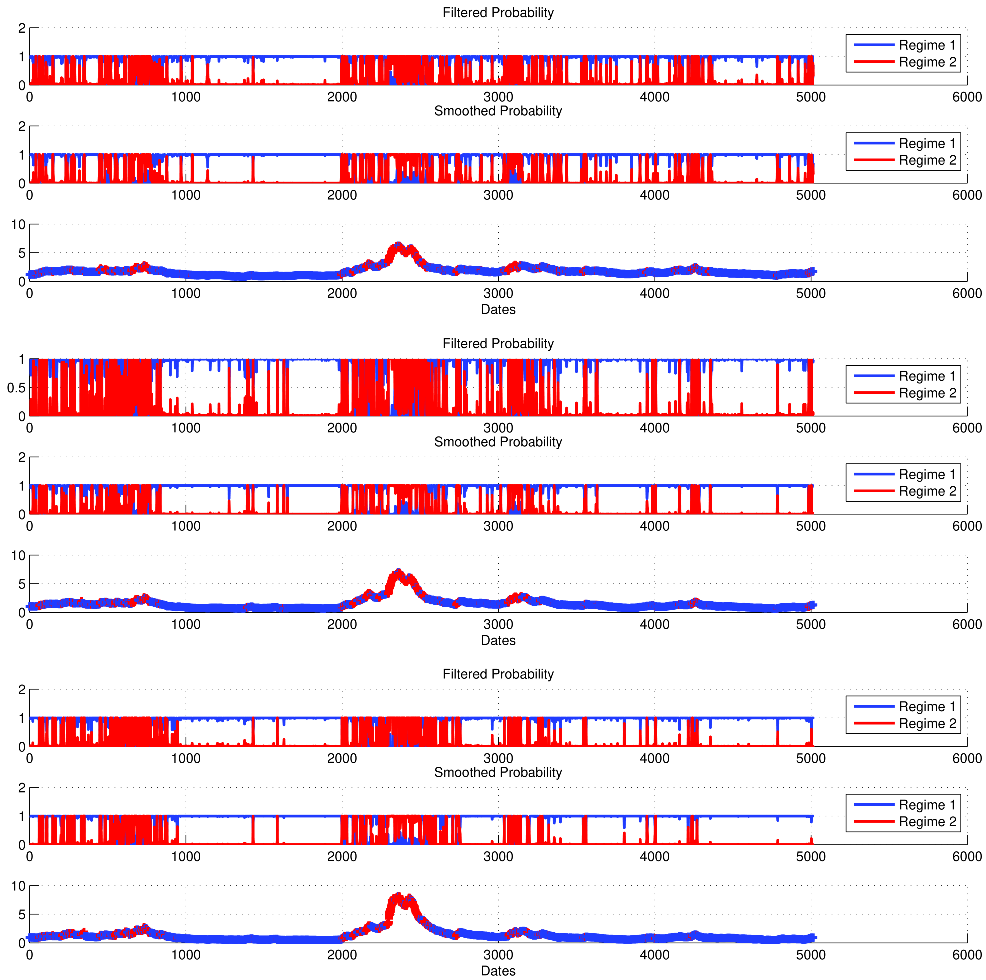


Figure 7: CBOE Volatility Index: VIX.