Time-Varying Jump Tails with Lévy Processes

Tidsafhængig Spring Intensitet med Lévy Processer

Marcus Frank Olesen (S111218)
Master's Thesis

MSc in Business Adminstration and Mathematical Business Economics

Copenhagen Business School

May 17, 2021

Supervisor: Professor Anders Bjerre Trolle
Number of Pages: 80 Number of Characters: 139,937

Resumé

Formålet med denne kandidatafhandling er at undersøge dynamikken for tidsafhængige risikoneutrale spring i afkastet på finansielle produkter gennem minimale antagelser om strukturen i disse spring. Fokuset er på haleelementet af intensiteten for spring under $\mathbb{Q}$.

Først er der en introduktion til Lévy processer og nogle af deres vigtigste egenskaber, samt hvordan de kan karakteriseres. Fokus er på at forstå Lévy-Itôs berømte sætning og især hvilke dynamikker, der bliver udvist af de to uafhængige dele med spring. Som led i dette bliver sammenhængen mellem stokastiske Poisson mål og Lévy mål forklaret, hvilket er essentielt, da det antages, at dynamikkerne for halerne bliver fuldt karakteriseret af et Lévy mål.

Efterfølgende er der en gennemgang af den model, der vil blive anvendt i denne afhandling. Først gennemgås hvilke parametre, der vil blive anvendt til at vurdere halen for intensiteten af spring. Desuden vises, at netop variation i den venstre hale kan bruges som en proxy for den frygt, der er i markedet på et givent tidspunkt. Herefter følger en gennemgang af hvilke egenskaber, der er ønskelige for de parametre, der karaktiserer dynamikken i den risiko-neutrale hale samt et teoretisk fokus på hvorfor, det er essentielt, at halerne udviser tidsafhængig dynamik. Derpå gennemgås det, hvordan estimatorerne for de førnævnte parametre opnås ud fra data på optioner. Gennemgangen af estimatorerne starter med tilfældet, hvor det antages, at halerne ikke er tidsafhængige, hvorefter der generaliseres til tilfældet med tidsafhængige haler, som vil blive benyttet i resten af afhandlingen.

Dette leder til en gennemgang af de empiriske resultater baseret på tidsperioden januar 1996 til og med december 2020, hvorunder især haleformen er et omdrejningspunkt, da det vises, at denne indeholder størstedelen af informationen. Ved hjælp af metodik fra tidsserier modelleres dynamikken for halens form. I denne process er der fokus på at fjerne periodicitet og trend i estimatet for derefter at modellere middelværdien med en ARIMA model og variansen med en GARCH model, så den kombinerede model er stationær.

Afslutningsvis gennemgås det, hvordan formen på halen, variansen på venstrespring og VIX indekset kan bruges til prædiktion af fremtidige afkast ved hjælp af univariat og multivariat regression. Dette undersøges både for den aggregerede markedsportefølje og for yderligere fem porteføljer, der er sorteret henholdsvis efter Fama-French tre-faktor model, momentum, betting-against-beta og quality-minus-junk. Ved at sammenligne regressionerne fra VIX og variansen på venstrespring, fremgår det, at markedet behandler variansen på store negative spring særligt.

Contents

1 Introduction

The return of financial assets has always been at the very core of economic history, where it has undergone significant development from a case-by-case comparison to modern stochastic models. Perhaps the biggest steppingstone in this development was the famous Black-Scholes-Merton model for the pricing of European options. Firstly, published by Fischer Black and Myron Scholes in Black & Scholes (1973), where Robert C. Merton expanded upon it with Merton (1973). With a couple of somewhat reasonable assumptions, the model could price options consistently most of the time.

Nonetheless, the application of inappropriate assumptions about the market and the dynamics of assets can be detrimental to pricing and risk management. One of the most classic examples was the collapse of Long-Term Capital Management, see Edward (1999), where one of the most prominent hedge funds, with partners such as Robert C. Merton and Myron Scholes, collapsed under the effects of the 1998 Russian financial crisis partly due to high leverage, which helps to show the great importance of tail risk management.

However, the Black-Scholes-Merton model's assumption had proved itself faulty earlier than the collapse of Long-Term Capital management. Under the market crash in 1987 and with the appearance of volatility smiles, it became evident that the assumption of constant volatility was detrimental. These volatility smiles are very pronounced for short maturity options, which leads to the fact that financial asset returns are not conditionally normally distributed, but instead exhibit decaying tails, which are fatter than expected from a normal distribution. These are attributable to infrequent significant price changes, which will be modelled through Lévy Processes, and are very clear in periods with financial distress such as the financial crisis and the Covid-19 crisis, where the latter will be a focal point in this dissertation.

The choice of dynamics of asset prices is of great importance, and the most common process used in modelling in finance is the Brownian motion, which was also one of the first proposed uses of the Brownian motion. A Danish astronomer first described the mathematics of the Brownian motion in Thiele (1880). However, he is commonly not credited for being the first person to model the stochastic process, which instead goes to Bachelier (1900), where he employed it to value stock options and was the most advanced use of mathematics in finance at that time. His work builds upon the hypotheses in Regnault (1863), which is often credited as being the first paper to use a random walk model to model price changes, where Regnault observed that the standard deviation of a price change over a time interval scaled with the square root of the length of the interval. The Brownian motion has proven itself extremely useful at describing the continuous part of price changes, but empirical data has proven that the tails of a normal distribution are not thick enough to describe the return of assets, which is where Lévy processes came into use. The foundation of Lévy processes can be found in the foundational works of Bachelier (1900) & Bachelier (1901), concerning the use of Brownian motion in financial mathematics, and Lundberg (1903), concerning the use of the Poisson process within the context of insurance mathematics. By combining continuous and jump processes under a common process, Lévy processes have become crucial in risk management, which fittingly is the core of financial mathematics and insurance mathematics.

As tails by definition are related to infrequent episodes, it is natural to use the tail distribution to construct a proxy for market fear. A method for this was proposed in Bollerslev et al. (2015), and I will review how this fear component developed during 2020 and how quickly it reacted to macro- and fiscal policy changes.

In the discussion of market fear, it is natural to include the *VIX*. The *VIX* is representing the market's expectations for volatility over the coming 30 days. It is also naturally related to the method used in this paper as both the *VIX* and the method used here are based on the price of *SPX* options close to expiration. However, a significant difference is that I will only focus on deep out-of-the-money options and discard options that are at-the-money or close by. This limit will vary over time, related to the implied volatility in the market and will be described in detail in Section 4. A comparison of the predictability of returns on the aggregate market portfolio between left jump tail variation and the VIX shows that the left jump tail variation has a stronger predictability and is a cleaner proxy for fear.

1.1 Thesis statement

The field of investigation for this thesis is partly arguing the need for time-dependent jump tails and partly showing that the left risk-neutral jump tail variation helps predict future market returns, indicating that investors demand special compensation for bearing jump tail risk. This leads to the following thesis statement:

How can one model time-dependent jump tails, both from a theoretical standpoint and through a time series model for the tail shape, and how can the estimates be used to show that the market has a special treatment for bearing *the risk of large jumps. Reviewed on the index S&P 500 and the U.S. aggregate market portfolio.*

To answer the thesis statement, the focal point is to answer the following sub-questions:

- Define a Lévy process and which main attributes they possess.
- Describe the main differences between the two elements in the Lévy-Itô decomposition that concern minor and large jumps.
- Derive that the left jump variation measure can be used as a proxy for fear.
- Assess the implications of the limit for what constitutes a significant jump.
- Assess the need for time-variant tail shapes from a theoretical and empirical standpoint.
- Derive an ARIMA/GARCH model for the time-variant tail shapes.
- Discuss whether the left jump variation or the VIX is a stronger predictor for market returns, and examine the implications on the fear proxy.

1.2 Methodology

In order to investigate the specified thesis statement, the thesis has a dual construction consisting of a theoretical and empirical part, respectively. The theoretical part is concerned with the underlying theory about Lévy measures and jumps and how these can be applied to construct our estimators. The empirical part is concerned with estimation and performing predictive analysis on the estimates.

The dominant methodology throughout the thesis is logical positivism with its core idea that reality exists independent from our realisation and the idea that scientific work is centred around verifying and confirming theories from an empirical basis. Being a realistic ontology, it is a natural choice, as this thesis attempts to map and model theories based on observable data in reality. Naturally, it is assumed that the data set employed is complete and representative of the U.S. market. If this assumption proves faulty, then the conclusions herein should be reconsidered. The evaluation of the fear proxies included in this thesis is based on how well each model predicts future returns. The evaluation is based solely on objective values observable in the market, a classical positivist trait in order that it remains free of subjective values.

2 Introduction to Lévy processes

In preparation for estimating the relevant jump risk premia and tail risk premia, it is crucial to describe the theory used in constructing the models before advancing to the estimation of these. This chapter will give an introduction to certain core Lévy processes and their use in this case. The following section has been based around Kyprianou (2014), Papapantoleon (2008), private lecture notes for the course ST426: Applied Stochastic Processes at London School of Economics, and own results.

2.1 Introduction to Lévy processes and definition

Lévy processes are now playing a central role in several fields of science that are of importance in this article, from the use in continuous time-series models in economics; for the calculation of insurance and re-insurance risk in actuarial science; and, of course, in mathematical finance.

For the definition of the Lévy process, it is natural to start with two familiar processes, which are both special cases of a Lévy process, a Brownian motion and a Poisson process. Let $(\Omega, \mathcal{F}, \mathbf{F}, \mathbb{P})$ denote a stochastic basis with filtration $\mathbf{F} = (\mathcal{F}_t)_{t \geq 0}$.

Then a real-valued process, $B = \{B_t : t \geq 0\}$, defined on said probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is said to be a Brownian motion if the following hold:

i. The paths of B are $\mathbb{P}$-almost surely continuous.

ii. $\mathbb{P}(B_0 = 0) = 1$.

iii. For $0 \leq s \leq t, B_t - B_s$ is equal in distribution to B_{t-s}.

iv. For $0 \leq s \leq t, B_t - B_s$ is independent of $\{B_u : u \leq s\}$.

v. For each $t > 0$, B_t is equal in distribution to a normal random variable with zero mean and variance t.

Where the third requirement is known as *stationary increments* and the fourth requirement is known as *independent increments*, which both are two of the requirements for a Lévy process.

The other side of Lévy processes are their jump part, and as such it is natural to start with the Poisson process.

A process valued on the non-negative integers, $N = \{N_t : t \geq 0\}$, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, is said to be a Poisson process with intensity $\lambda > 0$ if the following hold:

i. The paths of N are $\mathbb{P}$-almost surely right-continuous with left limits.

ii. $\mathbb{P}(B_0 = 0) = 1$.

iii. For $0 \leq s \leq t, N_t - N_s$ is equal in distribution to N_{t-s}.

iv. For $0 \leq s \leq t, N_t - N_s$ is independent of $\{N_u : u \leq s\}$.

v. For each $t > 0, N_t$ is equal in distribution to a Poisson random variable with parameter λt.

It is evident that despite their substantial differences, after all, one process is continuous with unbounded variation over finite time horizons, and the other is a non-decreasing jump process with bounded variation over finite time horizons; there are many similarities in their definitions. Using these common properties, it is possible to define a general class of one-dimensional stochastic processes called Lévy processes.

Definition 2.1 (Lévy Process)**.** A process $X = \{X_t : t \geq 0\}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, is said to be a Lévy process if it posses the following properties:

i. The paths of X are $\mathbb{P}$-almost surely right-continuous with left limits. This is also known as càdlàg paths.

ii. $\mathbb{P}(X_0 = 0) = 1$.

iii. For $0 \leq s \leq t, X_t - X_s$ is equal in distribution to X_{t-s}.

iv. For $0 \leq s \leq t, X_t - X_s$ is independent of $\{X_u : u \leq s\}$.

2.2 Infinitely divisible distributions and the Lévy-Khintchine formula

However, the simplicity of Definition 2.1 can be deceiving in showing just how rich the class of Lévy processes is. The mathematician de Finetti (1929) introduced the concept of infinitely divisible distributions and showed their relationship to Lévy processes.

Definition 2.2 (Infinite Divisibility)**.** A random variable X is infinitly divisible if, for all $n \in \mathbb{N}$, there exist i.i.d. random variables $X_1^{(n)}, ..., X_n^{(n)}$ such that

$$X \overset{d}{=} X_1^{(n)} + ... + X_n^{(n)}$$

Which can also be generalised to the probability measure

Definition 2.3. A probability measure ρ is infinitely divisible if, for all $n \in \mathbb{N}$, there exists another probability measure ρ_n such that

$$\rho = \rho_n * ... * \rho_n$$

To understand the deep connection between infinitely divisible distributions and Lévy processes, it is natural to prove that all Lévy processes have infinitely divisible laws.

Lemma 2.1. *Let $X = (X_t)_{t \geq 0}$ be a Lévy process. Then the random variable $X_t, t \geq 0$, are infinitely divisible.*

Proof. Let $X = (X_t)_{t \geq 0}$ be a Lévy process. For any $n \in \mathbb{N}$ and any $t > 0$ X_t can be rewritten as

$$X_t = X_{\frac{t}{n}} + \left(X_{\frac{2t}{n}} - X_{\frac{t}{n}}\right) + ... + \left(X_t - X_{\frac{(n-1)t}{n}}\right) \tag{1}$$

as all the terms cancel out except from X_t. By recalling Definition 2.1 it is given that increments of a Lévy process are stationary. By using this definition it gives

$$X_{\frac{tk}{n}} - X_{\frac{(k-1)t}{n}} \overset{d}{=} X_{\frac{t}{n}}$$

for any $k \geq 1$. From the independence of the increments it yields that the random variable $X_{\frac{tk}{n}} - X_{\frac{(k-1)t}{n}}, k \geq 1$ are independent of each other. From this it gives that each parentheses in (1), $\left(X_{\frac{tk}{n}} - X_{\frac{(k-1)t}{n}}\right)_{k \geq 1}$ is an i.i.d. sequence of random variables, and by Definition 2.2 it can be concluded that X_t is infinitely divisible, therefore a general Lévy process is infinitely divisible. $\square$

Before progressing on to the Lévy-Khintchine theorem, the notation of characteristic functions and characteristic exponents used in this thesis will be denoted to sort out the ambiguity regarding the negative exponent present in the literature. Denote the characterising function by φ, its law by P_X, and its moment generating function by M_X, hence

$$\varphi_X(u) = \int_{\mathbb{R}} e^{iux} P_X(dx) = M_X(iu).$$

The characteristic exponent ψ for a real-valued random variable is given by

$$\int_{\mathbb{R}} e^{iux} P_X(dx) = e^{\psi(u)}.$$

Lemma 2.2. *Let X_t be a random variable with infinitely divisible law. Then it holds true that $\psi_t(u) = t\psi_1(u)$.*

Proof. Let X_t be a random variable with infinitely divisible law, and as such, it is a Lévy process. Then by (1) characteristic function can be expanded to the characteristic function of the divisible form. By assuming that m, n are any two positive integers, it is given that:

$$\begin{aligned}
\psi_m(u) &= \log E(e^{iuX_m}) \\
&= \log E\left[e^{iu\left(X_{\frac{m}{n}} + \left(X_{\frac{2m}{n}} - X_{\frac{m}{n}}\right) + \ldots + \left(X_m - X_{\frac{(n-1)m}{n}}\right)\right)}\right] \\
&= \log\left(E\left[e^{iuX_{\frac{m}{n}}}\right] E\left[e^{iu\left(X_{\frac{2m}{n}} - X_{\frac{m}{n}}\right)}\right] \ldots E\left[e^{iu\left(X_m - X_{\frac{(n-1)m}{n}}\right)}\right]\right) \\
&= \log\left(E\left[e^{iuX_{\frac{m}{n}}}\right] \ldots E\left[e^{iuX_{\frac{m}{n}}}\right]\right) \\
&= n\psi_{\frac{m}{n}}(u) = m\psi_1(u)
\end{aligned}$$

Where the second equals come from inserting the divided form, the third and fourth comes from the fact that each parenthesis is i.i.d. random variables, and the last equals come from assuming that $n = m$.

As such it holds for any rational $t > 0$,

$$\psi_t(u) = t\psi_1(u). \tag{2}$$

If, however, t is not rational then m, n cannot be selected to construct t. In this case a decreasing sequence of rationals $\{t_n : n \geq 1\}$ such that $t_n \downarrow t$ as $n \to \infty$.

As it is obviously dominated by the initial t_1, as it is a decreasing sequence, the dominated convergence theorem gives that the almost sure right-continuity of X, which is given by càdlàg paths in the Definition 2.1, implies right-continuity of $e^{\psi_t(u)}$ and hence (2) holds true for all $t \geq 0$. □

The following result provides a complete characterization of infinitely divisible distributions and links them to the concept of Lévy triplets. Paul Lévy and Aleksandr Khinchin both proved the result independently, and as such, the theorem is named the Lévyy–Khintchine theorem. Nevertheless, first, the Lévy measure will be described, as it is the function that describes the jump process, and it plays a crucial role in this thesis. Intuitively, the Lévy measure describes the expected number of jumps of a certain height in a time interval of length 1. As such, a large Lévy measure is expected in times of financial distress, when a Brownian motion cannot explain the market movements.

Definition 2.4 (Lévy measure)**.** Let ν be a measure on $\mathbb{R}$. ν is called a *Lévy measure* if it satisfies

$$\nu(\{0\}) = 0 \text{ and } \int_{\mathbb{R}} (|x|^2 \wedge 1)\nu(dx) < \infty$$

The Lévy measure has no mass at the origin, as this would indicate an expected number of jumps with size zero, and the mass away from the origin is finite. Thus only a finite number of large jumps can occur. Singularities can, nevertheless, occur around the origin.

And as such the theorem is stated:.

Theorem 2.1. *The law P_X of a random variable X is infinitely divisible iff there exists a triplet (b, c, ν), also known as the Lévy or characteristic triplet, with $b \in \mathbb{R}, c \in \mathbb{R}^+$ and the measure ν, such that*

$$E\left[e^{iuX}\right] = \exp\left[ibu - \frac{u^2 c}{2} + \int_{\mathbb{R}} (e^{iux} - 1 - iux 1_{\{|x|<1\}})\nu(dx)\right]$$

b is notated as the drift characteristic and c the Gaussian or diffusion characteristic.

The proof of this theorem is outside the scope of this paper, but the structure of the formula gives much intuition about the structure of Lévy processes, as it splits it up into a drift component, a Brownian component, and lastly, a jump component, which is again split up into smaller and larger jumps.

From Lemma 2.2 one can conclude the next results follows

Corollary 2.1. *The infinitely divisible random variable X_t has the Lévy triplet $(bt, ct, \nu t)$.*

On this basis I will now continue on the Lévy-Itô decomposition.

2.3 The Lévy-Itô Decomposition

Where the previous subsection was concerned with constructing the Lévy triplet $(bt, ct, \nu t)$ for an infinitely divisible random variable X_t this section will be focusing on the reverse path. Starting from a Lévy triplet (b, c, ν) one can construct a Lévy process $X = (X_t)_{t\geq 0}$. The Lévy-Itô decomposition accomplishes this by describing the structure of a general Lévy process in terms of four independent Lévy processes with their unique path behaviours. This leads to the theorem

Theorem 2.2. *Let ρ be an infinitely divisible distribution with Lévy triplet (b, c, ν), where $b \in \mathbb{R}, c \in \mathbb{R}_+$ and ν is a Lévy measure satisfying Definition 2.4. Then, there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which four independent Lévy processes exist, $X^{(1)}, ..., X^{(4)}$, where $X^{(1)}$ is a constant drift, $X^{(2)}$ is a Brownian motion, $X^{(3)}$ is a compound Poisson process and $X^{(4)}$ is a square integrable pure jump martingale with a almost surely countable number of jumps of magnitude less than 1 on each finite time interval. Setting $X = X^{(1)} + ... + X^{(4)}$ then that there exists a probability space on which a Lévy process $X = (X_t)_{t\geq 0}$ is defined, with characteristic exponent*

$$\psi(u) = iub - \frac{u^2 c}{2} + \int_{\mathbb{R}} \left(e^{iux} - 1 - iux1_{\{|x|<1\}}\right) \nu(dx)$$

for all $u \in \mathbb{R}$, and path, or Lévy-Itô, decomposition

$$X_t = bt + \sqrt{c}W_t + \int_0^t \int_{|x|\geq 1} x\mu^X(ds, dx) + \int_0^t \int_{|x|<1} x(\mu^X - \nu^X)(ds, dx)$$

where $\nu^X(ds, dx) = \nu(dx)ds$ and μ is the Poisson random measure and $\int_0^t \int_{|x|\geq 1} x\mu^X(ds, dx)$ is a compound Poisson random variable with intensity $t\nu(\mathbb{R} \setminus (-1, 1))$.

A random measure is a convenient tool at pooling sources of randomness into a single one. Consider a set $A \in \mathcal{B}(\mathbb{R} \setminus \{0\})$ such that $0 \notin \bar{A}$ and let $0 \leq t \leq T$; define the random measure of the jumps of the process X by

$$\begin{aligned} \mu^X(\omega; t, A) &= \#\{0 \leq s \leq t; \Delta X_s(\omega) \in A\} \\ &= \sum_{s\leq t} 1_A(\Delta X_s(\omega)) \end{aligned}$$

Hence the measure $\mu^X(\omega; t, A)$ counts the jumps of the process X of size in A up to time t. It satisfies the following properties.

$$\mu^X(t, A) - \mu^X(s, A) \in \sigma(\{X_u - X_v | s \leq v \leq u \leq t\})$$

hence $\mu^X(t, A) - \mu^X(s, A)$ is independent of $\mathcal{F}$ and as such has independent increments. It is also clear that $\mu^X(t, A) - \mu^X(s, A)$ equals the number of jumps in $X_{s+u} - X_s$ in A for $0 \leq u \leq t - s$; hence it can be conculded that the $\mu^X(\cdot, A)$ has stationary increments. Therefore, $\mu^X(\cdot, A)$ is a Poisson process with intensity $\nu(A) = E[\mu^X(1, A)]$ and μ^X is a Poisson random measure.

The full proof is outside the scope of this thesis, but an outline will be given with a focus on path variation. By splitting up the characteristic exponent into

$$\psi^{(1)}(u) = iub \tag{3}$$

$$\psi^{(2)}(u) = -\frac{u^2 c}{2} \tag{4}$$

$$\psi^{(3)}(u) = \int_{|x|\geq 1} (e^{iux} - 1)\nu(dx) \tag{5}$$

$$\psi^{(4)}(u) = \int_{|x|<1} (e^{iux} - 1 - iux)\nu(dx) \tag{6}$$

it can be controlled which known process they relate to by calculating the characteristic exponents. Let $X = bt$. Then the characteristic exponent is as follows $e^{t\psi(u)} = E[e^{iuX_t}] = E[e^{iubt}] = e^{iubt} = e^{t\psi(u)}$, where $\psi(u) = iub$. As such, the first part corresponds to a deterministic drift with parameter b.

Then let $X = \sqrt{c}W_t$, where W_t is a standard Brownian Motion. $E[e^{iuX_t}] = E[e^{iu\sqrt{c}W_t}] = e^{-\frac{1}{2}cu^2 t} = e^{t\psi(u)}$, where $\psi(u) = -\frac{u^2 c}{2}$, where it was used that $W_t \sim N(0, t)$. Therefore, the second corresponds to a Brownian motion with coefficient $\sqrt{c}$.

For the third term, let X_t be a compound Poisson process. Instead of finding this directly, the characteristic function of a compensated compound Poisson process will be found, as the characteristic function for the compound Poisson process will be found in the process, and the compensated compound Poisson process will be helpful for the fourth term. Let $X_t = \sum_{k=1}^{N_t} J_k - t\lambda\kappa$, where N is a Poisson process with parameter λ, so $E[N_t] = \lambda t$, J is an i.i.d. sequence of random variables with probability F, and $E[J] = \kappa < \infty$. Clearly, F describes with distribution of the jumps, which arrive according to the Poisson process.

$$\begin{aligned} E\left[e^{iuX_t}\right] &= E\left[\exp\left(iu\left(\sum_{k=1}^{N_t} J_k - t\lambda\kappa\right)\right)\right] \\ &= E\left[\exp\left(iu\sum_{k=1}^{N_t} J_k - iut\lambda\kappa\right)\right] \end{aligned} \tag{7}$$

To find the characteristic function of the compensated compound Process, I start with the compound Process. The characteristic function of the compound Poisson process can be found by conditioning on the number of jumps at a given time, using the tower rule, using independence and the moment generating function of a Poisson random variable.

$$
\begin{aligned}
E\left[\exp\left(iu\sum_{k=1}^{N_t} J_k\right)\right] &= E\left[E\left[\exp\left(iu\sum_{k=1}^{N_t} J_k\right)\middle| N_t = N_1\right]\right] \\
&= E\left[E\left[\exp\left(iuJ_1\right)\right]^{N_1}\right] \\
&= E[\varphi_J(u)^{N_1}], \varphi_J(u) = E\left[\exp\left(iuJ_1\right)\right] \\
&= E\left[e^{\log \varphi_J(u)^{N_1}}\right] \\
&= E\left[e^{N_1 \log \varphi_J(u)}\right] \\
&= M_{N_1}[\log \varphi_J(u)], M_{N_1}(t) = E[e^{tN_1}] = \exp(\lambda t(e^t - 1)) \\
&= \exp\left(\lambda t(e^{\log E[\exp(iuJ_1)]} - 1]\right) \\
&= \exp\left(\lambda t(E\left[\exp\left(iuJ_1\right) - 1\right]\right) && (8) \\
&= \exp\left(\lambda t \int_{\mathbb{R}} (e^{iux} - 1)F(dx)\right) && (9)
\end{aligned}
$$

The characteristic function of the compound Poisson process is now (9) and the compensated compound Poisson process can now be found by inserting (8) into (7), and by using $E[J_1] = \kappa$.

$$
\begin{aligned}
E\left[e^{iuX_t}\right] &= E\left[\exp\left(iu\sum_{k=1}^{N_t} J_k - iut\lambda\kappa\right)\right] \\
&= \exp\left(\lambda t(E\left[\exp\left(iuJ_1\right) - 1 - iuJ\right])\right) \\
&= \exp\left(\lambda t \int_{\mathbb{R}} (e^{iux} - 1 - iux)F(dx)\right) && (10)
\end{aligned}
$$

where the final step used that the distribution of J is F.

By setting the arrival rate in (9) equal to $\lambda := \nu(\mathbb{R} \setminus (-1,1))$ and the jump magnitude $F(dx) := \frac{\nu(dx)}{\nu(\mathbb{R}\setminus(-1,1))} 1_{\{|x|\geq 1\}}$ it becomes evident that this is equal to (5).

It seems that a natural choice for the process for $\psi^{(4)}(u)$ would be the compensated compound Poisson process. However, from Definition 2.4 it follows that

$$
\int_{\mathbb{R}} \min(1, x^2)\nu(dx) < \infty \Rightarrow \nu(\mathbb{R} \setminus [-1,1]) = \int_{\mathbb{R}\setminus[-1,1]} \nu(dx) < \infty
$$

this indicates that $\nu([-1,1])$ can be ∞, which is not allowed for a compound Poisson process. As such, there is a process that resembles a compensated compound Poisson process, but it is not precisely so. To get a deeper insight into the field of interest, one can split up

$$[-1,1] \setminus \{0\} = \left[-1,-\frac{1}{2}\right) \cup \left[-\frac{1}{2},-\frac{1}{4}\right) \cup ... \cup \left(\frac{1}{4},\frac{1}{2}\right] \cup \left(\frac{1}{2},1\right]$$

such that $[-1,1] \setminus \{0\} = \bigcup_{k=0}^{\infty} A_k, A_k = \{x : 2^{-(k+1)} < x \leq 2^{-k}\}$. This result will be used in the characteristic function from (6)

$$\exp\left(\int_{[-1,1]} (e^{iux} - 1 - iux)\nu(dx))\right) = \exp\left(\sum_{k=0}^{\infty} \int_{A_k} (e^{iux} - 1 - iux)\nu(dx))\right).$$

For this to hold it is important to claim, and show, that $\nu(A_k)$ is finite, otherwise this has not solved the previous issue.

Lemma 2.3. *For the set* $[-1,1]\setminus\{0\} = \bigcup_{k=0}^{\infty} A_k, A_k = \{x : 2^{-(k+1)} < x \leq 2^{-k}\}$ *it holds true that* $\nu(A_k)$ *is finite for any* $\{\nu : \int_{\mathbb{R}} \min(1,x^2)\nu(dx) < \infty\}$*.*

Proof. For any $0 < \varepsilon < 1 : \nu((\varepsilon,\infty)) < \infty$ because $\nu([1,\infty]) < \infty$ and

$$\nu((\varepsilon,1)) = \int_{\varepsilon}^{1} \nu(dx) = \frac{1}{\varepsilon^2}\int_{\varepsilon}^{1} \varepsilon^2\nu(dx) \leq \frac{1}{\varepsilon^2}\int_{\varepsilon}^{1} x^2\nu(dx) < \infty$$

A corresponding argument can be made for $\nu((-\infty,\varepsilon))$. Since A_k is a positive distance away from 0 for all k, it holds true that $\nu(A_k) < \infty \forall k$. □

Since $\nu(A_k)$ is finite, $X_t^{(4)}$ can be considered as an infinite sum of compensated compound Poisson processes, which can be shown to converge uniformly. This presents an issue as there are infinite sources of randomness, where the Poisson random measure comes in. In order to convert this informal result into a precise mathematical statement, it requires results on Poisson random measures and square integrable martingales, which are outside the scope of this thesis, but the main ideas of the proof are captured here. For a full result of the proof, the reader is referred to Kyprianou (2014).

3 Setup and model assumptions

The continuous-time dynamic no-arbitrage framework underlying the following empirical investigation is based on a minimal amount of assumptions about the structure and dynamics.

3.1 Notation and variance risk premium

The underlying asset price X_t is defined on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $(\mathcal{F}_t)_{t\geq 0}$ denotes the standard filtration. The instantaneous arithmetic return of X is assumed to have the following continuous representation

$$\frac{dX_t}{X_{t-}} = \alpha_t dt + \sigma_t dW_t + \int_{\mathbb{R}} (e^x - 1)\tilde{\mu}^{\mathbb{P}}(dt, dx) \tag{11}$$

where W_t is a Brownian motion, μ is a counting measure, as described in the previous section, for the jumps in X with compensator $dt \otimes \nu_t(dx)$, so that $\tilde{\mu}^{\mathbb{P}}(dt, dx) = \mu(dt, dx) - dt\nu_t^{\mathbb{P}}(dx)$ denotes the corresponding martingale measure under $\mathbb{P}$. Recall that $\mu([0,t], A) = \sum_{s\leq t} 1_{\{\log(\Delta X_s)\in A\}}$ for any measurable $A \in \mathbb{R}\setminus\{0\}$. It is assumed that the drift and volatility processes, α_t and σ_t, follows càdlàg paths, as described in Definition 2.1. The quadratic variation is found by splitting the function up into the continuous part and the pure jump part.

$$\begin{aligned} [X, X]_{[t,t+\tau]} &= \int_t^{t+\tau} \sigma_s^2 ds + \sum_{s\leq t+\tau} \Delta X_t^2 \\ &= \int_t^{t+\tau} \sigma_s^2 ds + \int_t^{t+\tau} \int_{\mathbb{R}} x^2 \mu(ds, dx) \end{aligned} \tag{12}$$

Where the volatility of the diffusive price increments, represented by the first term, is the variation due to small price moves. The second term measures the variation of the jumps through the measure μ that captures the amount of jumps in the process. Even though both terms contribute to the quadratic variation, they do so in drastically different ways. Diffusive risks can be hedged by a dynamic portfolio continuously controlling the exposure, and the same thing cannot be done to hedge the risk of jumps as they are in nature unpredictable.

I will assume that the risk-neutral probability measure $\mathbb{Q}$ exists, and that X takes the following dynamic under $\mathbb{Q}$,

$$\frac{dX_t}{X_{t-}} = (r_{f,t} - \delta_t)dt + \sigma_t dW_t^{\mathbb{Q}} + \int_{\mathbb{R}} (e^x - 1)\tilde{\mu}^{\mathbb{Q}}(dt, dx), \tag{13}$$

where $r_{f,t}$ and δ_t refer to the instantaneous risk-free rate, assumed to be the one month treasury bill rate, and the dividend yield of X_t. $W_t^{\mathbb{Q}}$ is a Brownian motion under $\mathbb{Q}$ and $\tilde{\mu}^{\mathbb{Q}}(dt,dx) = \mu(dt,dx) - dt\nu_t^{\mathbb{Q}}(dx)$ where again $dt \otimes \nu_t^{\mathbb{Q}}(dx)$ is the compensator of the jumps, this time under $\mathbb{Q}$.

The fact that a Lévy process drives the price process makes the market, in general, incomplete. As such, there exists a large set of equivalent martingale measures, but this will not be an issue in this thesis, as it is still possible to identify the tail uniquely.

For the jump compensator to be a valid jump compensator, or Lévy measure, $\nu_t^{\mathbb{Q}}(dx)$ must satisfy Definition 2.4,

$$\nu(\{0\}) = 0 \text{ and } \int_{\mathbb{R}} (x^2 \wedge 1)\nu(dx) < \infty \forall t \in \mathbb{R}_+.$$

A deeper decomposition of the variance risk premium (VRP) will now follow. The following definition will mirror the definition used in Bollerslev & Todorov (2011*b*). The variance risk premium on X is defined by,

$$VRP_{t,\tau} = \frac{1}{\tau}\left(E_t^{\mathbb{P}}(QV_{[t,t+\tau]}) - E_t^{\mathbb{Q}}(QV_{[t,t+\tau]})\right), \tag{14}$$

which correspond to the expected payoff on a (long) variance swap on the market portfolio, and historically it has been negative on average.

As the area of interest in this thesis is on jump and tail risk, it is possible to decompose the variance risk premium further into a continuous and jump part, as is shown in Bollerslev et al. (2015). The total continuous variation over $[t, t+\tau]$ was shown to be

$$CV_{[t,t+\tau]} = \int_t^{t+\tau} \sigma_s^2 ds,$$

which corresponds to the continuous part of (12). In the calculation of the total predictable jump variation under the $\mathbb{P}$ and $\mathbb{Q}$ probability measure it is important to recall that $\nu(A) = E[\mu(1, A)]$ as defined in Section 2.3. As such,

$$JV_{[t,t+\tau]}^{\mathbb{P}} = \int_t^{t+\tau}\int_{\mathbb{R}} x^2 v_s^{\mathbb{P}}(dx)ds, JV_{[t,t+\tau]}^{\mathbb{Q}} = \int_t^{t+\tau}\int_{\mathbb{R}} x^2 v_s^{\mathbb{Q}}(dx)ds. \tag{15}$$

By inserting these three terms into the variance risk premium it may be decomposed as,

$$
\begin{aligned}
VRP_{t,\tau} &= \frac{1}{\tau}\left(E_t^{\mathbb{P}}(CV_{[t,t+\tau]} + JV^{\mathbb{P}}_{[t,t+\tau]}) - E_t^{\mathbb{Q}}(CV_{[t,t+\tau]} + JV^{\mathbb{Q}}_{[t,t+\tau]})\right) \\
&= \frac{1}{\tau}\Bigg[\left(E_t^{\mathbb{P}}(CV_{[t,t+\tau]}) - E_t^{\mathbb{Q}}(CV_{[t,t+\tau]})\right) \\
&+ \left(E_t^{\mathbb{P}}(JV^{\mathbb{P}}_{[t,t+\tau]}) - E_t^{\mathbb{Q}}(JV^{\mathbb{P}}_{[t,t+\tau]})\right)\Bigg] \\
&+ \frac{1}{\tau}\left(E_t^{\mathbb{Q}}(JV^{\mathbb{P}}_{[t,t+\tau]}) - E_t^{\mathbb{Q}}(JV^{\mathbb{Q}}_{[t,t+\tau]})\right). \qquad (16)
\end{aligned}
$$

It might seem irrelevant to add and subtract the $E_t^{\mathbb{Q}}(JV^{\mathbb{P}}_{[t,t+\tau]})$ term, but it allows for a nice intuitive interpretation. The variance risk premium consists of the difference between the $\mathbb{P}$ and $\mathbb{Q}$ expectations of the continuous variation, the difference between the $\mathbb{P}$ and $\mathbb{Q}$ expectations of the same $\mathbb{P}$ jump variation, and the last is the difference between the $\mathbb{P}$ and $\mathbb{Q}$ jump variations under the same risk-neutral measure $\mathbb{Q}$.

As such, the first two terms is the variance risk premium accounts for the temporal variation of the jump intensity process, under the physical measure, and for the diffusive risk σ_t^2. For the market portfolio, this premium is caused by investors' willingness to hedge against changes in the investment opportunity set.

The last term is, however, different from this. It includes the difference between the $\mathbb{P}$ and $\mathbb{Q}$ jump variation measures under the same probability measure $\mathbb{Q}$ and stems from the fact that jumps may occur, and therefore it does not have a direct analogue for the diffusive price component. Moreover, as seen in Section 2.3 the jump measure includes both small and big jumps, which are inherently different. The inclusion of both small and big jumps poses an issue empirically, as our data is sampled discretely, and therefore one cannot distinguish between a minor jump or a continuous price change, and a time-variable limit will have to be placed for what is a small jump, and deemed a continuous change, and what is a significant jump.

3.2 Premium for tail risk

In the following part a focus will be placed on the premium related to tail risk and the methodology will follow Bollerslev et al. (2015). Apart from the separation of continuous variation and jump variation in the previous section, it is intuitive to also separate the jumps into negative and positive jumps, with the result that $\nu_t^{(\mathbb{Q},+)}((-\infty,0)) = 0, \nu_t^{(\mathbb{Q},-)}((0,\infty)) = 0$, and $\nu_t^{\mathbb{Q}} = \nu_t^{(\mathbb{Q},+)} + \nu_t^{(\mathbb{Q},-)}$.

Then the jump tail variation in (15) can also be separated into left tail jump variation and right tail jump variation over the interval $[t, t+\tau]$ by,

$$
\begin{aligned}
LJV^{\mathbb{Q}}_{[t,t+\tau]} &= \int_t^{t+\tau} \int_{x<-k_t} x^2 \nu_s^{\mathbb{Q}}(dx)ds \\
RJV^{\mathbb{Q}}_{[t,t+\tau]} &= \int_t^{t+\tau} \int_{x>k_t} x^2 \nu_s^{\mathbb{Q}}(dx)ds.
\end{aligned}
\tag{17}
$$

As mentioned previously, it becomes impossible to distinguish between change caused by a Brownian motion and a change caused by a jump on a discrete scale, and as such k_t is used as a time-varying cutoff that is related to the log-jump size and the Black-Scholes at-the-money implied volatility at time t.

Following the definition of the variance risk premium, based on the quadratic variation, in (14), it is natural that the left and right jump tail risk premia are defined in the same pattern by,

$$
\begin{aligned}
LJP_{t,\tau} &= \frac{1}{\tau}\left(E_t^{\mathbb{P}}(LJV^{\mathbb{P}}_{[t,t+\tau]}) - E_t^{\mathbb{Q}}(LJV^{\mathbb{Q}}_{[t,t+\tau]})\right) \\
RJP_{t,\tau} &= \frac{1}{\tau}\left(E_t^{\mathbb{P}}(RJV^{\mathbb{P}}_{[t,t+\tau]}) - E_t^{\mathbb{Q}}(RJV^{\mathbb{Q}}_{[t,t+\tau]})\right)
\end{aligned}
\tag{18}
$$

By subtracting this from the variance risk premium in (16) it should remove all premium related to jumps, and as such, can be interpreted as the part of the VRP attributable to continuous variation, or empirically to changes in $[-k_t, k_t]$. This can easily be checked by,

$$
\begin{aligned}
&VRP_{t,\tau} - (LJP_{t,\tau} + RJP_{t,\tau}) = \frac{1}{\tau}\left(E_t^{\mathbb{Q}}(JV^{\mathbb{P}}_{[t,t+\tau]}) - E_t^{\mathbb{Q}}(JV^{\mathbb{Q}}_{[t,t+\tau]})\right) \\
&+ \frac{1}{\tau}\left[\left(E_t^{\mathbb{P}}(CV_{[t,t+\tau]}) - E_t^{\mathbb{Q}}(CV_{[t,t+\tau]})\right) + \left(E_t^{\mathbb{P}}(JV^{\mathbb{P}}_{[t,t+\tau]}) - E_t^{\mathbb{Q}}(JV^{\mathbb{P}}_{[t,t+\tau]})\right)\right] \\
&- \frac{1}{\tau}\left[\left(E_t^{\mathbb{P}}(LJV^{\mathbb{P}}_{[t,t+\tau]}) - E_t^{\mathbb{Q}}(LJV^{\mathbb{Q}}_{[t,t+\tau]})\right) + \left((E_t^{\mathbb{P}}(RJV^{\mathbb{P}}_{[t,t+\tau]}) - E_t^{\mathbb{Q}}(RJV^{\mathbb{Q}}_{[t,t+\tau]})\right)\right] \\
&= \frac{1}{\tau}\left(E_t^{\mathbb{Q}}(JV^{\mathbb{Q}}_{[t,t+\tau]}) + E_t^{\mathbb{P}}(CV_{[t,t+\tau]}) - E_t^{\mathbb{Q}}(CV_{[t,t+\tau]}) - E_t^{\mathbb{Q}}(JV^{\mathbb{Q}}_{[t,t+\tau]})\right) \\
&= \frac{1}{\tau}\left(E_t^{\mathbb{P}}(CV_{[t,t+\tau]}) - E_t^{\mathbb{Q}}(CV_{[t,t+\tau]})\right),
\end{aligned}
$$

and the results are as expected. Again, by mimicking the work done on the variance risk premium in (16), the left and right tail premium may be decomposed further into terms of the physical and risk-neutral measure,

$$
\begin{aligned}
LJP_{t,\tau} = &\frac{1}{\tau}\left[E_t^{\mathbb{P}}(LJV_{[t,t+\tau]}^{\mathbb{P}}) - E_t^{\mathbb{Q}}(LJV_{[t,t+\tau]}^{\mathbb{P}})\right] \\
&+ \frac{1}{\tau}\left[E_t^{\mathbb{Q}}(LJV_{[t,t+\tau]}^{\mathbb{P}}) - E_t^{\mathbb{Q}}(LJV_{[t,t+\tau]}^{\mathbb{Q}})\right], \qquad (19)
\end{aligned}
$$

and correspondingly,

$$
\begin{aligned}
RJP_{t,\tau} = &\frac{1}{\tau}\left[E_t^{\mathbb{P}}(RJV_{[t,t+\tau]}^{\mathbb{P}}) - E_t^{\mathbb{Q}}(RJV_{[t,t+\tau]}^{\mathbb{P}})\right] \\
&+ \frac{1}{\tau}\left[E_t^{\mathbb{Q}}(RJV_{[t,t+\tau]}^{\mathbb{P}}) - E_t^{\mathbb{Q}}(RJV_{[t,t+\tau]}^{\mathbb{Q}})\right]. \qquad (20)
\end{aligned}
$$

Again, as with (16), the first term in both the LJP and RJP involves the difference between the physical and risk-neutral expectation of the same measure. This can be seen as a mirror to the first term in (16), where it was the difference between the physical and risk-neutral expectation of the future diffusive risk $CV_{[t,t+\tau]}$.
However, the second term is of more interest as it involves the difference between the risk-neutral expected values of the $\mathbb{P}$ and $\mathbb{Q}$ jump tail variation measures, which reflects the special treatment of jump tail risk.

In order to model the time-variant proxy for investors fears, proposed in Bollerslev & Todorov (2011*b*), $LJP_{t,\tau} - RJP_{t,\tau}$, it is helpful to assume that the distribution of large jumps are roughly symmetric under the $\mathbb{P}$ jump intensity process, such that $LJV_{[t,t+\tau]}^{\mathbb{P}} \approx RJV_{[t,t+\tau]}^{\mathbb{P}}$. Note that it is not assumed that $LJV_{[t,t+\tau]}^{\mathbb{Q}} \approx RJV_{[t,t+\tau]}^{\mathbb{Q}}$ holds true, as the market reacts differently to positive and negative jumps. Then the fear index approximately becomes,

$$
\begin{aligned}
LJP_{t,\tau} - RJP_{t,\tau} = &\frac{1}{\tau}\left[E_t^{\mathbb{P}}(LJV_{[t,t+\tau]}^{\mathbb{P}}) - E_t^{\mathbb{Q}}(LJV_{[t,t+\tau]}^{\mathbb{P}})\right] \\
&+ \frac{1}{\tau}\left[E_t^{\mathbb{Q}}(LJV_{[t,t+\tau]}^{\mathbb{P}}) - E_t^{\mathbb{Q}}(LJV_{[t,t+\tau]}^{\mathbb{Q}})\right] \\
&- \left(\frac{1}{\tau}\left[E_t^{\mathbb{P}}(RJV_{[t,t+\tau]}^{\mathbb{P}}) - E_t^{\mathbb{Q}}(RJV_{[t,t+\tau]}^{\mathbb{P}})\right]\right. \\
&\left.+ \frac{1}{\tau}\left[E_t^{\mathbb{Q}}(RJV_{[t,t+\tau]}^{\mathbb{P}}) - E_t^{\mathbb{Q}}(RJV_{[t,t+\tau]}^{\mathbb{Q}})\right]\right) \\
\approx &\frac{1}{\tau}\left[E_t^{\mathbb{Q}}(LJV_{[t,t+\tau]}^{\mathbb{P}}) - E_t^{\mathbb{Q}}(LJV_{[t,t+\tau]}^{\mathbb{Q}})\right] \\
&- \frac{1}{\tau}\left[E_t^{\mathbb{Q}}(RJV_{[t,t+\tau]}^{\mathbb{P}}) - E_t^{\mathbb{Q}}(RJV_{[t,t+\tau]}^{\mathbb{Q}})\right] \\
\approx &\frac{1}{\tau}\left(E_t^{\mathbb{Q}}(RJV_{[t,t+\tau]}^{\mathbb{Q}}) - E_t^{\mathbb{Q}}(LJV_{[t,t+\tau]}^{\mathbb{Q}})\right),
\end{aligned}
$$

where the four lines after the first equal sign are just the definition, after the first approximation it is the difference between the expectations of the $\mathbb{P}$ and $\mathbb{Q}$ jump tail variation measures under $\mathbb{Q}$, which correlated to the special treatment of jump tail risk. The last part expresses the fear component of the tail risk premia as a function of the risk-neutral measure alone, both in the jump tail measure and in the expectation.

This fits neatly into standard asset pricing in finance because it is only dependent on $\mathbb{Q}$. In the following sections with a focus on estimation, this also poses a significant advantage over the functions in (19) and (20), as this avoids any tail estimation under $\mathbb{P}$. As with all extreme value theory, this is vulnerable in sampling, especially on a short time horizon where there might be a complete lack of large jumps, which would lead to a physical estimate of $\nu_t^{\mathbb{P}}((-\infty, \infty)) = 0$, which is not correct.

As mentioned before it will be shown that $LJP_{t,\tau}$ is orders of magnitude larger than $RJP_{t,\tau}$. Therefore, it will be assumed that empirically it holds that $RJP_{t,\tau} \approx 0$, such that the fear proxy, $LJP_{t,\tau} - RJP_{t,\tau}$ is approximately equal to the risk-neutral expectation of the negative left jump variation only,

$$LJP_{t,\tau} - RJP_{t,\tau} \approx -\frac{1}{\tau} E_t^{\mathbb{Q}}(LJV_{[t,t+\tau]}^{\mathbb{Q}}), \tag{21}$$

which leaves a simple expression for the fear component. Intuitively it also makes sense that the fear component is located mainly around negative jumps for the aggregate market portfolio; this is the primary source of fear for investors.

If the focus is on a subset of the market portfolio, e.g. hedge funds or especially short funds, this assumption will likely be in contrast with reality as a significant positive jump can be detrimental to a short position, especially in a margin situation.

At the core of asset pricing is the idea that investors require a positive long-term return for holding investments that are not risk-free. In contrast, it is natural to expect that investors should accept a low return for risk-free investments, such as the one-month Treasury bill for dollar-denominated investors, or even negative long-term returns for products that deliver positive returns in the worst times, just as everyone is used to paying a premium on insurance. This is an intuitive concept and was confirmed in Ilmanen (2012), where it was shown that buying catastrophe insurance delivered poor long-run rewards. This is also in line with Goetzmann et al. (2016) where a survey of individual and institutional investors assessed the probability of a severe single-day stock market crash to be much higher than the historical rate. As such, I expect to find a negative premium on the tail risk.

3.3 The need for time-varying jump tails

As stated in Bollerslev & Todorov (2014), most of the early literature on asset pricing assumed that the distribution of jumps was invariant, which seems unlikely given the clustering of jumps and the fact that they are distributed according to an extreme value distribution. When they were not assumed to be invariant, the dynamics of the risk-neutral tail were often assumed to be in the form of a predictable scaling process $\phi_t^{\pm}$ and a Lévy measure $\lambda^{\pm}$ that satisfies Definition 2.4, such that it is on the form of,

$$v_t^{\mathbb{Q}}(dx) = \phi_t^+ \times \lambda^+(dx)1_{\{x>0\}} + \phi_t^- \times \lambda^-(dx)1_{\{x<0\}}, \tag{22}$$

which also satisfies Definition 2.4 for any $\phi_t^{\pm}$ that is finite on $\mathbb{R} \setminus [-1, 1]$. As an example of this, one can see Bollerslev & Todorov (2011*b*) in section II.A, where it is claimed to be a very weak assumption.

A common assumption is that the jump distribution changes symmetrically for positive and negative jumps, though they might have different intensities, such that $\phi_t^+ = \phi_t^-$.

In advance of the following analysis, it is convenient to review some additional notation.

Let $\psi^+(x)$ and $\psi^-(x)$ be the functions that transform jumps in the log-price into jumps in the price level, such that

$$\psi^+(x) = \begin{cases} e^x, & x > 0 \\ 0, & x \leq 0 \end{cases}, \psi^-(x) = \begin{cases} e^-x, & x < 0 \\ 0, & x \geq 0, \end{cases} \tag{23}$$

which obviously leads to a change in the set, as $x > 0 \Rightarrow \psi^+(x) > 1$ and $x < 0 \Rightarrow \psi^-(x) > 1$. By using this transformation on the continuous relative price change on X_t it yields,

$$\frac{\Delta X_t}{X_{t-}} + 1 = \psi^+(\Delta \log(X_t))1_{\{\Delta \log(X_t)>0\}} + [\psi^+(\Delta \log(X_t))]^{-1}1_{\{\Delta \log(X_t)>0\}}.$$

As usual, $\Delta X_t = X_t - X_{t-}$ and $\Delta \log(X_t) = \log(X_t) - \log(X_{t-})$. The images of the measure $\nu_t^{\mathbb{Q}}$ under the mappings $x \rightarrow \psi^+(x)$ and $x \rightarrow \psi^-(x)$ may be expressed as,

$$\begin{aligned}
\nu^{\mathbb{Q},+}_{t,\psi}(x) &= \frac{\nu^{\mathbb{Q}}_t(\log(x))}{x} \\
\nu^{\mathbb{Q},-}_{t,\psi}(x) &= \frac{\nu^{\mathbb{Q}}_t(-\log(x))}{x}, \forall x > 1.
\end{aligned} \tag{24}$$

Lastly, let the tail integral of a given measure η on $\mathbb{R}$ be given as

$$\begin{aligned}
\bar{\eta}^{+}(x) &= \int_x^{\infty} \eta(du), \forall x > 0 \\
\bar{\eta}^{-}(x) &= \int_{\infty}^{x} \eta(du), \forall x < 0.
\end{aligned} \tag{25}$$

Even though it might seem convenient to scale a time-invariant Lévy measure with a time-variant process there is one significant issue with it, which can be seen in the following:

Let,

$$\begin{aligned}
\bar{\nu}^{\mathbb{Q},+}_t(x) &= \phi^{+}_t \times \int_x^{\infty} \lambda^{+}(du) = \phi^{+}_t \times \bar{\lambda}^{+}(x), \forall x > 0 \\
\bar{\nu}^{\mathbb{Q},-}_t(x) &= \phi^{-}_t \times \int_{-\infty}^{x} \lambda^{-}(du) = \phi^{-}_t \times \bar{\lambda}^{-}(x), \forall x < 0
\end{aligned}$$

This leads to,

$$\begin{aligned}
\frac{\bar{\nu}^{\mathbb{Q},+}_t(x)}{\bar{\nu}^{\mathbb{Q},+}_t(y)} &= \frac{\phi^{+}_t \times \bar{\lambda}^{+}(x)}{\phi^{+}_t \times \bar{\lambda}^{+}(y)} = \frac{\bar{\lambda}^{+}(x)}{\bar{\lambda}^{+}(y)}, \forall x > y > 0, \\
\frac{\bar{\nu}^{\mathbb{Q},-}_t(x)}{\bar{\nu}^{\mathbb{Q},-}_t(y)} &= \frac{\phi^{-}_t \times \bar{\lambda}^{-}(x)}{\phi^{-}_t \times \bar{\lambda}^{-}(y)} = \frac{\bar{\lambda}^{-}(x)}{\bar{\lambda}^{-}(y)}, \forall x < y < 0, \\
\frac{\bar{\nu}^{\mathbb{Q},+}_t(x)}{\bar{\nu}^{\mathbb{Q},-}_t(y)} &= \frac{\phi^{+}_t \times \bar{\lambda}^{+}(x)}{\phi^{-}_t \times \bar{\lambda}^{-}(y)}, \forall x > 0 \text{ and } \forall y < 0.
\end{aligned} \tag{26}$$

Then the relative difference between the jump measure for different sized jumps in the same direction becomes time-invariant. As such, there cannot be periods with a high (low) intensity for large jumps without also increasing (decreasing) the intensity for small jumps. There can, however, be different intensities for positive and negative jumps. If the additional assumption $\phi = \phi^{+} = \phi^{-}$ is made, then this optionality is also lost,

$$\frac{\bar{\nu}^{\mathbb{Q},+}_t(x)}{\bar{\nu}^{\mathbb{Q},-}_t(y)} = \frac{\phi_t \times \bar{\lambda}^{+}(x)}{\phi_t \times \bar{\lambda}^{-}(y)} = \frac{\bar{\lambda}^{+}(x)}{\bar{\lambda}^{-}(y)}, \forall x > y > 0, \tag{27}$$

which clearly is in violation with the behavior in the market.

The modelling scheme in (22), therefore implies that the distribution of tails depends exclusively on the tail behaviour of the time-invariant Lévy measure λ, where singularities may occur around the origin, as mentioned in Section 2.1. However, on the set limited to the real numbers outside of zero, it behaves like a probability measure. These facts combined shows that there is a need for an alternative to (22), and it will be found around the assumptions on $\bar{\lambda}_\psi^\pm$. Following part one of Assumption A2 in Bollerslev & Todorov (2011*a*),

Assumption 3.1.
$\bar{\lambda}_\psi^\pm(x)$ *are regularly varying at infinity functions, i.e.,*

$$\bar{\lambda}_\psi^\pm(x) = x^{-\alpha^\pm} L^\pm(x), \alpha^\pm > 0, \tag{28}$$

and $L^\pm(x)$ *are slowly varying at infinity, i.e.,* $\lim\limits_{x\to\infty} \frac{L(ux)}{L(x)} = 1 \forall u > 0.$

This assumption is key in both Bollerslev & Todorov (2011*a*) and Bollerslev & Todorov (2014). Therefore it is also key for the foundation of this paper. For large numerical jumps, the ratio between jumps will exhibit a linear behaviour, as seen in the following. Assume that x is near to ∞, such that $L^\pm(x+u)/L^\pm(x) = 1, \forall x > \xi$ where ξ is such a "large" enough value.

Then the limit of the ratio is,

$$\lim_{x\to\infty} \frac{\bar{\lambda}_\psi^+(x+u)}{\bar{\lambda}_\psi^+(x)} = \left(\frac{x+u}{x}\right)^{-\alpha^\pm}, x > 0, u > 0, \tag{29}$$

which has the first order derivative

$$\frac{\partial}{\partial x}\left(\frac{x+u}{x}\right)^{-\alpha^\pm} = -\alpha^\pm \left(\frac{1}{x} - \frac{x+u}{x^2}\right)\left(\frac{x+u}{x}\right)^{-\alpha^\pm - 1}.$$

The limit of the first-order derivative, $\lim\limits_{x\to\infty} \frac{\partial}{\partial x}\left(\frac{x+u}{x}\right)^{-\alpha^\pm} = 0$, indicating the measure displays linear behavior as the jump size goes to infinity. In line with Bollerslev & Todorov (2011*a*) it rules out Lévy measures with light tails, which are not in the area of interest here, and instead, restricts to the Fréchet distribution that has an infinite right endpoint, and as such, it allows for extreme jumps. It is a special case of the generalized extreme value distribution.

Definition 3.1. The CDF of the generalized extreme value distribution satisfies

$$H_\xi(x) = \begin{cases} e^{-(1+x/\xi)^{-\xi}}, & \xi \neq 0 \\ e^{-e^{-x}}, & \xi = 0 \end{cases}, \tag{30}$$

where $1 + \xi x > 0$.

The Fréchet distribution is the special case where $\xi > 0$.

By inserting this Lévy measure into (22) it follows that,

$$\frac{\bar{\nu}_t^{\mathbb{Q},\pm}(x)}{\bar{\nu}_t^{\mathbb{Q},\pm}(y)} = \frac{\phi_t^{\pm} \times \bar{\lambda}^{\pm}(x)}{\phi_t^{\pm} \times \bar{\lambda}^{\pm}(y)} \approx \left(\frac{x}{y}\right)^{-\alpha^{\pm}}, \forall x > y > 1, \tag{31}$$

where the right hand side holds true in the limit. It is evident that the ratio is still time-invariant with a power law decay determined by the maximum domain of attraction of $\bar{\lambda}^{\pm}$, which is dependent on $L^{\pm}(x)$, as long as it is restricted to the Fréchet distribution.

Even though the maximum domain of attraction (MDA) here is rich, it can be helpful to use the following theorem in showing just how general it is.

Theorem 3.1. *For $\xi > 0$,*

$$F \in MDA(H_\xi) \Leftrightarrow \bar{F}(x) = x^{-\xi} L(x)$$

for some function, L, that is slowly varying at ∞ and where $\bar{F}(x) = 1 - F(x)$.

By this theorem it follows easily that e.g. the double-exponential jump model from Kou (2002) belongs to this,

$$\begin{aligned} &\lambda^+(x) = c^+ e^{-\alpha^+ x} 1_{\{x>0\}} \\ &\lambda^-(x) = c^- e^{\alpha^- x} 1_{\{x<0\}}, \alpha^{\pm} > 0, c^{\pm} \geq 0. \end{aligned} \tag{32}$$

By applying the transformation in (24) and taking the tail integral it yields,

$$\begin{aligned} \lambda_\psi^+(x) &= \frac{\lambda^+(\log(x))}{x} = \frac{c^+}{x} e^{-\alpha^+ \log(x)} 1_{\{x>0\}} = \frac{c^+}{x} x^{-\alpha^+}, \forall x > 1, \\ \lambda_\psi^-(x) &= \frac{\lambda^-(-\log(x))}{x} = \frac{c^-}{x} e^{\alpha^- \log(x)} 1_{\{x<0\}} = \frac{c^-}{x} x^{-\alpha^-}, \forall x > 1, \\ \bar{\lambda}_\psi^+(x) &= \int_x^\infty \frac{c^+}{u} u^{-\alpha^+} du = \frac{c^+}{\alpha^+} x^{-\alpha^+}, \forall x > 1, \\ \bar{\lambda}_\psi^-(x) &= \int_x^\infty \frac{c^-}{u} u^{-\alpha^-} du = \frac{c^-}{\alpha^-} x^{-\alpha^-}, \forall x > 1. \end{aligned} \tag{33}$$

As such, by Theorem 3.1 it is given that the CDF of the double-exponential model is in the maximum domain of attraction of Fréchet, with $\xi = \alpha^{\pm}$ and $L^{\pm}(x) = \frac{c^{\pm}}{\alpha^{\pm}}$.

Despite the seeming richness of this setup and its ability to accommodate a range of the most popular models applied in empirical finance and risk management, it still has a significant problem with its ability to manage temporal variation in jump tails. The assumption of constant shape tails, just like the assumption of constant volatility in the Black-Scholes-Merton setup, is in clear violation of how tails behave during times of financial calm and times of financial crisis. Therefore, (22) will be used as a basis for the model, but where the Lévy measure is replaced with a time variable process inspired by the shape in Assumption 3.1,

$$\nu_t^{\mathbb{Q}}(dx) = \left(\phi_t^{+} \times e^{-\alpha_t^{+} x} 1_{\{x>0\}} + \phi_t^{-} \times e^{-\alpha_t^{-}|x|} 1_{\{x<0\}}\right) dx \tag{34}$$

Instead of having one source of time variance, there is a time-variant intensity shift across all jump levels, as seen in $\phi_t^{\pm}$. The advantage compared to before is the rate of decay of the tails governed by $\alpha_t^{\pm}$, which is both time-variant but also varies across jump sizes and, as such, allows for periods with adjusting intensity for large jumps without changing the intensity for small jumps. It is also evident that the tail measure $\bar{\lambda}_{t,\psi}^{\pm}$ is now time-variant and proportional to the rate of decay in the tail, such that

$$\bar{\lambda}_{t,\psi}^{\pm}(x) \propto x^{-\alpha_t^{\pm}}, \alpha_t^{\pm} > 0, x \to \infty. \tag{35}$$

This makes it evident that the ratio in (31) is also time variant,

$$\frac{\bar{\nu}_t^{\mathbb{Q},\pm}(x)}{\bar{\nu}_t^{\mathbb{Q},\pm}(y)} = \frac{\phi_t^{\pm} \times \bar{\lambda}_t^{\pm}(x)}{\phi_t^{\pm} \times \bar{\lambda}_t^{\pm}(y)} \approx \left(\frac{x}{y}\right)^{-\alpha_t^{\pm}}, \forall x > y > 0, \tag{36}$$

where it is evident that the expression in (34) fulfils the wishes for a jump measure that is time-variant across both level shifts and the cross-section of jumps.

Following this, there will now be a review of the main data employed in this thesis and some of its main features. An extensive and relatively computer-intensive cleaning procedure is also employed, which will be covered in its main steps.

4 Data

The estimation is based on the case of S&P500, or the Standard & Poor's 500 Index, which is a market-capitalization-weighted index that primarily includes 500 of the largest U.S. companies by market capitalization and as of the end of 2018, it covered 83.3% of the market capitalization of U.S. Equities.
Inclusion into the S&P 500 index have historically had a positive effect on the stock price, but as covered in Bennett et al. (2020) the positive announcement effect on the stock price of index inclusion has disappeared, and the long-run impact of index inclusion has become negative, which seems surprising considering the added demand for the stock from passive investment funds that track the S&P 500.

The data used here are primarily option data traded on the Chicago Board of Options Exchange (CBOE) on the S&P 500 Index ranging from 1996-01-04 to 2020-12-31 for a total of 6293 dates. The data is obtained from OptionMetrics through WRDS and includes the future price, and the Black-Scholes implied volatility as calculated by OptionMetrics. The period includes both the financial crisis of 2007-2008 and the impact of COVID-19 on the S&P 500 in the year 2020.

As described in Section 5, the estimates for the jump tail parameters in (48) and (51) are reliant on either an increasing number of out-of-the-money options or an increasing time horizon to eliminate the impact of the diffusive price component, which is not of interest. The modelling scheme is employed to hopefully only model price change due to jumps, which is more likely for a short time-to-maturity.
As such, the analysis is restricted to options with no more than 45 days until expiration. Super short-lived options pose an issue with market microstructure complications, which are not of interest either, and as such, a minimum time-to-expiry of 8 days is also imposed. Some diffusive risk will always be present, but this issue will be handled in Section 6

As is evident from Figure 1, there has been tremendous growth in the cross-section of option starting in around 2014 and seemingly plateauing around 2019-20. As several of the methods used in the estimation in the previous section were dependent on the number of options going to infinity, this is a significant difference compared to the previous papers, such as Bollerslev & Todorov (2014) and Bollerslev et al. (2015), where the data were far more limited.
It is important to note that the number of options in Figure 1 are based on the entire cross-section and before any cleaning has been done, and as such, the number of options employed in this thesis is far more limited.

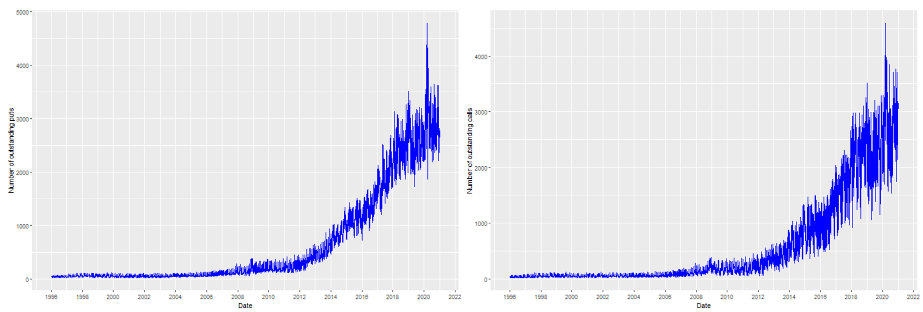


Figure 1: Left: Number of outstanding puts.
Right: Number of outstanding calls.

Apart from the closing bid and ask quotes for all S&P 500 options traded on the CBOE, the analysis directly employs the implied volatilities provided by OptionMetrics, as recommended in Carr & Wu (2009), and the relevant future prices.

The aggregate market return predictability regressions are based on a broad value-weighted portfolio of all Center for Research in Security Prices (CRSP), and the relevant time-series of daily and monthly returns are obtained from the AQR data library. The AQR data library is also the source for the Fama-French factor portfolios, the quality-minus-junk portfolio, and the betting-against-beta portfolio. The Fama-French daily factors data source in WRDS is the source for the risk-free rate.

4.1 Cleaning procedure

To avoid any obvious errors in the data, the cleaning procedure has been inspired by Carr & Wu (2003), and as such it follows,

i. The time to maturity is greater than 8 days,

ii. The time to maturity is less than 45 days,

iii. The bid option price is strictly positive,

iv. The implied volatility is valid,

v. The option is out-of-the-money,

vi. The ask price is no less than the bid price.

Apart from this, I have also applied a method to rule out arbitrage. Starting by sorting the options by date, time to expiration, and strike price. Then, for each pair of date and time to expiration, I begin with the closest at-the-money options and only keep the subsequent out-of-the-money options where the midquote is lower than the previous. If this condition is violated for a given pair of options, then the option with the highest volume at that date is retained. If the volume is identical, this often happened for a volume of zero, for a specific option on a given day, then the option that is closest to being at-the-money is retained.
If several options on a given date have the same strike and time to expiry, then they are merged by taking the midquote as an average.

Before the cleaning procedure, there were roughly 2.9 million data points on puts where they were out-of-the-money compared to the matching future. After the cleaning procedure, roughly 1.8 million data points were left. Approximately half of the removed options were due to the bid option price not being strictly positive. This leaves the following plot over the puts, where it seems to have

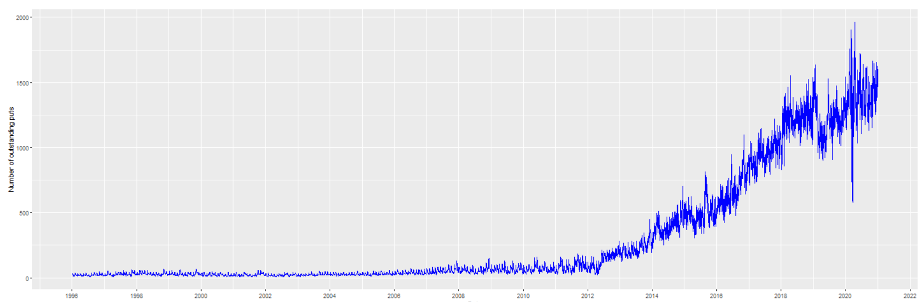


Figure 2: Number of outstanding puts after cleaning.

stabilized around 1300-1400 of outstanding puts at each date. In order to sustain such a significant increase in the number of outstanding puts, it would either indicate an increase in the width of expiry dates or an increase in how far out-of-the-money.

Below is a plot over the "depth" of the cross-section of the puts, where for each pair of date and expiration date, I have found the greatest difference between the strike price and the corresponding forward price. For each date, I have found the maximum and minimum depth over the corresponding expiration dates.

Figure 3 reveals an interesting picture. It seems reasonable that the maximum value of the depth is highly correlated with the number of outstanding puts at each date, as more outstanding options will typically indicate a broader

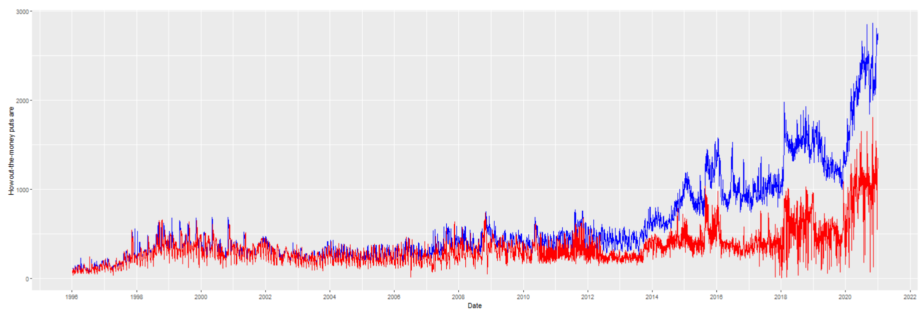


Figure 3: Blue: The maximum depth for each date.
Red: The minimum depth for each date.

range of expiration dates and especially an increase in the maximum of time to expiration. Options with a longer time to expiration will naturally have higher bids than a put with the identical strike but a shorter time to expiry.

It is, however, crucial to keep in mind that the options included here are limited to 45 days to expiry. Especially in the earlier years with a low number of outstanding puts, this limit was rarely reached where on the other hand it was reached. This can also be seen in the average maximum time to expiry at each date. Before 2015 this value is at 32.2, and after 2015 the value is at 42.8. After 2015 this value is relatively stable and does not explain the trend seen in Figure 3. As such, there must be an increasing demand for far out-of-the-money puts, perhaps in order to hedge tail risk.

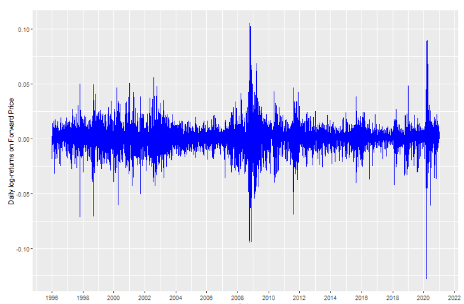
The minimum depth for each date reinforces this theory. With the minimum time to expiry limit of 8 days, one could expect a decreasing trend in the red line in Figure 3, perhaps causing the slightly decreasing trend from 2000 to 2004, as more and more puts with different expiration dates are available over the time. The average minimum time to expiration at each date before 2015 is at 19.5 and at 9.1 after 2015, showing that the range of expiration dates becomes far more detailed over time.
Nonetheless, this is in dire contrast to the trend line after 2015, but especially after 2018, where the minimum depth has been increasing quite drastically and thereby showing the need for deep out-of-the-money options and why it is such an exciting area to investigate.

4.2 Descriptive statistics

Table 1 shows that the average annualized log return for the period is roughly 7.2% with yearly volatility of nearly 20%. More interesting is that the skewness is -0.42, which is reasonably symmetrical but dominated towards negative jumps, and with a kurtosis of 13.3, or excess kurtosis of 10.3. The returns are roughly bell-shaped but with far heavier tails than expected in a normal distribution and longer tails towards the left. This reinforces the idea that there is a need for jumps and that the focus should be on negative jumps.

	Forward return
Mean (annualized)	0.029% (7.216%)
Volatility (annualized)	1.231% (19.546%)
Skewness	-0.418
Kurtosis	13.285
Minimum	-12.786%
Maximum	10.539%

Table 1: Descriptive statistics for the average log forward price and puts.



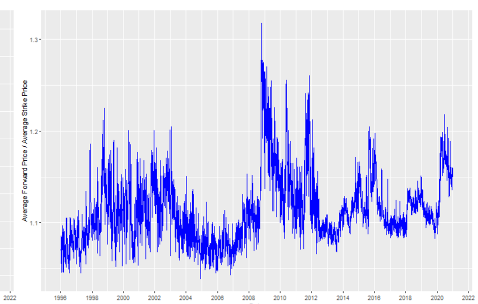


Figure 4: Left: Daily log returns for the forward price on S&P500 for the period of 1996-2020.
Right: Daily ratio of average forward price and average strike price of cleaned puts.

Figure 4 shows that the ratio of constant jump intensity is in contradiction with reality, as there are massive jumps both during the financial crisis, the European debt crisis, and the COVID-19 crisis. It is also evident that the original assumption of constant volatility in the Black-Scholes-Merton model is improbable as the most significant drop is more than ten standard deviations away from the mean. In a model with constant volatility and no jumps, this is highly unlikely.

The right part of Figure 4 is also interesting, as it shows that the ratio between average forward price on a given day over average strike price on a given day goes up in times of financial distress, indicating that even though the Forward Price goes down, the average strike price goes down even further. This could be the result of market makers offering options further out-of-the-money than usual, or that there were bids on deep out-of-the-money options, where they were usually zero-bid options and therefore dropped by the cleaning algorithm.
Keep in mind that this is a simple average and does not take open interest or daily volume into account.

A discussion that is always valid in the study of rare events is if 25 years of data is enough? It would have been preferable to have a more extended period of data, such as a century, as the number of extreme events is limited. However, it was not until 1983 that CBOE created options on broad-based stock indices, and as such, a century remains a dream. Data from 1983 up to 1996 could have been employed, but this is not included in OptionMetrics, the primary data source employed. However, over the 25 years, there have been exciting index options data on extreme events. There were significant market events in 1998, with the default of Russia and the crash of Long-Term Capital Management, 9/11 in 2001, the financial crisis in 2008, the European/Greek debt crisis around 2010/2011, and recently Covid in 2020.

The following section will cover tail approximations and how the data will be used in estimation procedures, and what the implications of the tail risk measures are on $LJP_{t,\tau}$.

5 Jump tail modelling

The estimation of the $\mathbb{Q}$ jump tail measures builds on the models of Bollerslev & Todorov (2011*b*) and Bollerslev & Todorov (2014). The use of deep out-of-the-money puts have been used as Crash Insurance historically, e.g. see Chen et al. (2019) or Ilmanen (2012) and the references herein, and are therefore a natural choice to study when the topic is tail risk. As mentioned in Section 3.2 it behaves in many ways like ordinary insurance, in that you pay a premium on it, and expectations inherent in the cross-section of puts relate directly to the risk-neutral distribution from a pricing perspective.

Another reason to use deep out-of-the-money options that are close-to-maturity is that they contain essential information about the risk-neutral jump measure and facilitates tail estimation without the need for a large number of crisis events. As mentioned in Section 4.2 the period is limited to just 25 years, and in Figure 4 it is evident that there have been roughly 5-7 periods of uncertainty, dependent on how the limit for a significant jump is set, which is a relatively limited amount of data to estimate multi-variable parameters. Here, the risk-neutral pricing of options allows for the estimation of the risk-neutral tail measure. Just as in 3.3 the discussion starts with the method concerning time-invariant jump tails, and then it expands to time-variant jump tails.

5.1 Modelling the time-invariant tail shape

The notation is as follows, let $O_{t,\tau}(k)$ denote the time t price of an out-of-the-money option on X_t with time to expiration τ and log-moneyness $k = \log(K/F_{t-,\tau})$, where $F_{t,\tau}$ refers to the futures price of X_t with the future date of τ, and K denotes the strike of the option. By Proposition 1 in Bollerslev & Todorov (2011*b*) it follows that for the time-invariant jump intensity process in (22) the model-free risk-neutral jump tail measures can be estimated with,

$$\frac{e^{r_{t,\tau}} O_{t,\tau}(k)}{F_{t-,\tau}} \approx \begin{cases} \int_t^{t+\tau} \int_{\mathbb{R}} (e^x - e^k)^+ E_t^{\mathbb{Q}}(\nu_s^{\mathbb{Q}}(dx)) ds, & \text{if } k > 0 \\ \int_t^{t+\tau} \int_{\mathbb{R}} (e^k - e^x)^+ E_t^{\mathbb{Q}}(\nu_s^{\mathbb{Q}}(dx)) ds, & \text{if } k < 0, \end{cases} \tag{37}$$

where $e^{r_{t,\tau}} = E_t^{\mathbb{Q}}(e^{\int_t^{t+\tau} r_s ds}$ denotes the $\mathbb{Q}$-expected risk-free interest rate over $[t, t+\tau]$. It is evident that this approximation on the right-hand side is only dependent on the jump measure, as the approximation relies on $t + \tau \to t$ and $k \to \pm\infty$. With $\tau \to 0$, the price change due to continuous will also go to zero, and for the limited period, there will be at most one large jump. As the data here is limited to a minimum of eight days to expiry, the $t + \tau \to t$ condition will not be complied with, and therefore, there will be a continuous component that affects $O_{t,\tau}(k)$. By the definition of large jumps and continuous drift, the continuous component is minor, and as such, it will be ignored.

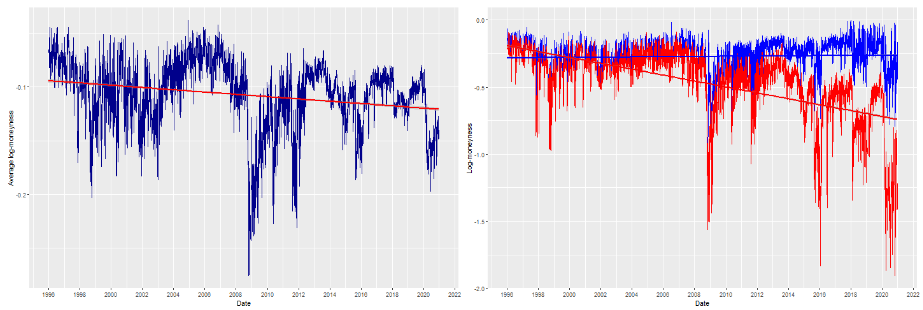


Figure 5: Left: Average log-moneyness on S&P500 for the period of 1996-2020. Right: Minimum (red)/maximum (blue) log-moneyness on each date over each maturity date.

The data set for the second condition, $k \to \pm\infty$, has dramatically improved over the recent years, which can be seen in Figure 5, where it is evident that both the average log-moneyness has decreased slightly over the period, but especially that the minimum log-moneyness has decreased drastically for puts in recent years, as seen in the red plot on the right. The minimum log-moneyness has gone from around -0.25 to around -1.25 at the end of 2020, which is a drastic drop.
The approximation error in (37) was researched by Monte Carlo simulation in Bollerslev & Todorov (2011*b*) and found to be relatively small for the maturity and moneyness employed in their paper. The same maturity limits are employed in this thesis, but as the moneyness is better suited here than in their paper, the approximation error will be of even less importance. In line with Bollerslev & Todorov (2014) the approximation error will be ignored in the upcoming sections.

By using the extreme value approximation for the baseline jump intensity process in Bollerslev & Todorov (2014) together with (37), it follows that

$$\frac{e^{r_{t,\tau}} O_{t,\tau}(k)}{\tau F_{t-,\tau}} \approx (\phi_t^+ 1_{\{k>0\}} + \phi_t^- 1_{\{k<0\}}) \Phi(\alpha^\pm, \mathrm{tr}, k), \tag{38}$$

where

$$\Phi(\alpha^\pm, \mathrm{tr}, k) = \begin{cases} \frac{\bar{\lambda}_\psi^+(\mathrm{tr})}{\alpha^+ - 1} \frac{(e^k)^{1-\alpha^+}}{\mathrm{tr}^{-\alpha^+}}, & e^k \geq \mathrm{tr} > 1, \\ \frac{\bar{\lambda}_\psi^-(\mathrm{tr})}{\alpha^- + 1} \frac{(e^{-k})^{-1-\alpha^-}}{\mathrm{tr}^{-\alpha^-}}, & e^{-k} \geq \mathrm{tr} > 1, \end{cases} \tag{39}$$

for some threshold tr > 1. Again it is interesting to review the ratio of the measure under two differently sized jumps at the same time. For simplicity assume that it is a put, such that $|k_2| > |k_1|$ and that $e^{-k_2} > e^{-k_1} \geq \text{tr} > 1$,

$$\begin{aligned}
\frac{\frac{e^{r_{t,\tau}} O_{t,\tau}(k_2)}{\tau F_{t-,\tau}}}{\frac{e^{r_{t,\tau}} O_{t,\tau}(k_1)}{\tau F_{t-,\tau}}} &= \frac{O_{t,\tau}(k_2)}{O_{t,\tau}(k_1)} \\
&\approx \frac{(\phi_t^+ 1_{\{k>0\}} + \phi_t^- 1_{\{k<0\}})\Phi(\alpha^\pm, \text{tr}, k_2)}{(\phi_t^+ 1_{\{k>0\}} + \phi_t^- 1_{\{k<0\}})\Phi(\alpha^\pm, \text{tr}, k_1)} = \frac{\Phi(\alpha^\pm, \text{tr}, k_2)}{\Phi(\alpha^\pm, \text{tr}, k_1)} \\
&= \frac{\frac{\bar{\lambda}_\psi^-(\text{tr})}{\alpha^-+1} \frac{(e^{-k_2})^{-1-\alpha^-}}{\text{tr}^{-\alpha^-}}}{\frac{\bar{\lambda}_\psi^-(\text{tr})}{\alpha^-+1} \frac{(e^{-k_1})^{-1-\alpha^-}}{\text{tr}^{-\alpha^-}}} = \frac{(e^{-k_2})^{-1-\alpha^-}}{(e^{-k_1})^{-1-\alpha^-}}, e^{-k_2} > e^{-k_1} \geq \text{tr}.
\end{aligned} \tag{40}$$

This can be simplified even further by taking the logarithm, such that

$$\log\left(\frac{O_{t,\tau}(k_2)}{O_{t,\tau}(k_1)}\right) = (-1-\alpha^-)(-k_2) - (-1-\alpha^-)(-k_1) = (1+\alpha^-)(k_2 - k_1), \tag{41}$$

for $e^{-k_2} > e^{-k_1} \geq \text{tr}$. This leaves a simple expression for the time-invariant tail decay parameter, which can be estimated through an increasing number of short-maturity puts with decreasing strike or an increasing number of puts over an increasing sample of span T, or both. If just two options are available then it is clear that a simple estimator of α^- would be,

$$\frac{\log\left(\frac{O_{t,\tau}(k_2)}{O_{t,\tau}(k_1)}\right)}{(k_2 - k_1)} - 1 = \hat{\alpha}^-,$$

which clearly can be calculated as the entire left hand side is known. For several options it is desirable to minimize it as a function of the errors or the absolute deviation. For a single day the target to minimize naturally becomes,

$$\sum_{i=2}^{N^-} g(\alpha^- - \hat{\alpha}^-) = \sum_{i=2}^{N^-} g\left(\alpha^- - \left(\frac{\log\left(\frac{O_{t,\tau}(k_i)}{O_{t,\tau}(k_{i-1})}\right)}{(k_i - k_{i-1})} - 1\right)\right),$$

where the function $g : \mathbb{R} \to \mathbb{R}_+$ such that $g(x) = 0$ if and only if $x = 0$. As each term here is clearly positive by the definition of g then the sum will be normalized by the number of options minus 1. Then $\hat{\alpha}$ is the value that minimizes that, and then the one day estimation becomes,

$$\hat{\alpha}^- = \underset{\alpha^-}{\operatorname{argmin}} \frac{1}{N^- - 1} \sum_{i=2}^{N^-} g\left(\alpha^- - \left(\frac{\log\left(\frac{O_{t,\tau}(k_i)}{O_{t,\tau}(k_{i-1})}\right)}{(k_i - k_{i-1})} - 1\right)\right), \tag{42}$$

where N^- denotes the total number of puts on the day with log-moneyness $0 < -k_1 < -k_2 < ... < -k_{N^-}$. But, as the data set does not consist of a single day this will have to be modified slightly. The time invariant $\hat{\alpha}^-$ follows,

$$\hat{\alpha}^- = \underset{\alpha^-}{\operatorname{argmin}} \frac{1}{\sum_{t=1}^{T}(N_t^- - 1)} \sum_{t=1}^{T} \sum_{i=2}^{N_t^-} g\left(\alpha^- - \left(\frac{\log\left(\frac{O_{t,\tau}(k_{t,i})}{O_{t,\tau}(k_{t,i-1})}\right)}{(k_{t,i} - k_{t,i-1})} - 1\right)\right). \tag{43}$$

5.2 Modelling the time-variant tail shape

The approximations underlying (43) are based on (22) and as such it assumes that the tail decay parameter is time invariant. The use of (34) complicates the situation, even though the outcome resembles the time-invariant case greatly. The time variant alternative to (37) is as follows,

$$\frac{e^{r_{t,\tau}} O_{t,\tau}(k_t)}{\tau F_{t-,\tau}} \approx \begin{cases} \frac{\phi_t^+ e^{k_t(1-\alpha_t^+)}}{\alpha_t^+(\alpha_t^+ - 1)}, & \text{if } k_t > 0 \\ \frac{\phi_t^- e^{k_t(1+\alpha_t^-)}}{\alpha_t^-(\alpha_t^- + 1)}, & \text{if } k_t < 0, \end{cases} \tag{44}$$

for $\tau \downarrow 0$. The approximation steps in proving this are outside the scope of this thesis, and the reader is referred to Lemma 1 and the proof hereof in Bollerslev & Todorov (2014).

It does behave in the same manner as in the time-invariant case and as such the same logic can be applied. A requirement for this should be that (44) equals (38) for $\alpha_s^\pm = \alpha_t^\pm$ and $\phi_s^\pm = \phi_t^\pm$ for $s \in [t, t+\tau]$. I will show it for the case of puts, but calls follows in an identical matter. Let the right hand side of (44) equal the right hand side of (38),

$$\begin{aligned} \frac{\phi_t^- e^{k(1+\alpha_t^-)}}{\alpha_t^-(\alpha_t^- + 1)} &= \phi_t^- \frac{\bar{\lambda}_\psi^-(\mathrm{tr})}{\alpha^- + 1} \frac{(e^{-k})^{-1-\alpha^-}}{\mathrm{tr}^{-\alpha^-}}, k < 0, \\ \Rightarrow \frac{e^{k(1+\alpha_t^-)}}{\alpha_t^-} &= \bar{\lambda}_\psi^-(\mathrm{tr}) \frac{e^{k(1+\alpha_t^-)}}{\mathrm{tr}^{-\alpha^-}}, k < 0, \\ \Rightarrow \frac{1}{\alpha_t^-} &= \bar{\lambda}_\psi^-(\mathrm{tr}) \frac{1}{\mathrm{tr}^{-\alpha^-}}, k < 0, \\ \Rightarrow \bar{\lambda}_\psi^-(\mathrm{tr}) &= \frac{\mathrm{tr}^{-\alpha^-}}{\alpha_t^-} = \mathrm{tr}^{-\alpha^-} L^-(\mathrm{tr}), k < 0, \end{aligned} \tag{45}$$

which is clearly on the form of Assumption 3.1 with $L^-(\mathrm{tr}) = \frac{1}{\alpha^-}$ being constant.

In the same pattern as before, the estimator is found through the ratio of the logarithmic prices,

$$\log\left(\frac{O_{t,\tau}(k_{t,2})}{O_{t,\tau}(k_{t,1})}\right) = \log\left(\frac{\left(\frac{\phi_t^- e^{k_{t,2}(1+\alpha_t^-)}}{\alpha_t^-(\alpha_t^-+1)}\right)}{\left(\frac{\phi_t^- e^{k_{t,1}(1+\alpha_t^-)}}{\alpha_t^-(\alpha_t^-+1)}\right)}\right) = (k_{t,2}-k_{t,1})(1+\alpha_t^-), k_{t,i}<0, \tag{46}$$

which is identical as before except the tail decay parameter is no longer time-invariant. As such, the simple estimator follows in the same pattern as (42),

$$\hat{\alpha}_t^- = \underset{\alpha_t^-}{\operatorname{argmin}} \frac{1}{N_t^- - 1}\sum_{i=2}^{N_t^-} g\left(\alpha_t^- - \left(\frac{\log\left(\frac{O_{t,\tau}(k_{t,i})}{O_{t,\tau}(k_{t,i-1})}\right)}{(k_{t,i}-k_{t,i-1})} - 1\right)\right), \tag{47}$$

estimates α_t^- for $t = 1, 2, .., T$ when $N_t^- \to \infty$. In practice, this is flawed as the number of options available at each date is finite and not infinite. However it has been growing in later years. As a compromise, for the empirical results reported below, the results are estimated on a time-varying weekly basis by summing the right-hand side of (47) over a weekly basis. In practice, there is also a need to take care of different maturity dates for each date, and as such, the employed smoothed estimator is,

$$\hat{\alpha}_t^- = \underset{\alpha_t^-}{\operatorname{argmin}} \frac{1}{\sum_{t=1}^{\tau}(N_t^- - 1)}\sum_{t=1}^{\tau}\sum_{k=1}^{K}\sum_{i=2}^{N_{k,t}^-} g\left(\alpha_t^- - \left(\frac{\log\left(\frac{O_{t,\tau}(k_{t,i})}{O_{t,\tau}(k_{t,i-1})}\right)}{(k_{t,i}-k_{t,i-1})} - 1\right)\right), \tag{48}$$

where $N_{1,t}^- + ... + N_{k,t}^- = N_t$, K is the number of unique maturity dates at each date, and τ is the number of days in the smoothing period. In practice, the last sum is done through vector operations as they are far more efficient.

5.3 Modelling the time-variant level shift

In order to completely characterize the $\mathbb{Q}$ jump intensity process, it is necessary also to model the time-variant level shift, not just the tail decay $\alpha_t^{\pm}$. It is important to notice that neither (47) nor (48) puts any restrictions on the parameter $\phi_t^{\pm}$. As such, $\phi_t^{\pm}$ will be modelled through utilizing (44). With the case of the put and ignoring the approximation error,

$$\frac{e^{r_{t,\tau}}O_{t,\tau}(k_t)}{\tau F_{t-,\tau}} \approx \frac{\phi_t^- e^{k_t(1+\alpha_t^-)}}{\alpha_t^-(\alpha_t^- + 1)}$$
$$\Leftrightarrow \hat{\phi}_t^- = \frac{e^{r_{t,\tau}}O_{t,\tau}(k_t)}{\tau F_{t-,\tau}} \frac{\hat{\alpha}_t^-(\hat{\alpha}_t^- + 1)}{e^{k_t(1+\hat{\alpha}_t^-)}}.$$

In order to stabilize it, the logarithm is taken,

$$\log(\hat{\phi}_t^-) = \log\left(\frac{e^{r_{t,\tau}}O_{t,\tau}(k_t)}{\tau F_{t-,\tau}}\right) + \log(\hat{\alpha}_t^-) + \log(\hat{\alpha}_t^- + 1) - k_t(\hat{\alpha}_t^- + 1). \quad (49)$$

From this the estimator follows directly,

$$\begin{aligned}\hat{\phi}_t^- &= \underset{\phi_t^-}{\operatorname{argmin}} \frac{1}{N_t^-}\sum_{i=1}^{N_t^-} g(\log(\hat{\phi}_t^-) - \log(\phi_t^-)) \\ &= \underset{\phi_t^-}{\operatorname{argmin}} \frac{1}{N_t^-}\sum_{i=1}^{N_t^-} g\Bigg(\log\left(\frac{e^{r_{t,\tau}}O_{t,\tau}(k_t)}{\tau F_{t-,\tau}}\right) \\ &\quad + \log(\hat{\alpha}_t^-) + \log(\hat{\alpha}_t^- + 1) - k_t(\hat{\alpha}_t^- + 1) - \log(\phi_t^-)\Bigg). \end{aligned} \quad (50)$$

As with tail decay, it is also necessary to adapt it to smooth over a week instead of daily data and adapt it to handle different maturities. The employed smooth estimator is,

$$\begin{aligned}\hat{\phi}_t^- &= \underset{\phi_t^-}{\operatorname{argmin}} \frac{1}{\sum_{t=1}^{\tau} N_t^-}\sum_{t=1}^{\tau}\sum_{k=1}^{K}\sum_{i=i}^{N_{k,t}^-} g\Bigg(\log\left(\frac{e^{r_{t,\tau}}O_{t,\tau}(k)}{\tau F_{t-,\tau}}\right) \\ &\quad + \log(\hat{\alpha}_t^-) + \log(\hat{\alpha}_t^- + 1) - k_t(\hat{\alpha}_t^- + 1) - \log(\phi_t^-)\Bigg), \end{aligned} \quad (51)$$

where again, in practice, the last sum and $\sum_{t=1}^{\tau} N_t^-$ are solved through vector operations.

Before proceeding with the empirical modelling of the jump tail, it is crucial to control that Definition 2.4 holds true such that $\nu_t^-(dx)$ is an actual Lévy measure. The first part holds true trivially since the jump measure is zero for jumps of size zero as (34) is defined with sharp inequalities. The second part is controlled by calculating the integrals,

$$
\begin{aligned}
&\int_{\mathbb{R}} (|x|^2 \wedge 1)\nu(dx) < 0 \\
&\Rightarrow \int_{-\infty}^{-1} \nu(dx) + \int_{-1}^{1} |x|^2 \nu(dx) + \int_{1}^{\infty} \nu(dx) \\
&= \frac{\phi_t^-}{\alpha_t^-} e^{-\alpha_t^-} + \frac{\phi_t^-}{(\alpha_t^-)^3} e^{-\alpha_t^-} + \frac{\phi_t^+}{(\alpha_t^+)^3} e^{-\alpha_t^+} + \frac{\phi_t^+}{\alpha_t^+} e^{-\alpha_t^+} < 0.
\end{aligned}
\tag{52}
$$

As such, it is a requirement that $\alpha_t^{\pm}$ is strictly positive and $\phi_t^{\pm}$ is finite for $\nu_t^{\mathbb{Q}}(dx)$ being a Lévy measure.

Now the $\mathbb{Q}$ jump intensity process is completely characterized, and the focus can shift to the empirical side.

6 Empirical jump tail modelling

The estimators for the $\mathbb{Q}$ jump intensity process in (48) and (51) are reliant on short-lived out-of-the-money options in an attempt to assess the jump tail risk and remove diffusive risk, which is still present. The idea here is that if there is a fictive future price of \$1000 and a strike price of \$995, then the diffusive risk is significant. On the other hand, with a future price of \$1000 and a strike price of \$400, as long as the volatility is limited, the diffusive risk is insignificant. Therefore, instead of just looking at out-of-the-money options, I am looking specifically at deep-out-of-the-money options, where the threshold of the moneyness is dependent on the at-the-money Black-Scholes implied volatility as calculated in OptionMetrics.

6.1 Choice of log-moneyness limit

A lower limit fixed strike could be used instead of a variable strike, but in times of low volatility, it would be too conservative and filter out too many options. By using a flexible strike instead of a lower fixed strike, it conserves more options. In order to decide the limit for the relative log-moneyness, Figure 6 shows the average number of weekly bonds before 2015 and after 2015 as a function of the log-moneyness. It would not be unreasonable to employ a stricter limit in recent years if, for example, a focus is on 2020 and Covid. In

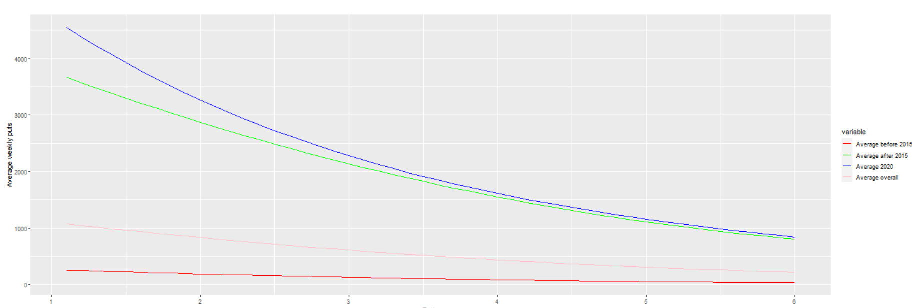


Figure 6: Average weekly number of puts above $k < -x \times \sigma_t^{\text{ATM}} \sqrt{\frac{\tau}{365}}$

order to accommodate at least a 100 weekly puts on average in the early period, I have placed the limit at $k < -3.5 \times \sigma_t^{\text{ATM}} \sqrt{\frac{\tau}{365}}$, where k is the log-moneyness defined as $k = \log(K/F_{t-,\tau})$. This gives an average of 30 puts weekly before 2010, 105 weekly before 2015, 1185 after 2015, 1915 in 2020, and 519 overall. It is also evident that the recent data can accommodate far stricter limits on which options to include, and as such, it can remove far more diffusive risk.

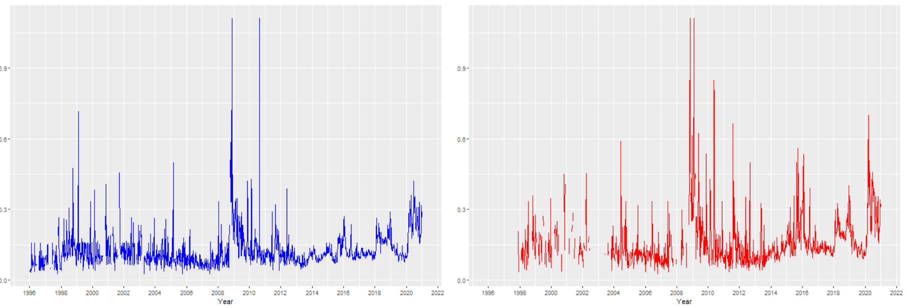


Figure 7: Left jump tail index estimates for $1/\alpha_t^-$ under the assumption that the shape of the left jump tails are constant over weekly horizons.
Left: Deep-out-of-the-money defined as $k < -3.5 \times \sigma_t^{\mathrm{ATM}} \sqrt{\frac{\tau}{365}}$.
Right: Deep-out-of-the-money defined as $k < -5 \times \sigma_t^{\mathrm{ATM}} \sqrt{\frac{\tau}{365}}$

In Figure 7 the estimates for $1/\alpha_t^-$ are plotted for a weekly horizon for a more lenient limit of $k < -3.5 \times \sigma_t^{\mathrm{ATM}} \sqrt{\frac{\tau}{365}}$ and a stricter limit of $k < -5 \times \sigma_t^{\mathrm{ATM}} \sqrt{\frac{\tau}{365}}$. The rough structure of the two plots are the same, except for the missing data points under the strict assumption, but $1/\alpha_t^-$ is also far more volatile and spikes out for the crisis periods in 2010-2012, where the lenient version is far more stable, apart from the financial crisis in 2007-08 and 2020 where it spikes out as expected. Interestingly, the tail decay is much steeper in 2020 compared to 2007-08. This might be due to an overall increase in all jumps, both smaller and larger. This will be reviewed in a later section.

6.2 Is it necessary to model time-variant tail shapes?

Whether or not it is necessary to model time-variant tail shapes instead of the simpler time-invariant shapes will be answered twofold.
I will cover the tail shapes modelled weekly, monthly, quarterly, and annually to see if the estimation time horizon is of importance and which conclusions stand out.

Note that plots in 8 are based such that a period is listed over the end time of the period, e.g. the period 2020-2021 is listed over 2021 in the bottom plot. The behaviour of the tail shapes resembles those of a mean reverting function with non-constant volatility Brownian motion and a jump function. Across the four time horizons, three of them are reasonably close to each other, apart from the naturally smoother estimates for the annual and quarterly pooling. Even at low frequencies, there are clear patterns in the estimates with fatter tails around the dot-com bubble, the financial crisis, and especially 2018 and 2020.

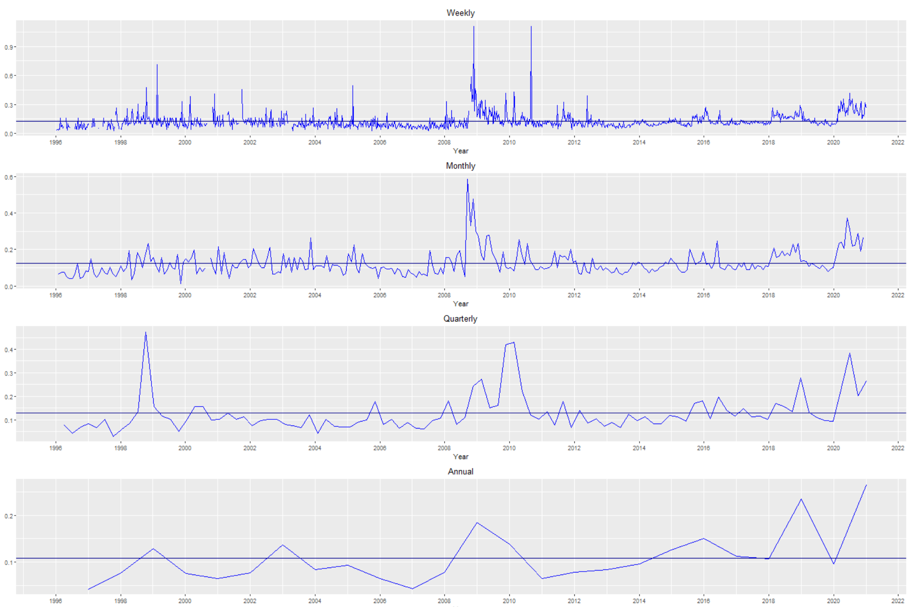


Figure 8: The estimates of $1/\alpha_t^-$ under the assumption that the shape of the left jump tails are constant over weekly (top), monthly (second), quarterly (third), and the annual (bottom) horizons respectively. The horizontal line is the mean.

The quarterly plot stands out as quite distinct from the other plots as its sharpest spikes are around the Russian crisis and the crash of Long-Term Capital Management in 1998 and centred around late 2009 and 2010 and not in 2008-09, where the financial crisis was at its worst. Allowing the tail parameters to change at a monthly or weekly frequency further reinforces the temporal variation in the shapes of the jump tails. The shape of the two are very similar to each other, but the magnitude of the spikes in the weekly estimates is far greater than the magnitude in the monthly estimate. The point estimates in the weekly estimates are noisier, as expected, and are more susceptible to very sharp spikes. An interesting result is that the quarterly process has the least amount of temporal dependencies, which can be seen directly through its first-order autocorrelations, where they are $0.44, 0.34, 0.18$, and 0.24 for weekly, monthly, quarterly, and annually respectively. For all plots, it is indisputable that none of them is constant, and they all exhibit behaviour that captures different periods of financial distress as expected, although the relative intensity between periods is captured differently.

In the next section, I will model the left jump tail index estimates by an ARIMA/GARCH model to determine the behaviour of the tail. The modelling will be based on the case where it is assumed that the left jump tail index is constant over a weekly basis as this is the most detailed version. In order to perform the analysis, it is crucial to understand the structure of the tail shape distribution. This paper will first cover the necessary definitions and theory for the time series employed in the analysis.

6.3 Notation and underlying time series theory

The exposition in this subsection largely follows the notes of ST422: Time Series at the London School of Economics in the year of 2019-2020 held by Professor Clifford Lam.
When working with discrete time series, there are a couple of definitions that must be introduced first. It is natural to start the topic of whether or not a time series is time-variant with the introduction of stationarity.

Definition 6.1 (Strictly stationary)**.** The time series $\{x_t\}$ is defined as being strictly stationary if for any h and time points $t_1, ..., t_m$ with $m > 0$, then

$$P(x_{t_1} \leq c_1, .., x_{t_m} \leq c_m) = P(x_{t_1+h} \leq c_1, ..., x_{t_1+h} \leq c_m) \tag{53}$$

holds true for any numbers $c_1, .., c_m$.

As such, it means that the probabilistic behaviour of every collection of the time series is identical to the time shifted set for any shift, and hence all the x_t's have the same distribution function F, the same marginal density f, if it exists, and is not time-variant. As it has the same density f, this means that the mean function exists and is time-invariant, such that $\mu_t = \mu_s$ for any s and t. As such, it is a constant. It also means that the autocovariance function of the process depends only on the time difference between t and s but not on the actual times themselves.

It is a relatively strict constriction to require, and it is often too strong an assumption to satisfy for actual data. Therefore weak stationarity comes into play.

Definition 6.2 (Weak stationarity)**.** A weakly/second-order stationary time series $\{x_t\}$ is such that

i. The mean $E(x_t) = \mu < \infty$;

ii. The autocovariance function $\gamma(s, t) = \text{cov}(x_s, x_t) = E[(x_s - \mu_s)(x_t - \mu_t)]$ depends on s and t only through their difference $|s - t|$. Moreover, $\text{Var}(x_t) < \infty$

Most strictly stationary processes are also weakly stationary process, except processes with infinite mean and variance. Often this is not a significant issue, but with extreme value theory, it can play an important part, as the mean of a Fréchet distributed random variable can be infinite dependent on the choice of parameters. The work on stationary and weakly stationary stochastic processes was also developed mainly by the Russian mathematician Khinthcine, just as with the Lévy-Khintchine Theorem for Lévy processes, in his paper Khintchine (1934).

When analyzing the stationarity of a time series, it is common to start with the autocovariance function, the autocovariance sequence (ACVS), and the autocorrelation function (ACF). Note that hereafter a stationary process means a weakly stationary process.

Definition 6.3 (Autocovariance function and sequence)**.** The autocovariance function of a stationary time series is denoted by

$$\gamma(h) = \mathrm{cov}(x_{t+h}, x_t) \tag{54}$$

for any h. For discrete and equal-spaced time series, the autocovariance sequence is defined as

$$s_\tau = \mathrm{cov}(x_{t+\tau}, x_t) \tag{55}$$

for any integer τ.

Which immediately leads to the ACF,

Definition 6.4 (Autocorrelation function)**.** The autocorrelation function (ACF) of a stationary time series is denoted by

$$\rho(h) = \frac{\gamma(t+h,t)}{\sqrt{\gamma(t+h,t+h)\gamma(t,t)}} = \frac{\gamma(h)}{\gamma(0)}. \tag{56}$$

The notation ρ_τ is used to denote the ACF for integer valued τ. Hence,

$$\rho_\tau = \frac{s_\tau}{s_0}. \tag{57}$$

Based on these two definitions one can define the most basic building block of complex time series which is the white noise process.

Definition 6.5 (White noise)**.** The time series $\{x_t\}$ is defined as being a white process iff the following holds true,

i. $E(x_t) = \mu < \infty$;

ii. $s_\tau = 0$ for $\tau \neq 0$.

The shorthand for saying $\{x_t\}$ is a white noise with mean μ and variance σ^2 is

$$x_t \sim WN(\mu, \sigma^2).$$

Hence white noise process is a collection of uncorrelated time series random variables with constant and finite mean. It is not necessarily a Gaussian white noise, but it can be a process with much fatter tails.

The two other main time series, the building blocks, are the moving average process of order q, MA(q), and the autoregressive process of order p, AR(p). These two functions can be combined both with each other but also with seasonal versions of them. The moving average process is a linear combination of white noise time series variables called a filtered white noise series.

Definition 6.6 (Moving average process)**.** A process $\{x_t\}$ is called a moving average process of order q, with shorthand $x_t \sim MA(q)$, if it can be written as

$$x_t = \mu + \theta_0 \varepsilon_t + \theta_1 \varepsilon_{t-1} + ... + \theta_q \varepsilon_{t-q}, \tag{58}$$

where μ and θ_j are constants with $\theta_q \neq 0$, and $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$.

The autoregressive process is defined as follows,

Definition 6.7 (Autoregressive process)**.** A process $\{x_t\}$ is an autoregressive process of order p, with shorthand $x_t \sim AR(p)$, if it can be defined as

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + ... + \phi_p x_{t-p} + \varepsilon_t, \tag{59}$$

where the ϕ_j's are constants with $\phi_p \neq 0$, and $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$.

For the classifications of time series into either an $AR(p)$, $MA(q)$, or a combined model, the use of the ACF and ACVS is very useful. It holds true that the ACVS and ACF for an $MA(q)$ process cuts of at q, and as such, the ACF can be used to select the parameter q if the process is a MA process. Nevertheless, for an $AR(p)$ process, the ACVS and ACF decays exponentially rather than cut off at a specific value, so the ACF can be used to indicate that it is a AR process, but not the parameter choice. As such, it would make sense to have a function that has a cutoff at p for the $AR(p)$ process. The partial autocorrelation function (PACF) serves this purpose.

Definition 6.8 (The Partial Autocorrelation function)**.** The PACF for a mean

0 stationary process $\{x_t\}$ is defined as

$$\begin{aligned}
&\pi(1) = \text{corr}(x_2, x_1)\\
&\pi(2) = \text{corr}(x_3 - E(x_3|x_2), x_1 - E(x_1|x_2))\\
&\pi(3) = \text{corr}(x_4 - E(x_4|x_3, x_2), x_1 - E(x_1|x_3, x_2))\\
&\text{etc...}
\end{aligned}$$

With this definition the PACF at lag j is the correlation of those portions of x_1, x_{j+1} which are unexplained by the intermediate variables $x_2, ..., x_j$.

Having defined the most central of the definitions and notation, the next section will be concerned with the actual time series analysis on the shape of the jump tails.

6.4 Time series analysis of the left jump tail shape

It is usually assumed that an economic time series will follow the additive model of time series components, such that it is made up of three components, the trend, the seasonality, and the regular component:

$$x_t = \mu_t + s_t + y_t, \tag{60}$$

where

$$\begin{aligned}
\mu_t &= \text{Trend component}\\
s_t &= \text{Seasonal/periodic component}\\
y_t &= \text{Regular/stationary component.}
\end{aligned}$$

The modelling of a time series to follow an ARMA process is typically done on the regular and stationary part $\{y_t\}$, after trends and periodic components are estimated and removed. Therefore, the trend component of the time series $1/\alpha_t^-$ will be investigated now on a weekly basis. To get an idea of the of the trend, it is natural to investigate the plot.

Figure 9 shows that the time series is not stationary as the mean is fluctuating over time. Note that some of the spikes in the earlier period and late 2010 might be due to a low amount of data in a specific week. The trend could either be a piecewise linear function or a higher-order polynomial, but as a high-order polynomial would lead to overfitting, it is not suited for this task. It is noteworthy that there is a debate about the appropriate choice between

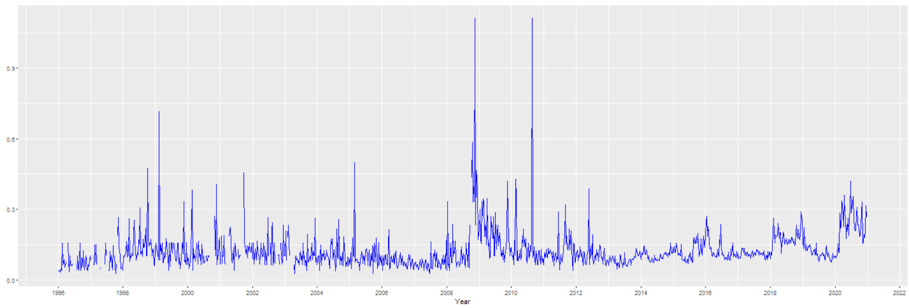


Figure 9: The estimates of $1/\alpha_t^-$ on a weekly basis.

differencing and detrending for financial time series as mentioned in Zhao & Wei (2003) and Bianchi et al. (1999). It is important to note that the dangers of a parametric approach, such as a polynomial, may cause misleading information and incorrect inference about the trend curve. For example, a sixth-order polynomial is significant at all estimates for this data, but it would indicate that the $1/\alpha_t^-$ would skyrocket in the coming years due to the elevated levels across 2020. Instead, a first-order differencing has been employed in order to remove any piecewise linear trend that is present, and the residuals look as following,

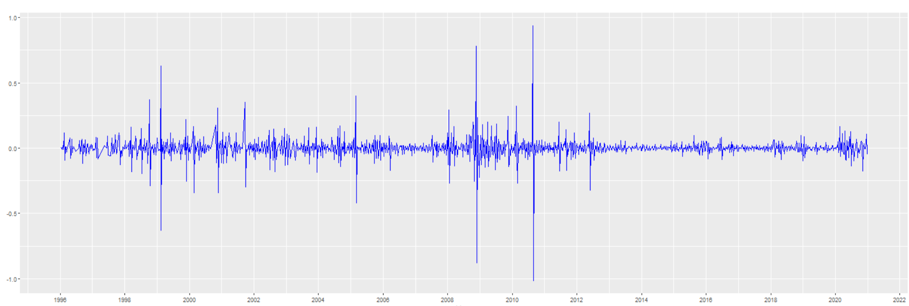


Figure 10: Residuals of a first order differencing on the estimates of $1/\alpha_t^-$ on a weekly basis.

It is now clear that any trend in the mean has been removed, and therefore, it satisfies half of the two conditions required for being called stationary. The second condition is less clear with differenced time series as the mean is relatively stable, but the autocovariance sequence might not only depend on the lag, which was the second requirement for weak stationarity in Definition 6.2. There is no clear seasonality present in the data, but that does not mean it is non-existent, and it is familiar with economic data to have a periodic component present. In order to detect an unknown periodic component, Frequency Domain Analysis as described in Koopmans (1995) will be used. It is assumed that the model is of the form

$$x_t - \mu_t = s_t + y_t = \beta \cos(2\pi f t) + \gamma \sin(2\pi f t) + \varepsilon_t, \tag{61}$$

which is a usual linear regression model with regression parameters β and γ, if f was known. However, it would not be interesting to perform this analysis as f is the target of interest. Instead, by following the method in the ST422 notes and fitting a saturated model, it can be shown that the estimated coefficients $\hat{\beta}_j$ and $\hat{\gamma}_j$ are essentially measures of covariance between the observations and a sinusoid oscillating at j cycles in T time points. As such, it can lead to a single measure for the presence for a frequency of oscillation of j cycles in T time points,

$$\begin{aligned} P(j/T) &= \hat{\beta}_j^2 + \hat{\gamma}_j^2 = \frac{4}{T} I(j/T), \\ I(j/T) &= \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} (s_t + y_t) \cos(2\pi t j/T) \right)^2 + \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} (s_t + y_t) \sin(2\pi t j/T) \right)^2, \end{aligned} \tag{62}$$

where the quantity $P(j/T)$ is called the scaled periodogram, while $I(j/T)$ is called the periodogram. The main conclusion from the periodogram is that if it is relatively "large" at frequency j/T compared to other frequencies, it means that the data has a high correlation with a particular oscillation at frequency j/T. In the application of frequency domain analysis, its similarity to the Fourier transformation, specifically the fast Fourier transformation as data is discrete, is applied as both models are equivalent and efficient algorithms for calculating the Fast Fourier Transformation exists. Figure 11 reveals that certain spike values stand out, especially the spike in the middle with frequency $f = 0.228$ corresponding to $1/f = 4.378$, which is roughly the number of weeks in a month which is 4.348.

For economic data collected on a weekly basis, it is relatively common to have a monthly seasonality. I have circled other choices of spikes that stand out, but especially for the higher frequency, corresponding to shorter periods, there is much subjectivity in the choice. The values selected here corresponds to, from left to the right: The appearance of especially $1/f = 13.3, 1/f = 4.4, 1/f = 2.1$ is expected as these represent the biweekly, monthly, and quarterly periods that are often represented in financial and economic time series. For example, see Sewell (2011) section on calendar effects for a more in-depth review.

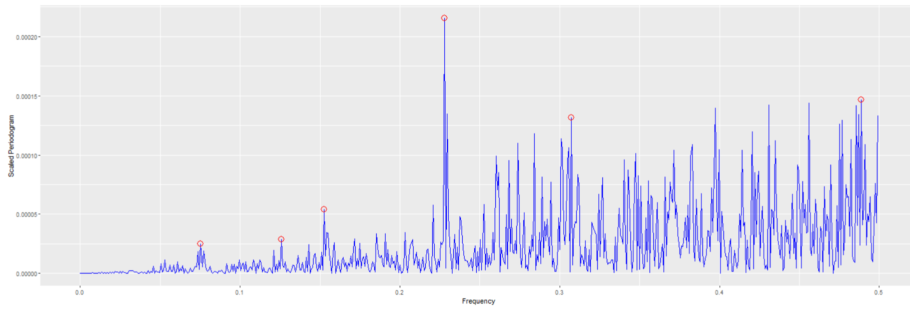


Figure 11: The scaled periodogram of the residual $1/\alpha_t^-$ series after a first order differencing.

f	$1/f$
0.075	13.255
0.126	7.936
0.153	6.558
0.228	4.387
0.307	3.253
0.489	2.046

Table 2: Frequencies of interesting from Figure 11.

In order to gain a more in-depth understanding of the rough peaks present, it is of interest to perform kernel smoothing. For the non-parametric kernel smoothing, it is of interest to choose the span of the kernel such that the bandwidth is sufficient to smooth the estimate but such that it does not remove any essential peaks. The kernels employed here are the modified Daniell kernel, defined such that for a span of $\{m : m \in \mathbb{N}_+\}$, the smoothing is

$$\hat{x}_t = \frac{x_{t-m} + 2x_{t-(m-1)} + ... + 2x_t + ... + 2x_{t+(m-1)} + x_{t+m}}{4m}$$

and the convoluted Daniell kernel. Let $I(j/T)$ denote the periodogram at frequency j/T. By employing a Daniell kernel with parameter m to smooth a periodogram, the smoothed value $\hat{I}(j/T)$ is a weighted average of the periodogram values for frequencies in the range $(j-m)/T$ to $(j+m)/T$.

On the two lower plots in Figure 12 the bandwidth is still too narrow to uniquely identify frequencies of interest. On the top plot, it is far more transparent, even though some subjectivity is included. One could discard the first top, or one could include the top at around $f = 0.425$. It is also clear that the smoothing has not removed the significant spikes included in Figure 11, but one effect of the smoothing is that the dominant peak in the unsmoothed version, the one that accounted for the monthly seasonality, is now significantly less distinct, as the height was so sharply defined in the unsmoothed version relative to its surrounding values.

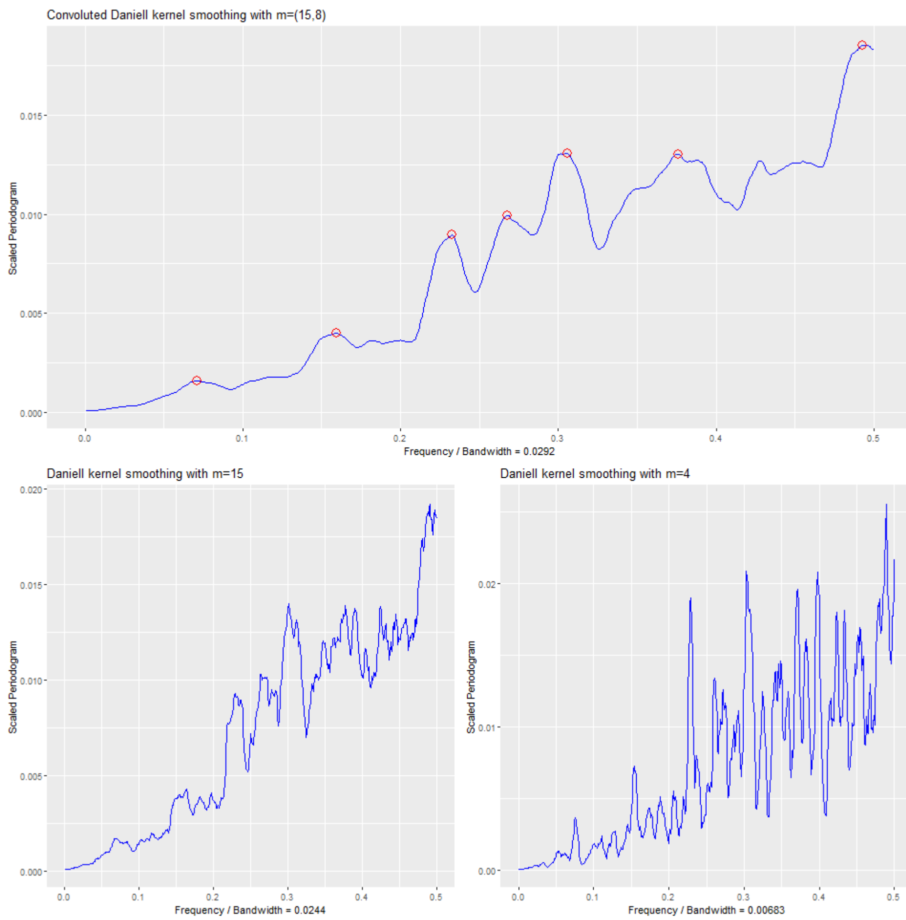


Figure 12: The smooth scaled periodogram of the residual $1/\alpha_t^-$ series after a first order differencing for three different choices of smoothing kernel.

When comparing 2 and 3 it becomes clear that most of the points of interest in the unsmoothed version have been preserved, which is a very desired trait, even if they have been moved slightly. The peak at $1/f = 7.936$ has been dropped, but instead, $1/f = 2.660$ and $1/f = 3.731$ have been included. These were also in the range of the unsmoothed periodogram where the volatility was extremely high, and it was inconvenient to determine which spikes were

f	$1/f$
0.071	14.045
0.159	6.281
0.233	4.296
0.268	3.731
0.306	3.272
0.376	2.660
0.493	2.029

Table 3: Frequencies of interesting from Figure 12.

relatively high compared to the rest.

It is also possible to make a parametric smoothing of the periodogram. It is built around the rigorous Spectral Representation Theorem, which is outside the scope of this thesis. However, it uses the fact that the periodogram is a sample estimate of the spectral density and that the spectral density of any stationary and invertible time series can be approximated by the spectral density of an AR model. The estimation in Figure 13 is based on the R function spec.ar from the Stats package in R.

f	$1/f$
0.071	14.056
0.114	8.754
0.154	6.481
0.188	5.309
0.229	4.358
0.269	3.724
0.305	3.283
0.347	2.884
0.381	2.626
0.444	2.253
0.491	2.029

Table 4: Frequencies of interesting from Figure 13.

By comparing 3 and 4 it is clear that the parametric smoothing has included all the peaks from the non-parametric smoothing, and on top of that, it has also made several other peaks visible. Now that there are two sets of frequencies of interest, from Figure 12 and 13, interest can return to the function in Equation 61 and fit it as a usual linear regression model and a model comparison.

An interesting result here is that for the non-parametric smoothing, just two frequencies are statistically significant at a 5% level which are the two

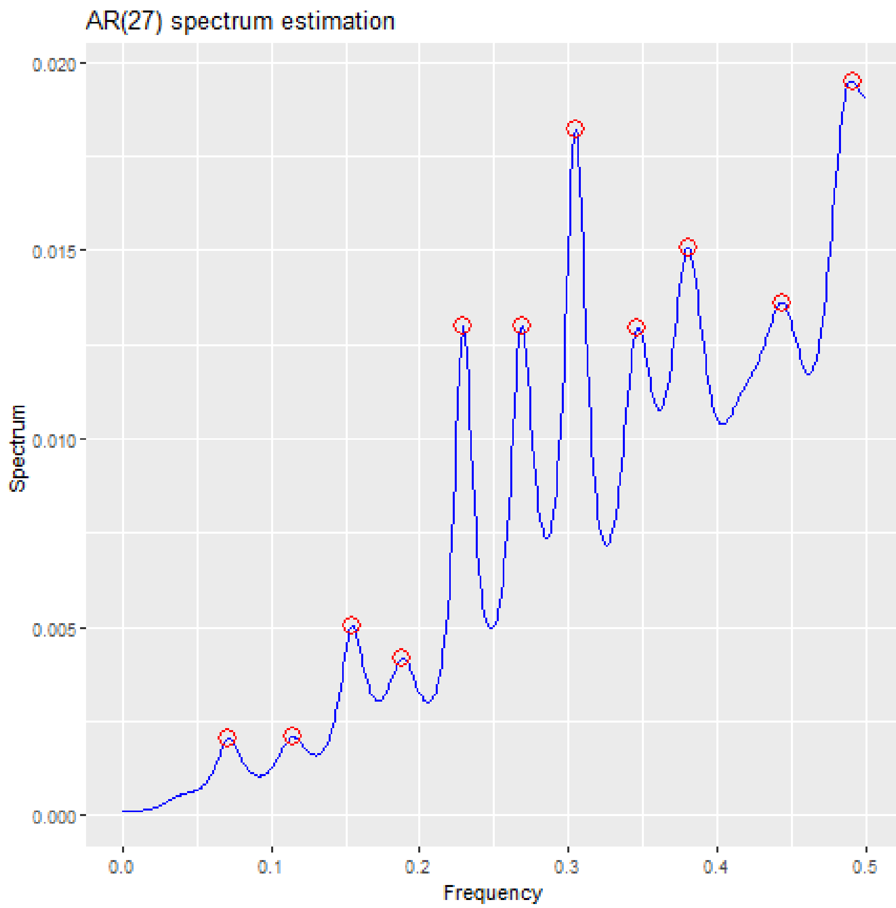


Figure 13: The estimated smoothed periodogram for the residual of the $1/\alpha_t^-$ series after a first order differencing. Estimated by an AR(27) time series.

frequencies at $f = 0.268$ and $f = 0.306$, corresponding to a period of roughly 3.3 weeks and 3.7 weeks. As such, the clear spike at the monthly level in the unsmoothed is not significant when the frequencies are based on smoothed periodogram done by a convoluted Daniell kernel smoothing with $m = (15, 8)$. On the other hand, the parametric smoothing leads to six significant frequencies at the 5% level. These are the $f = 0.154, f = 0.269, f = 0.305, f = 0.347, f = 0.381$, and $f = 0.491$ corresponding to periods of roughly 6.5, 3.7, 3.3, 2.9, 2.6, and 2.0 weeks, which also does not include the monthly level at 4.4 weeks. The non-parametric smoothing does not include this value due to the slight shifting towards the right caused by the upwards trend in the original periodogram or because the bandwidth is too high to capture the peaks of the unsmoothed periodogram.

The following set of linear regression models have been constructed to compare and select which frequencies are of interest. Let f_1 be the frequencies in Table 3 and f_2 be the frequencies in Table 4. Then,

$$x_i = \sum_{k=1}^{K_i} \beta_{k,i} \cos(2\pi f_{k,i} t) + \gamma_{k,i} \sin(2\pi f_{k,i} t), i = 1, 2, \tag{63}$$

where K_i is the total number of frequencies of interest in either periodogram. As most estimated β_k and γ_k turn out insignificant, four models have also been based on the significant variables, two based on a 5% level and two based on a 10% level.

$$\begin{aligned}
x_3 &= \gamma_{4,1} \sin(2\pi f_{4,1} t) + \beta_{5,1} \cos(2\pi f_{5,1} t) \\
x_4 &= \gamma_{4,1} \sin(2\pi f_{4,1} t) + \beta_{5,1} \cos(2\pi f_{5,1} t) + \gamma_{1,1} \sin(2\pi f_{1,1} t) \\
x_5 &= \beta_{1,2} \cos(2\pi f_{1,2}) + \beta_{8,2} \cos(2\pi f_{8,2}) + \gamma_{4,2} \sin(2\pi f_{4,2} t) \\
&\quad + \gamma_{7,2} \sin(2\pi f_{7,2} t) + \gamma_{9,2} \sin(2\pi f_{9,2} t) + \gamma_{10,2} \sin(2\pi f_{10,2} t) \\
x_6 &= \beta_{1,2} \cos(2\pi f_{1,2}) + \beta_{8,2} \cos(2\pi f_{8,2}) + \gamma_{4,2} \sin(2\pi f_{4,2} t) \\
&\quad + \gamma_{7,2} \sin(2\pi f_{7,2} t) + \gamma_{9,2} \sin(2\pi f_{9,2} t) + \gamma_{10,2} \sin(2\pi f_{10,2} t) \\
&\quad + \gamma_{6,2} \sin(2\pi f_{6,2} t) + \beta_{7,2} \cos(2\pi f_{7,2}).
\end{aligned} \tag{64}$$

In order to compare which of these six models to use, I've employed the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) to decide which of the models has the best trade-off between likelihood and number of parameters. It is noteworthy that even though all the parameters above are significant, at least at the 10% level, they explain a minimal amount of the variance in the first order differenced series of $1/\alpha_t^-$. The AIC and BIC turn out as,

	df	AIC	BIC
x_1	16	−2587.114	−2505.084
x_2	24	−2588.386	−2465.341
x_3	4	−2601.497	−2580.990
x_4	5	−2602.426	−2576.791
x_5	8	−2610.401	−2569.386
x_6	10	−2612.439	−2561.170

Table 5: Model selection criteria for smoothing frequencies.

Sadly, the two selection criteria are not agreeing on which model to use with the AIC preferring x_6, which is the model based on the parametric smoothing and all parameters that are significant at the 10% level, and the BIC prefers x_3, which is the most simple model from the non-parametric smoothing. This is due to the significant difference in degrees of freedom and how the BIC penalises more variables harder than the AIC.

As neither models include the clear monthly spike in the unsmoothed periodogram with its dominance in the original periodogram and the economic interpretation of having a monthly periodicity, I have decided to go forward with x_3 and the frequencies of interest herein, due to it being the simpler model and as such less prone to overfitting.

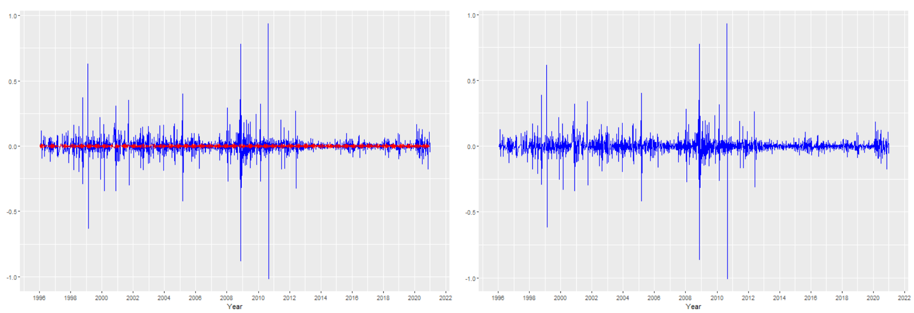


Figure 14: Left: The first order differenced $1/\alpha_t^-$ series and the estimated periodicity.
Right: The residual after deducting the periodicity from the first order differenced $1/\alpha_t^-$ series.

The results of this can be seen in Figure 14 where it is clear that even though the effect is statistically significant, it is also very minor in times of high volatility. In calmer periods, such as from 2013-2015 or in 2017, the volatility of the periodicity matches the volatility of the time series, but from the right plot, it is evident that it still explains a minor amount of the variance in these periods. The standard deviation has gone from 0.0854 to 0.0849, showing a minor drop in variance due to the extra part that is being explained.

It is also clear that the time series is still not a stationary time series as the volatility is fluctuating over time. It both exhibits volatility clustering, common with financial data, and greatly varying overall levels of tail decay rate. To confirm this suspicion, I will use the formal test for the whiteness of a series, the Ljung-Box-Pierce statistics, which is described in Ljung & Box (1978) &

Box & Pierce (1970), and is defined by

$$\mathcal{Q}^* = T(T+2)\sum_{j=1}^{k}\frac{\hat{\rho}_j^2}{T-j},$$

where $\hat{\rho}_j$'s are the ACF of the fitted residuals and T is the sample size. It will be compared with a χ^2_{k-m} where m is the number of parameters estimated in the fitted model. As such, $m = 3$ in this case from the parameters $\gamma_{5,1}, \beta_{6,1}$, and the intercept. It should be tested at a k that is reasonably large compared to m. It turns out that the choice of $\{k : k > m\}$ is irrelevant in this case as the p-value of the Ljung-Box-Pierce test is extremely small, confirming the non-whiteness of the residuals.

More importantly is that the regular part of the time series, as defined in (60), is still not a stationary time series after the trend and periodicity have been removed. This shows a need for time-variant tail shapes, and it cannot be assumed to follow an ARMA model even with non-Gaussian white noise. An alternative to this would be to use a periodically correlated ARMA model, also known as a PARMA model. These are outside the scope of this thesis, but the reader is referred to Franses et al. (1996) and Franses & Paap (2004), and the references therein.

An attempt can be made to model $1/\alpha_t$ by an ARIMA model, and it should be fairly accurate most of the time. To do this, it is common to start with the ACF and PACF of the differenced and periodically corrected time series in order to estimate the parameters of p and q. The danger here is that it assumes stationarity meaning that the ACF is only dependent on the time lag on not the starting positions. As such, it is necessary to assume stationarity for this method to work. Then the ACF and PACF are as follows, both for the start and for the entire period. It is immediately clear from the two top

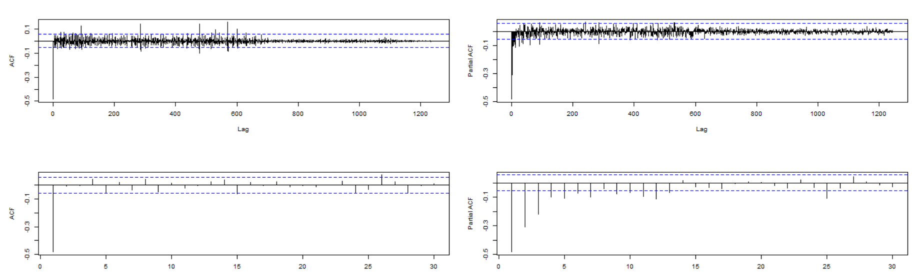


Figure 15: The ACF and PACF for the residual of the estimated model.

plots in Figure 15 that the assumption is flawed, but it also reveals a fascinating picture of how the autocorrelation between the initial position and a given lag decreases dramatically in size at around lag 700, which is in the middle of 2009 and the financial crisis. From the two bottom plots its clear that MA part of the process has a cutoff a $q = 1$. The cutoff for PACF is significantly less explicit, and one could argue that it is exponentially decreasing such that the process is $p = 0$. One could also argue that it drops dramatically off for values of $p \in \{1, 2, 3, 4, 7, 13\}$. To select which value of p is the most suited, the AIC and BIC are listed below. From Table 6 there are two possible candidates, either

p	AIC	BIC
0	−3175.37	−3165.03
1	−3176.64	−3161.13
2	−3178.69	−3158.01
3	−3184.96	−3159.10
4	−3195.16	−3164.13
7	−3189.48	−3142.93
13	−3200.25	−3122.68

Table 6: Model selection criteria values of p for $q = 1$ and $d = 1$.

$p = 4$ or $p = 13$. Here $p = 4$ seems like the most reasonable choice given it is the second-highest in AIC, where $p = 13$ is the worst model based on BIC due to its higher penalty to more complex models. As such, the most fitting model under the assumption of stationarity is an ARIMA(4,1,1).

One can also model an ARIMA model automatically through several functions and packages in R. One example of this is the auto.arima function in the package Forecast. This method proposes an ARIMA(5,1,0) model with $AIC = -3129.11$ and $BIC = -3098.08$, and as such, it is a significantly worse fit than what could be deducted from the ACF and PACF plots in 15.

In order to improve this model further, it is possible to model the volatility through a GARCH model on top of the ARMA model, such that the ARMA model models the mean and the GARCH model models the volatility. The ARIMA/GARCH has proven itself useful in a range of fields and can be combined with more advanced versions of GARCH, where Mohammadi & Su (2010) shows applications in oil price dynamics. This thesis is limited to the standard GARCH.

GARCH was proposed in Bollerslev (1986) as an improvement to the ARCH model proposed four years earlier in Engle (1982), which was the first systematic framework for volatility modelling. There are roughly four known properties of volatility seen in asset returns:

i. Volatility clustering. Volatility of asset returns tend to be high for a certain period of time, and low for other periods.

ii. Continuity. Volatility evolves continuously.

iii. Boundedness. Volatility varies within some fixed range. Hence it is usually stationary.

iv. Leverage effect. Volatility tends to react differently to big price increase or a big drop.

ARCH was an attempt to model these traits, but the volatility clustering of financial time series is more persistent than what an ARCH model can capture, and what can be seen in Figure 14. It was defined as follows,

Definition 6.9 (ARCH)**.** The autoregressive conditional heteroscedastic model of order p, or $ARCH(p)$, is defined by (shorthand $x_t \sim ARCH(p)$)

$$x_t = \sigma_t \varepsilon_t, \sigma_t^2 = \alpha_0 + \alpha_1 x_{t-1}^2 + ... + \alpha_p x_{t-p}^2, \tag{65}$$

where $\varepsilon \sim IID(0,1)$, with $\alpha_0 > 0$ and $\alpha_i \geq 0$ for $i \geq 0$.

GARCH overcame this issue by making the volatility dependent on its own lags on top of the lagged values of the underlying process. A significant issue with the GARCH model is that it did not address the leverage effect problem. This is not an issue here, as the focus is purely on negative jumps. The GARCH model is defined as,

Definition 6.10 (GARCH)**.** The generalised ARCH model of order p, q or GARCH(p, q) is defined by (shorthand $x_t \sim GARCH(p, q)$)

$$x_t = \sigma_t \varepsilon_t, \sigma_t^2 = \alpha_0 + \alpha_1 x_{t-1}^2 + ... + \alpha_p x_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + ... + \beta_q \sigma_{t-q}^2, \tag{66}$$

where $\varepsilon \sim IID(0,1)$, with $\alpha_0 > 0$ and $\alpha_i \geq 0, \beta_j \geq 0$ for $i > 0$ and any j, and $\sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) < 1$.

Model selection in the GARCH model is relatively close to model selection in the ARMA model. Instead of analysing the ACF and PACF of the underlying time series, one should analyse the ACF and PACF of the squared time series. These are plotted below. From 16 it is clear that a suitable choice is either the GARCH(1,1) or GARCH(4,4). Again, the choice is made by comparing the information criterion statistics under the assumption that the residuals are Gaussian. The fit is done through garchFit in the package fGarch.

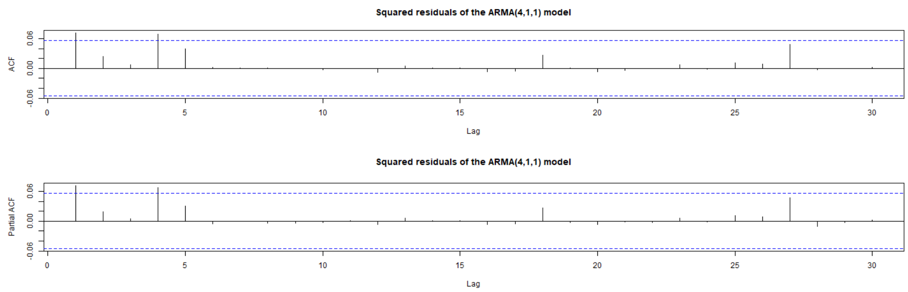


Figure 16: The ACF and PACF for the squared residuals of the estimated model.

(p,q)	AIC	BIC
$(1,1)$	-3640.716	-3599.701
$(4,4)$	-3648.420	-3576.644

Table 7: Information criterion statistics for (p,q).

Again the best model is ambiguous, dependent on whether AIC or BIC is preferred. There is, however, a significant issue with both models. The Jarque-Bera test and Shapiro-Wilk test on the standardised residuals are both rejected, and as such, it is not suited to model it according to a Gaussian distribution. As the tails are not heavy enough, it seems logical to try modelling the distribution of ε_t according to a standard Student's t with 5 degrees of freedom or a skew-standard Student's t with 5 degrees of freedom. This yields the results in Figure 8.

(p,q)	Skewed	AIC	BIC
$(1,1)$	N	-4344.208	-4298.066
$(1,1)$	Y	-4513.052	-4461.783
$(4,4)$	N	-4344.098	-4267.195
$(4,4)$	Y	-4522.584	-4440.554

Table 8: Information criterion statistics for (p,q) and whether ε follows a standard Student's t or a skew-standard Student's t distribution.

For all four models, the Ljung-Box statistics, lowest p-value is 0.85 dependent on lag, and the LM Arch test, with a p-value of 0.99, does not reject the null hypothesis for uncorrelated standardised residuals. This is a critical assumption for a GARCH-type model, and as such, it is a must that it is satisfied. The choice between the two models where ε_t follows a skew-standard Student's t is again ambiguous. By analysing the estimated parameters for the $ARMA(4,1,1) + GARCH(4,4)$, it becomes evident that a majority of the parameters are insignificant, and several parameters have identical estimates. As such, the $ARMA(4,1,1)+GARCH(1,1)$ is the most fitting for this data set, and the final best model is as follows:

$$\begin{cases} x_t & = y_t^{**} + y_{t-1}^{**} \\ y_t^{**} & = 1.758 * 10^{-4} + 6.7 * 10^{-3} \sin(0.268 * 2\pi t) \\ & -1.107 * 10^{-2} \cos(0.306 * 2\pi t) + y_t^* \\ y_t^* & = 1.725 * 10^{-1} y_{t-1}^* - 5.028 * 10^{-3} y_{t-2}^* + 1.142 * 10^{-1} y_{t-3}^* \\ & +4.765 * 10^{-2} y_{t-4}^* + y_t - 9.055 * 10^{-5} y_{t-1} \\ y_t & = \sigma_t \varepsilon_t \\ \sigma_t^2 & = 1.407 * 10^{-4} + 2.570 * 10^{-1} y_{t-1}^2 + 7.828 * 10^{-1} \sigma_{t-1}^2 \\ \varepsilon_t & \sim T_5(3.022, 1.603), \end{cases} \tag{67}$$

where $T_\kappa(\nu, \xi)$ is the skew-standard Student's distribution with shape parameter ν, skewness parameter ξ, and κ degrees of freedom.

In order to test the accuracy of the ARMA$(4,1,1)$+GARCH$(1,1)$ estimation above, Figure 17 shows the 1-day rolling forecasts for the entire period and for 2019-21. The rolling forecast is calculated using ugarchforecast from the package rugarch.

Figure 17 shows that the ARMA/GARCH model is well suited to capture the first order differenced series in times of lower volatility. It is also decently suited at capturing periods of higher volatility, but it lacks in times where there is a sudden shift, such as a jump. From the bottom plot and especially the top plot it is apparent that it is noteworthy better at capturing significant negative shifts than positive jumps, especially during the financial crisis in 2008-09, where it consistently underestimates the actual values in the most drastic weeks.

With an ARIMA/GARCH model ready for the jump tail shape, the next section will be concerning the overall jump intensity.

6.5 The jump intensity parameter

Having modelled and reviewed the importance of the tail shape parameter, it is time to move on to the second parameter in (34), videlicet, the jump intensity process governed by ϕ_t. Instead of just looking at ϕ_t alone, it is also of interest to look at the left jump intensity, as the values of ϕ_t are minimal in periods with few very sharp spikes. It is crucial to remember that the Lévy measure $\nu_t^{\mathbb{Q}}(dx)$ captures the jump intensity for jumps of size x. Therefore, to look at all

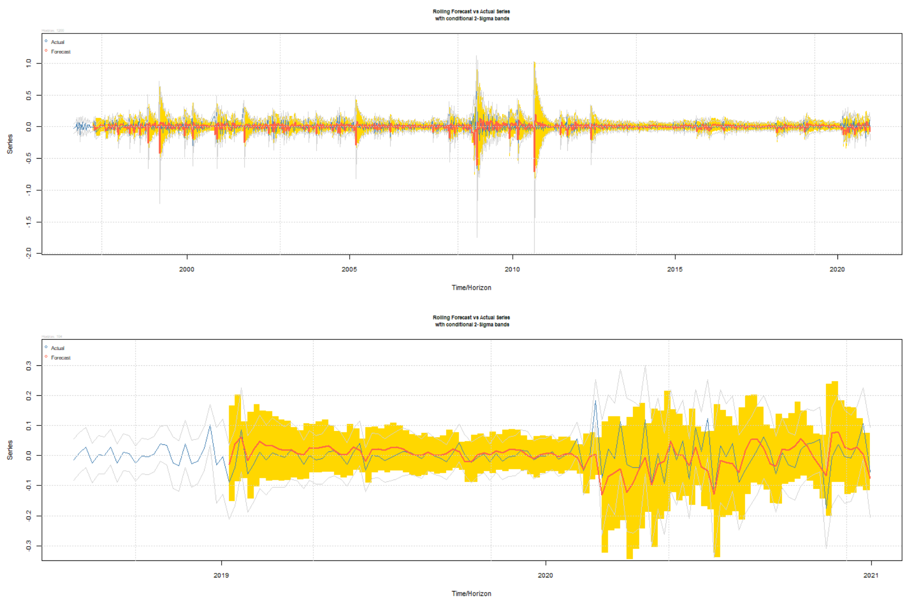


Figure 17: The one-period rolling forecast based on the ARIMA/GARCH model in (67).

significant jumps, such that $x < -|k_t|$ and where k_t is a time-variant limit for what is deemed significant, one should integrate $\nu_t^{\mathbb{Q}}(dx)$ on that range. Since $\nu_t^{\mathbb{Q}}(dx)$ as described in (34) behaves nicely, the integral can be computed as the Riemann Integral. Let LJI_t describe the left jump intensity for large jumps implied by the estimates $\hat{\alpha}_t^-$ and $\hat{\phi}_t^-$. Then,

$$LJI_t = \int_{x<-|k_t|} \nu_t^{\mathbb{Q}}(dx) = \int_{x<-|k_t|} \hat{\phi}_t^- e^{-\hat{\alpha}_t^- |x|} dx = \frac{\hat{\phi}_t^-}{\hat{\alpha}_t^-} e^{-\hat{\alpha}_t^- |k_t|}. \tag{68}$$

This calculation necessity a choice for what is a significant jump. The choice of the cutoff k_t is decisive in what constitutes a significant jump and when the start of the jump tails is. As there are periods with high volatility, it is clear that k_t should relate to Black-Scholes ATM volatility at time t in order to ensure that it is a relatively large jump instead of an absolute significant jump.

Several choices for the tail cutoff will be tested below to determine which one to go on with and of how great importance the value is. A natural baseline value is to use 3.5 times the time-normalized Black-Scholes ATM volatility at time t, as this was the choice for which puts are included in the analysis as covered in Section 6.1. The other choices are 6.868, as proposed in Bollerslev et al. (2015), $1500/925 * 6.868$, and $\log(x_t)/\log(925) * 6.868$, where x_t is the median strike price for the deepest OTM puts in each weak, 925 is the average strike price of the deepest OTM puts in the period in Bollerslev et al. (2015), and 1500 is the average strike price of the deepest OTM puts in the period in this analysis. The same specific cutoff will also be employed to calculate the critical measure in (17). Three of the four plots exhibit roughly the same

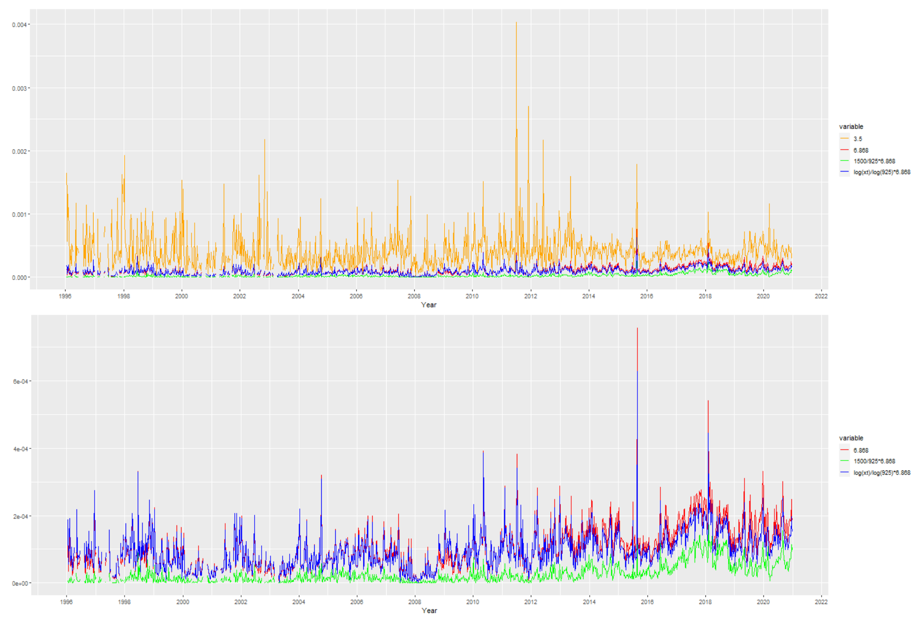


Figure 18: Plots of the estimated left jump intensities for large jumps beyond varying thresholds as defined in (68). The two plots are identical apart from the exclusion of the orange plot in the bottom plot.

dynamics, and it is clear that the jump intensity is monotonic with regard to the threshold. This is required as,

$$\int_{x<-|k_t^{(1)}|} \nu_t^{\mathbb{Q}}(dx) \leq \int_{x<-|k_t^{(2)}|} \nu_t^{\mathbb{Q}}(dx) \text{ for } k_2^{(2)} < k_2^{(1)}.$$

One can think of it as the jump intensity for jumps with sizes in the range $[0, 1]$ must be lower than the jump intensity for jumps with a range of $[0, 2]$, as the first is a subset of the second.

Additionally, the orange plot in the top panel displays a drastically different dynamic and order of magnitude than the others. This plot is based on 3.5 times the ATM Black-Scholes implied volatility, which was the same limit imposed on the data used. The significant difference here is that in the data estimation, the interest was on removing the continuous component. In this case, it is of interest to focus only on large jumps, and as such, the limit should be higher than 3.5, and the orange plot can be discarded.

There is a common issue in the bottom plots with the red and green line since they both relate to a constant value. Throughout the period, there has been significant growth in the strike price of the deepest OTM puts, as the price of S&P 500 has gone up dramatically. This is an apparent reason why the left jump intensity is at such high levels in more recent years compared to earlier in the period. Mainly the green plot exhibits very low jump intensities for the early period because the large jump threshold is relatively higher. This issue is improved with the blue line, which increases the intensity of the red line in the earlier years and decreases the intensity in the later years.

As the critical factor in the upcoming regression is the dynamics of LJI and not the order of magnitude, then all three models could be employed. This is easily shown by looking at the correlations matrix. Let x_1, x_2, and x_3 denote the red, green, and blue, respectively. Then the correlations matrix is, All

	x_1	x_2	x_3
x_1	1	0.86	0.97
x_2	0.86	1	0.79
x_3	0.97	0.79	1

Table 9: Correlation matrix for varying thresholds for large jumps.

correlations are very high, with x_1 and x_3 being nearly perfectly correlated, which is expected since x_3 is equal to x_1 times a minor adjusting factor. As such, the threshold for x_3 will be employed going forward.

6.6 The estimated jump tail variation

With the tail parameters and the level shift modelled and reviewed, the next section will be regarding the last step in order to estimate the measure for the fear component in (21). The left jump variation, $LJV_{t,\tau}$, will be modelled according to (17). It can be calculated directly from the integral under the assumptions that the measures ϕ_t^- and α_t^- are constant over the time horizon

$[t, t+\tau]$. Then the integral follows,

$$
\begin{aligned}
LJV^{\mathbb{Q}}_{[t,t+\tau]} &= \int_t^{t+\tau} \int_{x<-k_t} x^2 \nu_s^{\mathbb{Q}}(dx)ds \\
&= \hat{\phi}_t^- \int_t^{t+\tau} \int_{x<-k_t} x^2 e^{-\hat{\alpha}_t^-|x|} dx ds \\
&= \hat{\phi}_t^- \int_t^{t+\tau} \frac{1}{(\hat{a}_t^-)^3} e^{-\hat{a}_t^-|k_t|}(\hat{a}_t^- k_t(\hat{a}_t^- k_t + 2) + 2) ds \\
&= \frac{\hat{\phi}_t^- \tau}{(\hat{a}_t^-)^3} e^{-\hat{a}_t^-|k_t|}(\hat{a}_t^- k_t(\hat{a}_t^- k_t + 2) + 2), \qquad (69)
\end{aligned}
$$

where the last equation holds true due to the assumption that the parameters are constant on the interval $[t, t+\tau]$. The estimates for the weekly left $\mathbb{Q}$

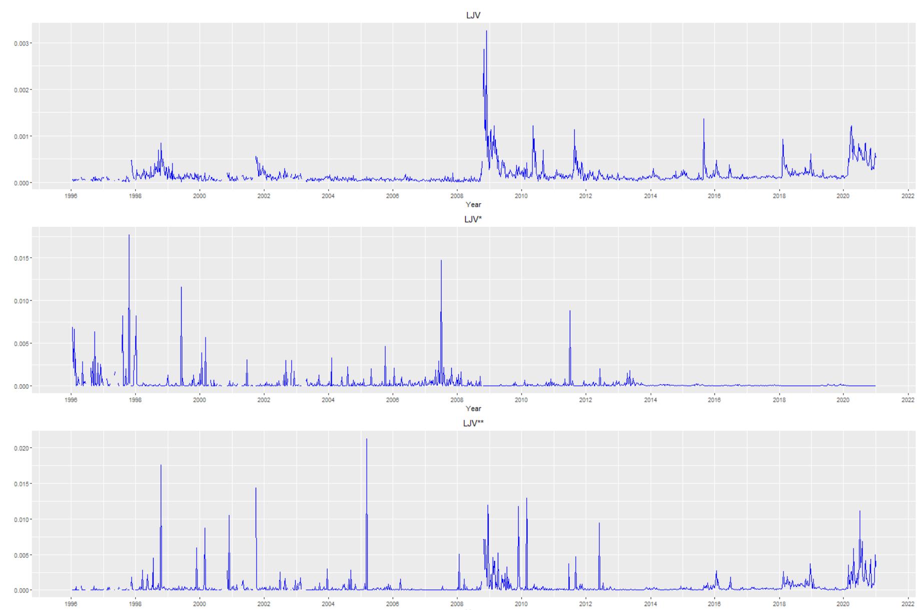


Figure 19: This figure plots the left jump variation, LJV_t, defined in (69). The middle panel plots the left jump variation LJV_t^* obtained by restricting the shape parameter α^- to be constant, but allowing ϕ_t^- to be time-variant. The bottom panel shows LJV_t^{**} obtained by restricting the level parameter ϕ^- to be time-invariant, but allowing α_t^- to be time-varying.

jump variation measure, as implied by (172), is depicted in Figure 19. Starting with the full time-variant LJV_t in the top panel, it is clear that this measure shares many of the same key dynamics as the estimates for tail shape, displayed in Figure 9, which is also why the modelling of the tail shape in Section 6.4 was more detailed than the modelling of the level parameter. One significant difference between the estimate for $1/\alpha_t^-$ and LJV_t is that LJV_t is less prone to jumps, such as in 1999 and 2010 in Figure 9. The relative variance in calm periods is also noticeably lower.

The middle and bottom plot in Figure 19 helps to underscore the great importance of having time-variant parameters over time-invariant parameters. The estimates in LJV_t^* are constructed by letting $\alpha^- = \bar{\alpha}_t^-$. This restriction was fairly common in the earlier literature, but it is clear that it mutes the temporal variation greatly. The jumps captured in LJV_t^* are also of less economic interpretation than in LJV_t, apart from the dot-com crash. There is no significant variation under the financial crisis, and the period following 2014, corresponding to when the number of outstanding options grew greatly, has been completely calm.
In contrast to this, it is more reasonable to assume that the level parameter is constant, such that ϕ^- is the median of ϕ_t^-, which is plotted in the bottom plot. By restricting the temporal variation to be solely driven by the shape of the jump tails, one gets a measure that is highly correlated with LJV_t but with far more frequent spikes and even more dramatic increases in magnitude. The increase in magnitude also holds true for LJV_t^*.

Before proceeding to the next section, a final step is to control that (52) holds true. The minimum value of α_t^- on a weekly basis is 0.9, which was on the week ending on 2010-08-30, such that it is strictly positive, and from Figure 18 it is clear that ϕ_t^- is finite at all times. Since the focus is on the negative jumps, it is assumed that $\phi_t^+ = 0 \forall t$ and therefore (2.4) holds true, and the estimated left jump measure is a Lévy measure.

The thesis will proceed to the next section, which is focused on the return predictability of the left jump tail measures across a range of portfolios. The results will also be compared with the VIX, another common fear proxy.

7 Return predictability

Table 10-12 reports the summary statistics for weekly returns, estimated jump tail parameters, and jump tail variation measures. The data is from January 1996 to the end of December 2020. All measures are recorded at the end of the week based on the Interpolation_Scheme in the appendix for the code. The SMB and HML portfolios are based on the Fama-French factor model in Fama & French (1993), but for the HML, I have used the HML Devil adjustment as proposed in Asness & Frazzini (2013). The UMD portfolio is from the adjusted three-factor model proposed in Carhart (1997), the QMJ is from Asness et al. (2019), and BAB is from Frazzini & Pedersen (2014). Returns are in weekly percentage form. All of the variation measures are in annualized percentage form. The sample correlation between the aggregate market portfolio for the United States and the different tail shapes and measures are all negative, apart from α_t^-, since this has an inverted relationship to the other measures. This effect confirms that the leverage effect is still present in the U.S. market. The leverage effect is the main shortcoming of the GARCH model in explaining the variance for the S&P500.

The contemporaneous in Table 11 between the tail variation measures and the weekly return on small-minus-big (SMB), high-minus-low, and betting-against-beta (BAB) are all negative or close to zero, but of a smaller magnitude. The correlations for the up-minus-down, or the so-called momentum factor (UMD), are all very close to zero but slightly positive. The factor that stands out the most is the quality-minus-junk factor (QMJ), where the contemporaneous correlations are of the same magnitude as the correlations for the market portfolio but with an opposite sign indicating an inflow into quality stocks in times of distress.

The picture for the correlations between the jump tail variation measures and the subsequent week, as reported in Table 12, is quite different. All the correlations for the aggregate market portfolio are now quite positive, while the correlations for small-minus-big, high-minus-low, and momentum all fluctuate around zero. One could expect that quality-minus-junk would have been negative, such that it mirrored the market portfolio again, but that is not the present picture. Instead, betting-against-beta is negatively correlated. This effect, especially on the market portfolio, indicates a volatility feedback effect, where the market overreacts to the variation measure by an immediate drop in order for a higher return in the following week as compensation for the higher risk. This would penalize the betting-against-beta strategy, since it is long stocks with a low correlation to the systematic risk and short stocks with a high correlation,

	MKT	SMB	HML	UMD	QMJ	BAB
Mean	0.16	0.03	0.00	0.09	0.09	0.16
Std. dev	2.47	1.19	1.88	2.34	1.28	1.75
Skewness	-0.67	-0.02	0.90	-1.15	0.20	-0.49
Kurtosis	8.18	5.52	11.56	11.33	7.53	7.68
Max	12.88	5.26	13.68	13.16	7.38	9.25
Min	-18.26	-7.29	-9.35	-16.41	-8.10	-10.82
AR(1)	-0.03	0.08	0.04	0.03	0.05	0.07
	α_t^-	LJI_t	LJV_t	LJV_t^*	LJV_t^{**}	
Mean	10.16	0.49	0.80	1.46	2.46	
Std. dev	4.78	0.29	0.11	0.53	0.76	
Skewness	1.45	1.64	6.14	10.06	7.38	
Kurtosis	7.02	11.09	64.9	133.51	73.34	
Max	38.75	3.27	16.9	91.99	110.3	
Min	0.90	0.00	0.01	0.00	0.00	
AR(1)	0.43	0.54	0.67	0.14	0.26	

Table 10: Univariate statistics.

and especially the high beta stocks should be rewarded in this manner. Nevertheless, it is hard to conclude from sample correlations whether or not the higher returns are associated with an increase in systematic risk or a change in attitude towards risk.

	MKT	SMB	HML	UMD	QMJ	BAB	α_t^-	LJI_t	LJV_t	LJV_t^*	LJV_t^{**}
MKT	1.00	0.20	0.13	-0.21	-0.59	-0.38	0.04	-0.08	-0.18	0.00	-0.04
SMB		1.00	0.06	-0.13	-0.39	-0.28	0.02	-0.02	-0.06	0.04	-0.04
HML			1.00	-0.78	-0.30	-0.19	0.00	-0.03	-0.10	-0.02	-0.06
UMD				1.00	0.31	0.43	0.03	0.01	0.02	0.03	0.04
QMJ					1.00	0.42	-0.07	0.01	0.17	-0.02	0.09
BAB						1.00	0.03	0.02	-0.05	0.01	0.00
α_t^-							1.00	-0.17	-0.47	0.61	-0.39
LJI_t								1.00	0.34	-0.05	0.00
LJV_t									1.00	-0.12	0.47
LJV_t^*										1.00	-0.08
LJV_t^{**}											1.00

Table 11: Contemporaneous correlations.

Another standard proxy for the fear factor is the VIX^2, where the VIX offers an approximation to the risk-neutral expectation of the total quadratic variation. Therefore, it is natural to compare our fear proxy to the VIX^2 process. In Figure 20, the two plots look fairly symmetrical at first glance. There are, however, a couple of significant dissimilarities. The LJV_t remains more steady in calmer times, where the jump variance is relatively insignificant. The VIX_t varies more here, as it also captures the continuous variance. This

	α_t^-	LJI_t	LJV_t	LJV_t^*	LJV_t^{**}
MKT	-0.02	0.05	0.11	0.00	0.03
SMB	-0.02	-0.01	0.02	-0.01	0.04
HML	-0.02	-0.05	0.01	-0.04	-0.01
UMD	0.04	0.02	-0.05	0.03	-0.05
QMJ	0.00	-0.02	-0.02	0.01	-0.01
BAB	0.04	0.02	-0.11	0.00	-0.09

Table 12: One-week-ahead return correlations.

is especially clear up to the start of the financial crisis in 2008-09, where LJV_t jumped straight up to a very elevated level, but the VIX^2 was on an increasing trend since 2007. Another interesting and highly relevant difference is how they reacted to Covid, where the effect on the LJV_t has been far more persistent than the effect on the VIX_t even though the effect on the VIX was more dramatic initially. Both the VIX_t and LJV_t are reported on a monthly scale, which will also be employed in the following section due to it preserving the dynamics of the weekly estimations but offering a more stable estimation.

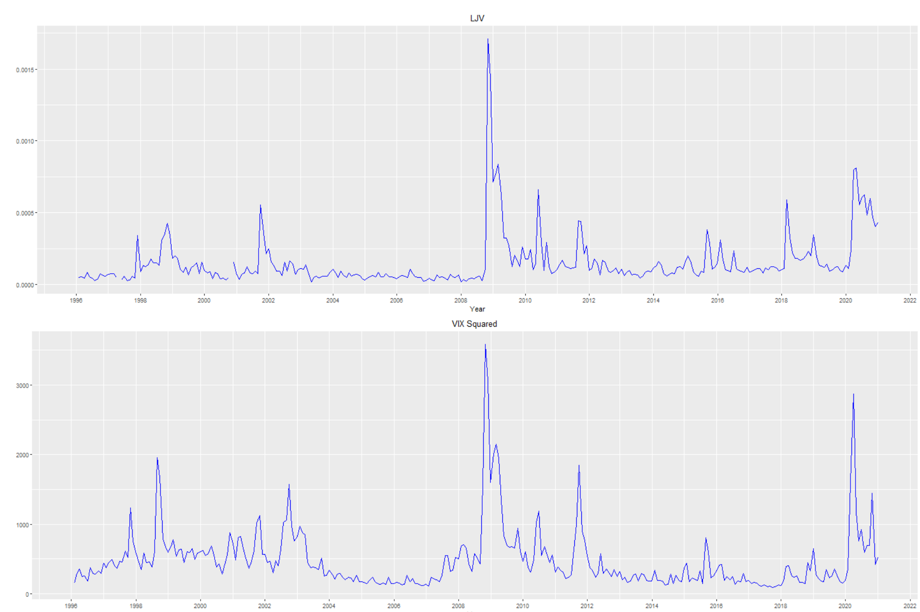


Figure 20: The two proxies for the fear component. Both series are plotted at a monthly frequency, and span the period from January 1996 till end of December 2020. The top panel shows the estimated left jump tail variation measure LJV_t and the second panel shows the CBOE VIX_t^2 volatility index.

The following section will be regarding predictability on the aggregate US aggregate market return.

7.1 US aggregate Market return predictability

Let the continuously compunded return from time t to $t+\tau$, say $r_{[t,t+\tau]} = \log X_{t+\tau} - \log X_t$, implied by the formulation in (11) may be expressed as,

$$r_{[t,t+\tau]} = \int_t^{t+\tau} (a_s + q_s)ds + \int_t^{t+\tau} \sigma_s dW_s + \int_t^{t+\tau} \int_{\mathbb{R}} x\tilde{\mu}^{\mathbb{P}}(ds, dx), \qquad (70)$$

where the unit interval corresponds to one month. The return linear regression may then be expressed as,

$$r_{[t,t+\tau]} = a_h + b_h V_t + \varepsilon_{t,t+h}, \qquad (71)$$

where V_t is a vector of the variation measures. The standard Newey-West t-statistics, with a lag length equal to twice the return horizon, are reported in the parentheses in Table 13.

Neither of the regressions is making a very good fit at predicting returns, which is expected. Especially at the three-months and 12-months, none of the variation measures is significant at the estimation. A Wald test for the 1-month nested models of $V_t = \alpha_t^{-*}$ vs $V_t = (LJV, VIX^{2*}, \alpha_t^{-*})$ is insignificant with a p-value of 0.80. The same is not true for the 6-month case comparing $V_t = LJV_t$ with $V_t = (LJV, VIX^{2*}, \alpha_t^{-*})$ or $V_t = (LJV, \alpha_t^{-*})$. Both of the nested models are significant in the Wald test over the simpler model, with $V_t = LJV_t$, with a p-value of 0.028 and 0.011 correspondingly. Comparing $V_t = (LJV_t, VIX^{2*}, \alpha_t^{-*})$ and $V_t = (LJV_t, \alpha_t^{-*})$ the p-value is 0.402 showing that VIX^{2*} is not needed. The simpler model for the 6-month, where $x_t = -0.0138 + 55.769 LJV_t + 0.126\alpha_t^{-}$ and all estimates are significant at a 5% level, is the most efficient model at predicting returns the aggregate market.

From Table 11-12 it was clear that there was also some prediction correlation with especially the betting-against-beta and quality-minus-junk portfolios. Therefore the one- and six-month return predictability regressions for QMJ and BAB are listed below in Table 14.

The results herein paint an interesting picture reinforcing the importance of the prediction horizon. For the one-month horizon, both LJV_t and VIX_t^{*2} are significant at the prediction at a 5% level, but VIX_t^{*2} can explain far more of the variance compared to LJV_t. By analyzing the Wald scores for $V_t =$

	1-month				3-months			
Constant	0.004	-0.001	0.018	0.013	0.005	-0.003	0.003	-0.011
	(1.102)	(-0.056)	(3.700)	(1.079)	(1.449)	(-0.399)	(0.439)	(-1.277)
LJV	18.767			2.212	8.527			-5.581
	(0.702)			(0.080)	(0.384)			(-0.183)
VIX^{2*}		0.184		0.099		0.246		0.330
		(0.662)		(0.343)		(1.141)		(1.274)
α_t^{-*}			-0.105	0.041			0.040	0.050
			(-2.744)	(-2.474)			(0.894)	(1.099)
R^2	0.566	0.400	1.730	2.015	0.113	0.707	0.242	1.162
	6-months				12-months			
Constant	0.000	-0.003	0.000	-0.008	0.006	0.004	0.006	0.003
	(0.092)	(-0.369)	(0.023)	(-0.936)	(1.457)	(0.925)	(1.774)	(0.706)
LJV	43.384			67.662	8.386			10.709
	(2.955)			(2.966)	(0.922)			(0.570)
VIX^{2*}		0.252		-0.193		0.058		0.000
		(1.188)		(-0.851)		(0.646)		(-0.001)
α_t^{-*}			0.067	0.126			0.008	0.018
			(2.394)	(3.355)			(0.222)	(0.459)
R^2	2.775	0.739	0.688	5.107	0.090	0.036	0.010	0.144

Table 13: The table reports one- to 12-month return predictability regressions for the aggregate market portfolio.

$(LJV_t, VIX^{2*}, \alpha_t^{-*})$, $V_t = (LJV_t, VIX^{2*})$, and $V_t = (VIX^{2*})$ it becomes clear that $V_t = (LJV_t, VIX^{2*})$ is the most suited at prediction the one-month leading returns for the betting-against-beta portfolio with a p-value of 0.045 over $V_t = (VIX^{2*})$, confirming that the additional parameter is significant at 5%, and p-value of 0.29 for using the additional parameter α_t^-. On the six-month horizon, none of the variables are significant at predicting the return.

The importance of the prediction horizon is the opposite for the quality-minus-junk portfolio. None of the parameters are significant at predicting the one-month return, and all the R^2 are by far the smallest of any prediction in this thesis. On the other hand, the six-month regressions can explain the highest amount of variance of any reviewed in this thesis. Both $V_t = (LJV_t)$, $V_t = (LJV_t, \alpha_t^{-*})$, and $V_t = (LJV_t, VIX^{2*}, \alpha_t^{-*})$ are highly significant with $V_t = (LJV_t, \alpha_t^{-*})$ being the preferred one according to the Wald test, showing that a significant portion of the return can be explained through six-months leading left jump variation and jump tail shape. By comparing VIX_t^2 and LJV_t, it also becomes clear that the market treats the risk of large jumps specially and, apart from the one-month prediction for BAB, large jump variance measures are more suited at predicting returns compared to continuous variance measures and the variance of small jumps.

Nevertheless, the left tail jump variation displays different dynamics in calm

	1-month			BAB		6-months		
Constant	0.011	0.031	0.001	0.030	0.008	0.011	0.014	0.019
	(3.546)	(6.202)	(0.159)	(4.313)	(2.025)	(2.316)	(2.293)	(2.391)
LJV	-26.412			21.223	-8.915			-15.758
	(-2.200)			(1.073)	(-0.868)			(-0.748)
VIX^{2*}		-0.586		-0.736		-0.094		-0.034
		(-4.174)		(-3.360)		(-0.654)		(-0.136)
α_t^{-*}			0.057	0.037			-0.066	-0.086
			(1.800)	(1.166)			(-1.204)	(-1.660)
R^2	1.476	5.317	0.678	6.307	0.153	0.134	0.870	1.493
	1-month			QMJ		6-months		
Constant	0.004	0.004	0.003	0.003	0.010	0.016	0.007	0.018
	(1.539)	(0.666)	(0.717)	(-0.332)	(3.561)	(2.448)	(1.884)	(3.073)
LJV	1.416			4.376	-38.568			-49.463
	(0.103)			(0.280)	(-4.731)			(3.073)
VIX^{2*}		0.009		-0.022		-0.283		0.043
		(0.059)		(-0.131)		(-1.887)		(0.240)
α_t^{-*}			0.012	0.015			-0.032	-0.080
			(0.359)	(0.409)			(-1.150)	(-4.441)
R^2	0.009	0.003	0.062	0.112	5.996	2.561	0.433	8.382

Table 14: The table reports one- to 12-month return predictability regressions for the aggregate market portfolio.

times and times of stress on the financial markets. Therefore, the next section will be concerning the predictability regression in times of distress.

7.2 Times of distress

The two main focuses for the next part will be on the financial crisis in 2007-08 and the Covid crisis in 2020. One could also include the European sovereign debt crisis and the dot-com crisis, but these are not as distinctive in the left jump variation as the two other crises.

7.2.1 The Financial Crisis of 2008-09

The focus here will be placed on the period from 2008-08-01 to 2009-08-01 for a total of 52 weeks in the sample. Due to the shorter period, the predictive regression will be done on a weekly basis instead of monthly. The regressors are all plotted in 21. Caution should be taken regarding α_t^- and LJV_t since they are very clearly not independent by nature, but the variance inflation factors between the two turn out below two. By comparing the middle and bottom plot, it is explicit that the VIX^2 was on a rising trend before LJV showing an increase in continuous variance and minor jumps leading up to the financial crisis. The tail is also fatter and more skewed for the VIX^2. An interesting result for the

Figure 21: The three regressors. All series are plotted at a weekly frequency, and span the period from August 2008 till start of August 2009. The top panel shows the estimated left jump tail shape parameter, the middle plot is the estimated left jump tail variation measure LJV_t, and the bottom panel shows the CBOE VIX_t^2 volatility index.

LJV_t are the two prominent spikes. They are at the end of the week at 2008-10-26 and 2008-11-23, both weeks after the main drop at the end of September and the start of October. The spikes are also trailing two weeks of rebounds where S&P500 climbed noticeably, indicating that the market feared a downwards jump following the correction. In the fourth column for each portfolio in Table 15 one can find the multivariate regression with the highest R^2 for significant variables. If none of the multivariate regressions had significant variables, then the multivariate regression based on all three regressors is listed. The Wald score is calculated by comparing the portfolio with its nested portfolio, where the regressor is significant, and R^2 was maximized. The regression for quality-minus-junk and high-minus-low are listed at a 1-week horizon since none of the variables was significant at a 13-week prediction horizon. The 5% and 1% significance level for the two-sided t-test are 2.03 and 2.72, respectively.

By comparing Table 15 with Table 13 and 14, it is evident that the significance of the variables and the R^2 has gone up for most of the regressions. Beforehand, the maximum R^2 for the aggregate market portfolio was at 5.1% at

	MKT			13-weeks		BAB		
Constant	0.020	0.025	-0.016	-0.049	-0.005	-0.027	0.013	0.019
	(3.121)	(1.349)	(-1.714)	(-0.847)	(-0.999)	(-4.614)	(3.149)	(3.978)
LJV	-18.525			-13.694	5.940			-4.371
	(-4.955)			(-3.993)	(1.800)			(-3.153)
VIX^{2*}		-0.209		0.448		0.291		
		(-1.155)		(1.062)		(5.839)		
α_t^{-*}			0.411	0.472			-0.263	-0.307
			(5.740)	(1.813)			(-6.833)	(-8.812)
R^2	6.200	1.310	11.200	14.650	1.867	7.460	13.400	14.190
Wald				0.6861				0.325
	UMD			13-weeks		SMB		
Constant	-0.023	-0.024	-0.005	-0.006	0.003	-0.004	-0.001	-0.024
	(-2.017)	(-0.885)	(-0.433)	(-0.108)	(0.914)	(-0.731)	(-0.585)	(-2.402)
LJV	14.885			18.536	-0.819			-2.067
	(2.132)			(3.546)	(-0.474)			(-1.825)
VIX^{2*}		0.129		-0.200		0.070		0.226
		(0.532)		(-0.496)		(1.313)		(2.608)
α_t^{-*}			-0.134	-0.044			0.055	0.132
			(-1.563)	(-0.231)			(-1.167)	(7.230)
R^2	3.012	0.374	0.893	3.461	0.194	2.334	3.259	15.830
Wald				0.079				2.540
	QMJ			1-week		HML		
Constant	0.005	-0.023	0.007	-0.037	0.005	0.044	-0.006	0.059
	(0.897)	(-2.180)	(0.966)	(-3.057)	(0.584)	(2.987)	(-0.593)	(3.523)
LJV	-2.365			-21.946	-1.156			23.774
	(0.897)			(-3.280)	(-0.108)			(2.218)
VIX^{2*}		0.334		0.667		-0.488		-0.849
		(2.667)		(3.935)		(-2.864)		(-3.606)
α_t^{-*}			-0.061				0.190	
			(-0.637)				(1.824)	
R^2	0.264	10.160	0.657	22.740	0.026	8.960	2.656	15.050
Wald				7.657				3.369

Table 15: The table reports either 1-week or quarterly return predictability regressions for varying portfolios, as described in 11, under the financial crisis of 2008-09.

the 6-months horizon, where LJV_t and α_t^- were significant. For the 3-months horizon, corresponding to the 13-weeks here, none of the parameters were significant for the overall period. The picture is drastically different from the financial crisis. The left jump tail variation measure was able to explain 6.2% of the variance of the market returns. In contrast, the jump tail shape was able to explain 11.2% of the variance, which is drastically higher than for the entire period showing that a significant part of the variance on returns on the aggregate market portfolio in crisis periods can be explained through the tail shape. This confirms the importance of distinguishing between the jump intensity for jumps of different sizes and not just the overall jump intensity.

Comparing the five other sub-portfolios, it is evident that their sensitivity to jumps varies greatly. The left jump variation measure is only significant in the univariate regressions for the portfolio sorted on momentum. However, it is significant in the multivariate regressions together with VIX^2 for the portfolios sorted on quality-minus-junk and high-minus-low, where they pull in opposite directions, indicating that an increase in the left jump variation measure predicts a negative effect on the return in the following week, where the opposite holds true for the VIX^2, which also includes minor jumps and continuous variance, reinforcing the idea that the market has a special treatment for large jumps, just as the Levy-Itô decomposition treats large and small jumps differently.

7.2.2 Covid-19 crisis

The focus in this section will be placed on the year 2020. It would be desirable to include data for the first part of 2021 to see if LJV_t remains at an elevated level, but this data is not accessible through OptionMetrics as of this date. In Figure 22 I have inserted vertical lines for important dates. The health dates are when WHO declared a global health emergency and when former President Trump declared a national emergency. The black lines are at the dates for when the Coronavirus Preparedness and Response Supplemental Appropriations Act of 2020, Families First Coronavirus Response Act of 2020, and the CARES Act went into law. The green lines were for when the Fed announced quantitative easing on Treasury securities and government-guaranteed mortgage-backed securities, when they adapted their policy to buying securities "in the amounts needed to support smooth market functioning and effective transmission of monetary policy to broader financial conditions." and lastly the Secondary Market Corporate Credit Facility.

The differences between the VIX^2 and LJV in Figure 22 are notable drastic. Whereas the VIX^2 has stabilized to a level close to pre-Covid times, the left jump tail variation measures are still several orders of magnitude larger than pre-Covid times. The relative standard variation is also far more considerable for LJV compared to the VIX^2 showing that smaller jumps and continuous variance have fallen to levels closer to pre-Covid times, but the market still prices in the risk-neutral probability for significant jumps. The second wave around June/July stands out quite distinctively for LJV_t.

One could have expected drops in the fear proxies following the dates where the fiscal and monetary policies were enacted, but that is not the picture that stands out on either of the two fear proxies. The opposite has proven itself to hold true, indicating that the market had either already priced the policy responses in or the worsening Covid crises dominated the policy responses with

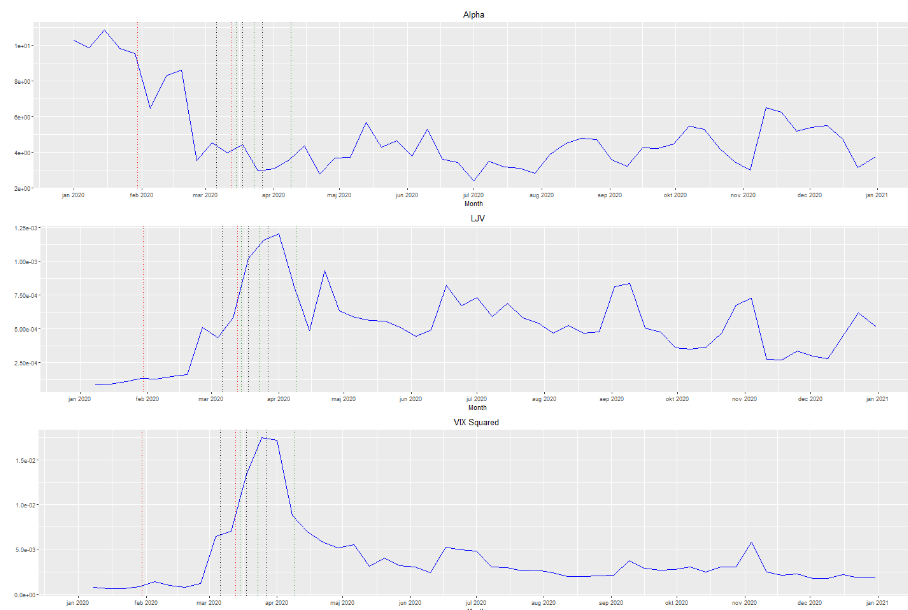


Figure 22: The three regressors. All series are plotted at a weekly frequency, and span the period from January 2020 till end of December 2020. The top panel shows the estimated left jump tail shape parameter, the middle plot is the estimated left jump tail variation measure LJV_t, and the bottom panel shows the CBOE VIX_t^2 volatility index. The red vertical lines indicate major public health events, black lines indicate governmental policy changes, and green lines are changes by the Federal reserve.

regards to the risk-neutral expectation of large negative jumps.

The return predictability for the 13-week horizon for the aggregate market portfolio dramatically resembles the one for the financial crisis. The left jump tail variation measure and the left jump tail shape are significant at predicting the return with approximately identical R^2 for both measures. The signs are also identical across both times of distress. An increase (decrease) in the left jump tail variation predicts a negative (positive) effect on the 13-week return.

Across the other sorted portfolios, the significance is drastically lower in 2020 compared to the period reviewed in the last section. There are just two portfolios where any parameter is significant at a 5% level, the betting-against-beta portfolio, where the multivariate regression for left jump tail variation and left jump tail shape is significant, and the quality-minus-junk, where the left jump tail variation is significant. Comparing the betting-against-beta portfolio to the financial crisis opens for an exciting result. The multivariate regression

	MKT			13-weeks		BAB		
Constant	0.029	0.020	-0.005	0.022	0.001	0.001	-0.002	-0.042
	(8.694)	(2.492)	(-0.923)	(5.152)	(0.117)	(0.111)	(-0.338)	(-2.554)
LJV	-29.733			-51.379	6.593			39.299
	(-4.906)			(-5.470)	(0.514)			(2.677)
VIX^{2*}		-0.114		0.307		0.056		
		(-1.229)		(3.101)		(0.373)		
α_t^{-*}			0.368				0.138	0.525
			(3.578)				(1.139)	(2.583)
R^2	9.682	1.217	9.398	13.440	3.367	3.369	0.833	4.968
Wald				1.563				1.768
	UMD			13-weeks		SMB		
Constant	-0.013	-0.004	0.008	-0.002	0.007	0.005	0.008	0.020
	(-2.267)	(-0.354)	(1.034)	(-0.099)	(2.376)	(1.328)	(1.708)	(1.464)
LJV	18.422			37.897	-2.005			-21.447
	(1.605)			(0.782)	(-0.426)			(-1.402)
VIX^{2*}		0.014		-0.319		0.021		0.108
		(0.076)		(-0.711)		(0.423)		(1.263)
α_t^{-*}			-0.218	-0.048			-0.048	-0.190
			(-1.699)	(-0.181)			(-0.553)	(-1.124)
R^2	0.848	0.004	0.751	1.813	0.097	0.094	0.350	2.919
Wald				0.172				0.463
	QMJ			1-week		HML		
Constant	-0.012	-0.017	0.002	-0.028	0.006	0.008	-0.006	0.029
	(-2.455)	(-2.128)	(0.512)	(-1.723)	(0.520)	(0.631)	(-0.500)	(1.233)
LJV	17.060			11.460	-17.851			-29.147
	(2.040)			(0.602)	(-0.108)			(-0.551)
VIX^{2*}		0.242		0.211		-0.192		-0.069
		(1.738)		(0.910)		(-0.940)		(-0.137)
α_t^{-*}			-0.103	0.143			0.074	-0.265
			(-1.454)	(1.010)			(0.293)	(-0.782)
R^2	7.073	11.530	1.664	11.560	1.242	1.171	0.127	1.997
Wald				0.3088				0.1811

Table 16: The table reports either 1-week or quarterly return predictability regressions for varying portfolios, as described in 11, under Covid-19.

is significant for all estimates in both periods, but the sign has turned to the opposite under Covid, and the R^2 has dropped to nearly a third. The Wald test comparing the multivariate regression to the nested model of just LJV_t does not reject the null hypothesis. Since the estimate is insignificant in the univariate regression, one should not place too much weight on the multivariate regression. There is a barely significant positive estimate for the quality-minus-junk portfolio but a somewhat limited R^2.

For the rest of the sorted portfolios, there was no significant linear predictability.

8 Further research

An issue with comparing the VIX and the LJV is that the VIX also includes the significant jumps. As such, it would be of interest to have an estimate of the variance risk premium and then remove the LJV_t in order to only proxy continuous risk premium. As proposed in Bollerslev et al. (2015), the variance risk premium can be calculated as the difference between the VIX_t^2 and the expected integrated variation over the time horizon. As proposed in Bollerslev & Todorov (2011b) the expected integrated variation can be calculated as a 22-order multivariate auto-regressive time series,

$$X_t = A_0 + A_1 X_{t-1} + A_5 \sum_{i=1}^{5} X_{t-i}/5 + A_{22} \sum_{i=1}^{22} X_{t-i}/22 + \varepsilon_i \tag{72}$$

for the four dimensional vector,

$$X_t \equiv (CV_t, RJV_t, LJV_t, (p_{t+\pi_t} - p_t)^2)'.$$

The significant issue with this methodology is that it is based on high-frequency data in order to distinguish between continuous variation and jump variation. The variation measures are as follows,

$$\begin{aligned} CV_t &\equiv \sum_{i=1}^{n-1} (\Delta_i^{n,t} f)^2 1_{\{|\Delta_i^{n,t} f| \leq \alpha \Delta_{n,t}^{\omega}\}} \\ RJV_t &\equiv \sum_{i=1}^{n-1} (\Delta_i^{n,t} f)^2 1_{\{\Delta_i^{n,t} f > \alpha \Delta_{n,t}^{\omega}\}} \\ LJV_t &\equiv \sum_{i=1}^{n-1} (\Delta_i^{n,t} f)^2 1_{\{\Delta_i^{n,t} f < -\alpha \Delta_{n,t}^{\omega}\}}, \end{aligned} \tag{73}$$

where $\Delta_{n,t} \equiv \frac{1}{n}$ and $\Delta_i^{n,t} f \equiv f_{t+i\Delta_{n,t}} - f_{t+(i-1)\Delta_{n,t}}$ for $i = 1, ..., n-1$.

The choice of truncation level is vital in practice since the variance of futures are dependent on the time of day, as the market is more active during typical trading hours. This diurnal pattern can be estimated nonparametrically by a time-of-day factor $TOD_i, i = 1, n, ..., n,$

$$
\begin{aligned}
&TOD_i \\
&= NOI_i \frac{\sum_{t=1}^{N} (f_{t-1+i\Delta_{n,t}} - f_{t-1+(i-1)\Delta_{n,t}})^2 1_{\{|f_{t-1+i\Delta_{n,t}} - f_{t-1+(i-1)\Delta_{n,t}}| \leq \bar{\alpha}\Delta_n^{\omega}\}}}{\sum_{t=1}^{N} \sum_{i=1}^{n-1} (f_{t-1+i\Delta_{n,t}} - f_{t-1+(i-1)\Delta_{n,t}})^2}, \\
&NOI_i = \frac{\sum_{t=1}^{N} \sum_{i=1}^{n-1} 1_{\{|f_{t-1+i\Delta_{n,t}} - f_{t-1+(i-1)\Delta_{n,t}}| \leq \bar{\alpha}\Delta_n^{\omega}\}}}{\sum_{t=1}^{N} 1_{\{|f_{t-1+i\Delta_{n,t}} - f_{t-1+(i-1)\Delta_{n,t}}| \leq \bar{\alpha}\Delta_n^{\omega}\}}}, \\
&\bar{\alpha} = 3\sqrt{\frac{\pi}{2}} \\
&\sqrt{\frac{1}{N} \sum_{t=1}^{N} \sum_{i=2}^{n-1} |f_{t-1+i\Delta_{n,t}} - f_{t-1+(i-1)\Delta_{n,t}}||f_{t-1+(i-1)\Delta_{n,t}} - f_{t-1+(i-2)\Delta_{n,t}}|}, \\
&\Delta_n \equiv \frac{1}{n}.
\end{aligned}
\tag{74}
$$

The significant issue I faced here was the lack of access to high-frequency data for the entire period. Through Bloomberg Terminal, it was possible to access 30-minute high-frequency data for the past year on a running basis, but higher frequency data was just available for three months. This led to a volatile time-of-day factor, as shown in Figure 23, which led the methodology above characterising almost all price changes as either a right or left jump. This is not a desirable trait, and as such, I have not included the VAR estimates, which was calculated using the package marima in R, nor the predictability regressions based on this. It would be beneficial to repeat this review using higher frequency data for the entire period.

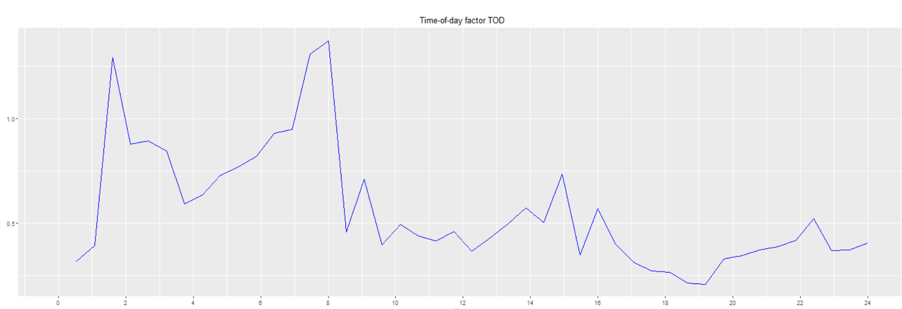


Figure 23: The estimates are based on 30-minute high-frequency S&P 500 futures data from 13-02-2020 till 13-02-2021.

Another exciting research area would be to compare the left jump variation measure across different countries and make a multivariate regression including other observable economic variables in order to gauge the effect of real economic shocks or rate changes. This could take a basis in this paper and in the recent paper Andersen et al. (2021) where the tail risk and return predictability for the Japanese equity market was reviewed.

9 Conclusion

In this thesis, the jump intensity process of risk-neutral returns has been analyzed under the assumption that the tail shape is either time-invariant or time-variant, where it was shown that both the overall jump intensity and tail shape was necessary for the risk-neutral jump variation measure. Nevertheless, the tail shape measure carried the majority of the information included in the jump variation measure, and as such, there was a focus on modelling the time-variant tail shape.

One of the few assumptions applied in this thesis is that the dynamics in the risk-neutral tails can be characterized fully by a Lévy measure. Therefore, to begin with, the fundamental properties of Lévy processes are considered. Based on a foundation of the Brownian motion and the Poisson Process, the class leads to a right-continuous stochastic process with jumps of varying activity. A strength of Lévy processes is that they can be completely characterized via their characteristic functions or Lévy triplet, which follows from the celebrated Lévy-Khintchine representation, and can be decomposed into four independent parts that yields information about the path properties of one's process.

Following the theoretical groundwork, a model was built around the idea that the variance risk premium is decomposed into two fundamentally different sources of market variance risk, and it is argued why the compensation for left jump tail risk is a valid proxy for market fear. This led to modelling the tail shape parameter and a level shift for the jump intensity, where the empirical estimation is built by using the option surface to get risk-neutral data as the physical jump estimation would be plagued by a dearth of extreme events.

A rigorous cleaning procedure is employed, and it is reviewed how deep out-of-the-money options should be in order for them to mimic large jumps and not smaller jumps and continuous changes. Further, it is shown that there is a dire need for time-variant tail shapes and that the assumption of constant tail shapes in earlier literature is problematic. A time series modelling of the tail shape follows. The time series is first cleaned of trend and seasonality, whereafter it is modelled by an ARIMA/GARCH in order to secure both stationarity in the mean and in the variance, and the estimate is controlled by a one-period rolling forecast that can replicate the majority of the actual tail shape.

Based on the estimated models, it was found that the left jump tail variation was a significant predictor with the aggregate market, which is consistent with the idea that jump tail variation is a proxy for fear in the market. By comparing with the predictive regression of the VIX, which includes both continuous and jump variation, it is clear that the explanatory power of return predictability regressions differs significantly between the VIX and LJV, which is persistent with the theory that jump risk factors represents state variables that drive the market risk premium. For the aggregate market portfolio, the predictability regressions were more substantial for the LJV, confirming a truer proxy for market fears.

Finally, the same process was repeated for other commonly studied portfolio sorts under the financial crisis and the Covid-19 crisis. The message conveyed for the aggregate market portfolio is consistent across the overall period and the times of distress, with the left jump tail variation measure being significant while the VIX is insignificant. By comparing the two periods of distress, it is clear that the explanatory power of return predictability regression is more substantial across the financial crisis compared to the Covid crisis, but the message is generally the same. The VIX, thereby continuous variance, shows stronger predictability on the short horizon for quality-minus-junk and high-minus-low, but weaker predictability across the other portfolios with the 13-week horizons.

The procedures mentioned above were implemented in R using both standard and non-standard packages. A specification of the packages used and the code can be found in the appendix.

10 Bibliography

Andersen, T. G., Todorov, V. & Ubukata, M. (2021), 'Tail risk and return predictability for the japanese equity market', *Journal of Econometrics* **222**(1), 344–363.

Asness, C. & Frazzini, A. (2013), 'The devil in HML's details', *The Journal of Portfolio Management* **39**(4), 49–68.

Asness, C. S., Frazzini, A. & Pedersen, L. H. (2019), 'Quality minus junk', *Review of Accounting Studies* **24**(1), 34–112.

Bachelier, L. (1900), Théorie de la spéculation, *in* 'Annales scientifiques de l'École normale supérieure', Vol. 17, pp. 21–86.

Bachelier, L. (1901), Théorie mathématique du jeu, *in* 'Annales Scientifiques de l'Ecole Normale Supérieure', Vol. 18, pp. 143–209.

Bennett, B., Stulz, R. M. & Wang, Z. (2020), Does joining the S&P 500 index hurt firms?, Technical report, National Bureau of Economic Research.

Bianchi, M., Boyle, M. & Hollingsworth, D. (1999), 'A comparison of methods for trend estimation', *Applied Economics Letters* **6**(2), 103–109.

Black, F. & Scholes, M. (1973), 'The pricing of options and corporate liabilities', *The Journal of Political Economy* **81**(3), 637–654.

Bollerslev, T. (1986), 'Generalized autoregressive conditional heteroskedasticity', *Journal of econometrics* **31**(3), 307–327.

Bollerslev, T. & Todorov, V. (2011*a*), 'Estimation of jump tails', *Econometrica* **79**(6), 1727–1783.

Bollerslev, T. & Todorov, V. (2011*b*), 'Tails, fears, and risk premia', *The Journal of Finance* **66**(6), 2165–2211.

Bollerslev, T. & Todorov, V. (2014), 'Time-varying jump tails', *Journal of Econometrics* **183**(2), 168–180.

Bollerslev, T., Todorov, V. & Xu, L. (2015), 'Tail risk premia and return predictability', *Journal of Financial Economics* **118**(1), 113–134.

Box, G. E. & Pierce, D. A. (1970), 'Distribution of residual autocorrelations in autoregressive-integrated moving average time series models', *Journal of the American statistical Association* **65**(332), 1509–1526.

Carhart, M. M. (1997), 'On persistence in mutual fund performance', *The Journal of finance* **52**(1), 57–82.

Carr, P. & Wu, L. (2003), 'What type of process underlies options? A simple robust test', *The Journal of Finance* **58**(6), 2581–2610.

Carr, P. & Wu, L. (2009), 'Variance risk premiums', *The Review of Financial Studies* **22**(3), 1311–1341.

Chen, H., Joslin, S. & Ni, S. X. (2019), 'Demand for Crash Insurance, Intermediary Constraints, and Risk Premia in Financial Markets', *The Review of financial studies* **32**(1), 228–265.

Edward, F. R. (1999), 'Hedge funds and the collapse of long-term capital management', *Journal of Economic Perspectives* **13**(2), 189–210.

Engle, R. F. (1982), 'Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation', *Econometrica: Journal of the econometric society* pp. 987–1007.

Fama, E. F. & French, K. R. (1993), 'Common risk factors in the returns on stocks and bonds', *Journal of Financial Economics* **33**, 3–56.

Franses, P. H. & Paap, R. (2004), *Periodic time series models*, OUP Oxford.

Franses, P. H. et al. (1996), 'Periodicity and stochastic trends in economic time series', *OUP Catalogue* .

Frazzini, A. & Pedersen, L. H. (2014), 'Betting against beta', *Journal of Financial Economics* **111**(1), 1–25.

Goetzmann, W. N., Kim, D. & Shiller, R. J. (2016), Crash beliefs from investor surveys, Technical report, National Bureau of Economic Research.

Ilmanen, A. (2012), 'Do financial markets reward buying or selling insurance and lottery tickets?', *Financial Analysts Journal* **68**(5), 26–36.

Khintchine, A. (1934), 'Korrelationstheorie der stationären stochastischen prozesse', *Mathematische Annalen* **109**(1), 604–615.

Koopmans, L. H. L. H. (1995), *The spectral analysis of time series*, Probability and mathematical statistics ; v. 22, [2nd ed.]. edn, Academic Press, San Diego.

Kou, S. G. (2002), 'A jump-diffusion model for option pricing', *Management science* **48**(8), 1086–1101.

Kyprianou, A. E. (2014), *Fluctuations of Lévy processes with applications: Introductory Lectures*, Springer Science & Business Media.

Ljung, G. M. & Box, G. E. (1978), 'On a measure of lack of fit in time series models', *Biometrika* **65**(2), 297–303.

Lundberg, F. (1903), Approximerad framställning av sannolikhetsfunktionen. aterförsäkring av kollektivrisker. Akad, PhD thesis, Afhandling. Almqvist och Wiksell, Uppsala.

Merton, R. C. (1973), 'Theory of rational option pricing', *The Bell Journal of economics and management science* pp. 141–183.

Mohammadi, H. & Su, L. (2010), 'International evidence on crude oil price dynamics: Applications of ARIMA-GARCH models', *Energy Economics* **32**(5), 1001–1008.

Papapantoleon, A. (2008), 'An introduction to Lévy processes with applications in finance', *arXiv preprint arXiv:0804.0482* .

Regnault, J. (1863), *Calcul des chances et philosophie de la bourse*, librairie Castel.

Sewell, M. (2011), 'Characterization of financial time series', *Rn* **11**(01), 01.

Thiele, T. N. (1880), *Om Anvendelse af mindste Kvadraters Methode i nogle Tilfaelde, hvor en Komplikation af visse Slags uensartede tilfaeldige Fejlkilder giver Fejlene en" systematisk" Karakter, af TN Thiele*, B. Lunos Kgl. Hof.-Bogtrykkeri.

Zhao, S. & Wei, G.-W. (2003), 'Jump process for the trend estimation of time series', *Computational Statistics & Data Analysis* **42**(1-2), 219–241.

11 Appendix

All of the calculations have been done in R through the use of several standard and non-standard packages.

The following packages were employed in the code:

- data.table: Used for converting data.frames to data.tables which are easier to manipulate and more efficient,
- dplyr: Allows for efficient manipluation,
- ggplot2: Convenient for plotting,
- lubridate: Efficient data manipulation,
- readr: Efficient reading of data tables,
- reshape2: Allows for melting of data which is used for plots,
- rlang: Part of the tidyverse,
- rlist: Used for working with lists,
- stats: Includes a wide range of statistical tools,
- stringr: Used in string manipulation,
- tideselect: Select from strings,
- zoo: Package for working with time series,
- TSA: Package for working with time series,
- astsa: Package for working with time series,
- forecast: Package for forecasting time series,
- tibble: Addition to data.tables,
- gridExtra: Additional control over plots,
- sandwich: Allows for the calculation of Newey-West t-statistics,
- lmtest: Used for Wald test.

```r
#=============================================================================
#
# Script for running thesis calculations
#
#=============================================================================

#rm(list=ls())            # Clear the workspace
packages <- c("data.table","dplyr","ggplot2","fredr",
              "jrvFinance", "lubridate","RcppRoll","readr",
              "reshape2","rlang","rlist","stats","stringr","tidyselect","zoo","progress","minpack.lm",
              "TSA","astsa","forecast","tibble","gridExtra","sandwich","lmtest")

#load library for all packages
lapply(X = packages, FUN = library, character.only = TRUE)

# Set working directory the folder in which the script is
setwd(dirname(rstudioapi::getActiveDocumentContext()$path))

# Go up one folder in working directory, so we go from "Term Structure of Funding / Code" to the main folder "Term Structure of Funding"
setwd("..")

# Which henceforth will be the root directory
root_dir   <- getwd()

# Changing working directory to the location of the script
work_dir   <- file.path(getwd(),"Code")

# Directory of data and output files
data_dir          <- file.path(root_dir,"Data")
output_dir        <- file.path(root_dir,"Results")

#Read CSV files
optionprices <- fread(file.path(data_dir,"OptionMetrics.csv"), sep = ",")
indexprice   <- fread(file.path(data_dir,"SPPrice.csv"), sep = ",")
future       <- fread(file.path(data_dir,"Future.csv"), sep = ",")
FamaFrench   <- fread(file.path(data_dir,"FamaFrench.csv"), sep = ",")

#Remove observations without implied volatility and mutate dates and strike
optionprices <- optionprices %>% filter(!is.na(impl_volatility)) %>%
  mutate(date = ymd(date),exdate = ymd(exdate)) %>%
  mutate(strike_price = strike_price/1000)
indexprice <- indexprice %>% mutate(date = ymd(date))
future <- future %>% mutate(date = ymd(date), expiration = ymd(expiration))
FamaFrench <- FamaFrench %>% mutate(date = ymd(date))
#Combine the data sets and drop irrelevant columns
combined <- merge(optionprices,indexprice,by="date")
drop_col <- c("secid.x","symbol","forward_price","index_flag","class","issuer","secid.y","cusip","volume.y","shrout","cfret")
combined <- combined %>% select(-one_of(drop_col))
combined <- combined %>% mutate(comdate = paste(combined$date,combined$exdate))
colnames(future)[3] <- "exdate"

future <- future %>% mutate(comdate = paste(future$date,future$exdate))
future <- future[comdate %in% combined$comdate]
future <- future[!duplicated(future$comdate)]
combined <- merge(combined,future[,c(4,5,10)],by = "comdate")


### Cleaning procdure ###
#Remove any zero-bid options
combined <- combined %>% filter(best_bid != 0)

#Only keep ATM or OTM points
combined <- combined %>% filter(cp_flag=="P" & strike_price<=ForwardPrice | cp_flag=="C" & strike_price>=ForwardPrice)

#Add a mid-quote
combined <- combined %>% mutate(midquote = (best_bid+best_offer)/2)
combined <- combined[order(date,exdate,cp_flag,strike_price)]
```

```r
#Check for arbitrage trades
combined_p <- combined %>% filter(cp_flag=="P") %>% arrange(comdate,-strike_price)
combined_p_diff <- NULL
k<-1
cont <- 1
while(k <= length(unique(combined_p$comdate))){
  print(k)
  i <- unique(combined_p$comdate)[k]
  if(cont == 1){
    combined_p_diff <- rbind(combined_p_diff,combined_p %>% filter(comdate==i) %>%
    mutate(diff = c(-1,diff((combined_p %>% filter(comdate==i))$midquote))))
    combined_p_diff[which(combined_p_diff$diff<0 &
    combined_p_diff$diff>-1*10^-5)]$diff=0
    combined_p_diff[which(combined_p_diff$diff>0 &
    combined_p_diff$diff<1*10^-5)]$diff=0
  }
  cont <- 1
  if(any(combined_p_diff$diff>0)){
    for(p in 1:length(which(combined_p_diff$diff>0))){
      index <- which(combined_p_diff$diff>0)
      j <- index[1]
      if(combined_p_diff[j]$volume.x>combined_p_diff[j-1]$volume.x){
        combined_p_diff <- combined_p_diff %>% slice(-(j-1))
        combined_p_diff[j]$diff = -1
      } else if(combined_p_diff[j]$volume.x<=combined_p_diff[j-1]$volume.x){
        combined_p_diff <- combined_p_diff %>% slice(-j)
      }
      index = index - 1
    }
    combined_p_diff[combined_p_diff$comdate==i] <-
    combined_p_diff[combined_p_diff$comdate==i]  %>% mutate(diff =
    c(-1,diff((combined_p_diff[combined_p_diff$comdate==i] %>%
    filter(comdate==i))$midquote)))
    combined_p_diff[combined_p_diff$comdate==i][which(combined_p_diff$diff<0 &
    combined_p_diff$diff>-1*10^-5)]$diff=0
    combined_p_diff[combined_p_diff$comdate==i][which(combined_p_diff$diff>0 &
    combined_p_diff$diff<1*10^-5)]$diff=0
    cont <- 0
  }
  if(any(combined_p_diff$diff>=0) & cont == 1){
    for(p in 1:length(which(combined_p_diff$diff>=0))){
      index <- which(combined_p_diff$diff>=0)
      j <- index[1]
      if(combined_p_diff[j]$volume.x>combined_p_diff[j-1]$volume.x){
        combined_p_diff <- combined_p_diff %>% slice(-(j-1))
        combined_p_diff[j-1]$diff = -1
      } else if(combined_p_diff[j]$volume.x<=combined_p_diff[j-1]$volume.x){
        combined_p_diff <- combined_p_diff %>% slice(-j)
      }
      index = index - 1
    }
  }
  if(cont == 1) k = k + 1
}
#fwrite(x=combined_p_diff,file.path(data_dir,"PutsCleanedFinal.csv"), sep = ",")
combined_final <- fread(file.path(data_dir,"PutsCleanedFinal.csv"))

##Fix issue with multiple bonds with same comdate and same strike
list <- NULL
k <- 1
for(i in unique(Option_P$comdate)[1:length(unique(Option_P$comdate))]){
  Option_sell <- Option_P %>% filter(comdate==i)
  if(any(duplicated((Option_sell$strike_price)))){
    list <- rbind(list,Option_sell)
  }
  k = k + 1
  print(k)
}
list2 <- NULL
for(i in 1:length(unique(list$comdate))){
```

```
    j = unique(list$comdate)[i]
    Option_sell <- Option_P %>% filter(comdate==j)
    size <- dim(Option_sell)[1]
    k = 1
    while(k<size){
      if(Option_sell[k,5]==Option_sell[k+1,5]){
        Option_sell[k,23] <- mean(c(as.numeric(Option_sell[k,23]),as.numeric(Option_sell[k+1,23])))
        Option_sell <- Option_sell %>% slice(-(k+1))
        size = size - 1
      }
      k = k + 1
    }
    list2 <- rbind(list2, Option_sell)
  }

  #Control
  for(i in unique(Option_P$comdate)){
    Option_sel <- Option_P %>% filter(comdate==i)
    if(any(duplicated((Option_sel$strike_price)))) print(i)
  }


  ##Find the ATM implied vol for each comdate
  for(i in unique(Option_P$comdate)[1:length(unique(Option_P$comdate))]){
    current[1,1] <- i
    current[1,2] <- as.numeric((Option_P %>% filter(comdate==i))[1]$impl_volatility)
    atmvol <- rbind(atmvol,current)
    k=k+1
    print(k)
  }
  #Merge atmvol with Option_P and save it

  ##Proceeds to next step, including filter. CleanedAndReady.csv is already cleaned for filter(cp_flag=="P" & log_moneyness < -5*atmvol)
  #Cleaned up option table
  OptionTable <- fread(file.path(data_dir,"CleanedAndReady.csv"), sep = ",")
  OptionTable <- OptionTable[,1:27]
  storage.mode(OptionTable$date) <- "integer"
  storage.mode(OptionTable$exdate) <- "integer"

  ##The two lines below are needed for the first run
  #OptionTable <- OptionTable %>% mutate(log_moneyness = log(strike_price/ForwardPrice),timetoexp = as.integer(exdate-date),atmvol=atmvol*sqrt(timetoexp/365))
  #OptionTable <- OptionTable %>% filter(cp_flag=="P" & log_moneyness < -5*atmvol)

  #Write table
  #fwrite(x=OptionTable,file.path(data_dir,"OptionTable.csv"), sep = ",")
  OptionTable <- setDT(OptionTable)[,.N, by = c("date","cp_flag")] %>% left_join(x = OptionTable, y = ., by = c("date","cp_flag"))
  colnames(OptionTable)[28] <- "N_date"
  OptionTable <- setDT(OptionTable)[,.N, by = c("comdate","cp_flag")] %>% left_join(x = OptionTable, y = ., by = c("comdate","cp_flag"))
  #Construction

  #Calculation of alpha and phi
  #Filter for at least two bonds on each day
  Option_P <- OptionTable %>% filter(cp_flag=="P" & N>2 & N_date >2)
  Option_P <- Option_P[order(date,exdate,-log_moneyness)]
  storage.mode(Option_P$date) <- "integer"
  storage.mode(Option_P$exdate) <- "integer"
  alpha <- NULL

  #Select smoothing
  freq="weekly"
  if(freq=="weekly"){     interpolation_scheme <- seq(from = min(Option_P$date), to = max(Option_P$date), length.out = as.numeric(max(Option_P$date)-min(Option_P$date))/7)
    }else if(freq=="daily"){interpolation_scheme <- seq(from = min(Option_P$date), to = max(Option_P$date), length.out = as.numeric(max(Option_P$date)-min(Option_P$date)))
    }else if(freq=="quarterly"){interpolation_scheme <- seq(from = min(Option_P$date), to = max(Option_P$date), length.out =
```

```
  as.numeric(max(Option_P$date)-min(Option_P$date))/91.3125)
  }else if(freq=="monthly"){interpolation_scheme <- seq(from = min(Option_P$date), to = max(Option_P$date), length.out = as.numeric(max(Option_P$date)-min(Option_P$date))/30.4375)
  }else if(freq=="yearly")interpolation_scheme <- seq(from = min(Option_P$date), to = max(Option_P$date), length.out = as.numeric(max(Option_P$date)-min(Option_P$date))/364)
Option_P <- Option_P[order(date,exdate,-log_moneyness)]
Option_P <- Option_P %>% select(-c(28:29))
Option_P <- setDT(Option_P)[,.N, by = c("date","cp_flag")] %>% left_join(x = Option_P, y = ., by = c("date","cp_flag"))

###Smoothing
alpha<-NULL
for(i in 1:(length(interpolation_scheme)-1)){
  Option_sel <- Option_P %>% filter(date>=interpolation_scheme[i]&date<interpolation_scheme[i+1])
  if(dim(Option_sel)[1]>0){
    alpha[i] <- optim(par = 10, fn=Find_alphaSmooth,method="Brent",lower=0.1,upper=100,Option=Option_sel)$par
  }
  print(i)
}
phi <- NULL
for(i in 1:(length(interpolation_scheme)-1)){
  Option_sel <- Option_P %>% filter(date>=interpolation_scheme[i]&date<interpolation_scheme[i+1])
  if(dim(Option_sel)[1]>0){
    phi[i] <- optim(par = 0.5, fn=Find_PsiSmooth,method="Brent",lower=0.000001,upper=200,Option=Option_sel,alpha=alpha_weekly$alpha[i],riskfree=testdate$rf[i])$par
  }
  print(i)
}
#fwrite(x=alpha1,file.path(data_dir,"Alpha.csv"), sep = ",")

#Functions applied above
#Non-smoothed
Find_alpha2 <- function(Option,par){
  alpha_minus_sum <- 0
  for(j in 1:length(unique(Option$exdate))){
    Option_sell <- Option %>% filter(exdate==unique(Option$exdate)[j])
    size <- dim(Option_sell)[1]
    alpha_minus_sum <- sum(abs(log(Option_sell[2:size,23]/Option_sell[1:(size-1),23])/(Option_sell[2:size,25]-Option_sell[1:(size-1),25])-(1+par)))
  }
  alpha_minus_sum <- alpha_minus_sum/Option[1,28]
  return(as.numeric(alpha_minus_sum))
}

#Smoothed
Find_alphaSmooth <- function(Option,par){
  alpha_minus_sum <- 0
  for(k in 1:length(unique(Option$date))){
    for(j in 1:length(unique(Option$exdate))){
      Option_sell <- Option %>% filter(exdate==unique(Option$exdate)[j])
      size <- dim(Option_sell)[1]
      alpha_minus_sum <- sum(abs(log(Option_sell[2:size,23]/Option_sell[1:(size-1),23])/(Option_sell[2:size,26]-Option_sell[1:(size-1),26])-(1+par)))
    }
  }
  alpha_minus_sum <- alpha_minus_sum/(dim(Option)[1])
  return(as.numeric(alpha_minus_sum))
}
#Smoothed
Find_PsiSmooth <- function(Option,par,alpha,riskfree){
  psi_minus_sum <- 0
  for(k in 1:length(unique(Option$date))){
    Option_sel <- Option %>% filter(date==unique(Option$date)[k])
    for(j in 1:length(unique(Option_sel$exdate))){
      Option_sell <- Option_sel %>% filter(exdate==unique(Option_sel$exdate)[j])
```

```
      size <- dim(Option_sell)[1]
      psi_minus_sum <-
      sum(abs(log(exp(riskfree*Option_sell[1:size,27])*Option_sell[1:size,23]/(Option_sell[1:size,27]*Option_sell[1:size,22]))-(1+alpha)*Option_sell[1:size,26]+log(alpha+1)+log(alpha)-log(par)))
    }
  }
  psi_minus_sum <- psi_minus_sum/(dim(Option)[1])
  return(as.numeric(psi_minus_sum))
}
#Non-smoothed
Find_Psi <- function(Option,par,alpha){
  psi_minus_sum <- 0
  for(j in 1:length(unique(Option$exdate))){
    Option_sell <- Option %>% filter(exdate==unique(Option$exdate)[j])
    size <- dim(Option_sell)[1]
    psi_minus_sum <-
    sum(abs(log(Option_sell[1:size,23]/(Option_sell[1:size,26]*Option_sell[1:size,22]))-(1+alpha)*Option_sell[1:size,25]+log(alpha+1)+log(alpha)-log(par)))
  }
  psi_minus_sum <- psi_minus_sum/Option[1,28]
  return(as.numeric(psi_minus_sum))
}

##Time series analysis
#First, test for polynomial trend
t = 1:372
fit = lm(prodn ~ t + I(t^2))
r.fit = fit$resid
I=abs(fft(r.fit))^2/372
P=(4/372)*I[1:186]
f=0:185/372
plot(f, P, type="l", xlab="Frequency", ylab = "Scaled Periodogram")


ts <- alpha_weekly_3$alpha[which(!is.na(alpha_weekly_3$alpha))]
t = 1:length(ts)
r.fit = diff(1/alpha_weekly$alpha[which(!is.na(alpha_weekly$alpha))],1)
fit = lm(ts~t+I(t^2)+I(t^3)+I(t^4)+I(t^5))
r.fit = fit$resid
I=abs(fft(r.fit))^2/1246
P=(4/1246)*I[1:(1246/2)]
f=0:(1246/2-1)/1246
plot(f, P, type="l", xlab="Frequency", ylab = "Scaled Periodogram")
plot(ts,type="l")
points(fit14$fitted,type="l")
f1 = 1/ts2[which(ts2$P == max((ts2 %>% filter(f<0.03))$P))]$f
f2 = 1/ts2[which(ts2$P == max((ts2 %>% filter(f>=0.03 & f<0.1))$P))]$f
f3 = 1/ts2[which(ts2$P == max((ts2 %>% filter(f>=0.1 & f<0.2))$P))]$f
f4 = 1/ts2[which(ts2$P == max((ts2 %>% filter(f>=0.2 & f<0.3))$P))]$f

c1 = cos(2*pi*t/f4); s1 = sin(2*pi*t/f4)
c2 = cos(2*pi*t/f3); s2 = sin(2*pi*t/f3)
c3 = cos(2*pi*t/f2); s3 = sin(2*pi*t/f2)
c4 = cos(2*pi*t/f1); s4 = sin(2*pi*t/f1)
fit2 = lm(ts~t+I(t^2)+I(t^3)+I(t^4)+I(t^5)+c1+s1+c2+s2+c3+s3+c4+s4)
plot(t,ts,type="l"); points(t,fit2$fitted, type="l", col="red")




#Linear fit
ts <- alpha_weekly$alpha[which(!is.na(alpha_weekly$alpha))]
t = 1:length(r.fit)
fit = lm(1/ts~t)
plot(t,1/ts,type="l"); points(t,fit$fitted, type="l", col="red")
summary(fit)
fit$residuals[2:1246]


#First-order differenced, non smoothed
r.fit = diff(1/alpha_weekly$alpha[which(!is.na(alpha_weekly$alpha))],1)
I=abs(fft(r.fit))^2/1246
```

```
P=(4/1246)*I[1:(1246/2)]
f=0:(1246/2-1)/1246
plot(f, P, type="l", xlab="Frequency", ylab = "Scaled Periodogram")
ggplot(data=Periodogram, aes(x=f,y=P))+geom_line(color="blue")
max1 <- max((Periodogram %>% filter(f<0.2))$P)
max2 <- max((Periodogram %>% filter(f<0.1))$P)
max3 <- max((Periodogram %>% filter(f<0.35 & f>0.3))$P)
max4 <- max((Periodogram %>% filter(f<0.15))$P)
max5 <- max((Periodogram %>% filter(f>0.3))$P)
ggplot(data=Periodogram, aes(x=f,y=P))+geom_line(color="blue")+geom_point(data=Periodogram[Periodogram$P==max(Periodogram$P),],pch=21, fill=NA, size=4, colour="red", stroke=1)+
  geom_point(data=Periodogram[Periodogram$P==max1,],pch=21, fill=NA, size=4, colour="red", stroke=1) +
  geom_point(data=Periodogram[Periodogram$P==max2,],pch=21, fill=NA, size=4, colour="red", stroke=1) +
  geom_point(data=Periodogram[Periodogram$P==max3,],pch=21, fill=NA, size=4, colour="red", stroke=1) +
  geom_point(data=Periodogram[Periodogram$P==max4,],pch=21, fill=NA, size=4, colour="red", stroke=1) +
  geom_point(data=Periodogram[Periodogram$P==max5,],pch=21, fill=NA, size=4, colour="red", stroke=1) +
  labs(x="Frequency",y="Scaled Periodogram")
f0 <- 1/Periodogram[Periodogram$P==max(Periodogram$P),]$f
f1 <- 1/Periodogram[Periodogram$P==max1,]$f
f2 <- 1/Periodogram[Periodogram$P==max2,]$f
f3 <- 1/Periodogram[Periodogram$P==max3,]$f
f4 <- 1/Periodogram[Periodogram$P==max4,]$f
f5 <- 1/Periodogram[Periodogram$P==max5,]$f

c1 = cos(2*pi*t/f4); s1 = sin(2*pi*t/f4)
c2 = cos(2*pi*t/f3); s2 = sin(2*pi*t/f3)
c3 = cos(2*pi*t/f2); s3 = sin(2*pi*t/f2)
c4 = cos(2*pi*t/f1); s4 = sin(2*pi*t/f1)
c5 = cos(2*pi*t/f0); s5 = sin(2*pi*t/f0)
fit2 = lm(r.fit~c1[2:1246]+s1[2:1246]+c2[2:1246]+s2[2:1246]+c3[2:1246]+s3[2:1246]+c4[2:1246]+s4[2:1246]+c5[2:1246]+s5[2:1246])
fit2 = lm(r.fit~c1[1:1245]+s1[1:1245]+c2[1:1245]+s2[1:1245]+c3[1:1245]+s3[1:1245]+c4[1:1245]+s4[1:1245]+c5[1:1245]+s5[1:1245])
fit3 = lm(r.fit~s5[1:1245])


#First-order differenced, non-parametric smoothing and parametric smoothing
m=5
l=2*m+1
m2=5
l2=2*m2+1
vals3 <- mvspec(r.fit,spans = c(15*2+1,8*2+1),log="no")
vals2 <- mvspec(r.fit,spans = c(15*2+1),log="no")
vals1 <- mvspec(r.fit,spans = c(4*2+1),log="no")

specvals<- spec.ar(r.fit,log="no")
specvals<- setDT(as.data.frame(cbind(specvals$freq,specvals$spec)))
plotsmooth <- setDT(as.data.frame(cbind(vals1$freq,vals1$spec,vals2$spec,vals3$spec)))

max1 <- max((plotsmooth %>% filter(V1<0.1))$V4)
max2 <- max((plotsmooth %>% filter(V1<0.2))$V4)
max3 <- max((plotsmooth %>% filter(V1<0.25))$V4)
max4 <- max((plotsmooth %>% filter(V1<0.28))$V4)
max5 <- max((plotsmooth %>% filter(V1<0.35))$V4)
max6 <- max((plotsmooth %>% filter(V1>0.35&V1<0.4))$V4)

per_1 <- ggplot(data=plotsmooth, aes(x=V1,y=V4))+geom_line(color="blue")+geom_point(data=plotsmooth[plotsmooth$V4==max(plotsmooth$V4),],pch=21, fill=NA, size=4, colour="red", stroke=1)+
  geom_point(data=plotsmooth[plotsmooth$V4==max1,],pch=21, fill=NA, size=4, colour="red", stroke=1) +
  geom_point(data=plotsmooth[plotsmooth$V4==max2,],pch=21, fill=NA, size=4, colour="red", stroke=1) +
  geom_point(data=plotsmooth[plotsmooth$V4==max3,],pch=21, fill=NA, size=4, colour="red", stroke=1) +
```

```
  geom_point(data=plotsmooth[plotsmooth$V4==max4,],pch=21, fill=NA, size=4, colour="red", stroke=1) +
  geom_point(data=plotsmooth[plotsmooth$V4==max5,],pch=21, fill=NA, size=4, colour="red", stroke=1) +
  geom_point(data=plotsmooth[plotsmooth$V4==max6,],pch=21, fill=NA, size=4, colour="red", stroke=1) +
  labs(x="Frequency / Bandwidth = 0.0292",y="Scaled Periodogram",title="Convoluted Daniell kernel smoothing with m=(15,8)")
per_2 <- ggplot(data=plotsmooth, aes(x=V1,y=V3))+geom_line(color="blue")+
  labs(x="Frequency / Bandwidth = 0.0244",y="Scaled Periodogram",title="Daniell kernel smoothing with m=15")
per_3 <- ggplot(data=plotsmooth, aes(x=V1,y=V2))+geom_line(color="blue")+
  labs(x="Frequency / Bandwidth = 0.00683",y="Scaled Periodogram",title="Daniell kernel smoothing with m=4")

png("smoothing.png",width=480*2,height = 480*2)
grid.arrange(per_1,per_2,per_3,layout_matrix=(matrix(c(1,1,2,3), 2, 2, byrow = TRUE)))
dev.off()


f1 <- plotsmooth[plotsmooth$V4==max1,]$V1
f2 <- plotsmooth[plotsmooth$V4==max2,]$V1
f3 <- plotsmooth[plotsmooth$V4==max3,]$V1
f4 <- plotsmooth[plotsmooth$V4==max4,]$V1
f5 <- plotsmooth[plotsmooth$V4==max5,]$V1
f6 <- plotsmooth[plotsmooth$V4==max6,]$V1
f0 <- plotsmooth[plotsmooth$V4==max(plotsmooth$V4),]$V1

c6 = cos(2*pi*t*f6); s6 = sin(2*pi*t*f6)
c5 = cos(2*pi*t*f5); s5 = sin(2*pi*t*f5)
c4 = cos(2*pi*t*f4); s4 = sin(2*pi*t*f4)
c3 = cos(2*pi*t*f3); s3 = sin(2*pi*t*f3)
c2 = cos(2*pi*t*f2); s2 = sin(2*pi*t*f2)
c1 = cos(2*pi*t*f1); s1 = sin(2*pi*t*f1)
c0 = cos(2*pi*t*f0); s0 = sin(2*pi*t*f0)

fit1 = lm(r.fit~c0+s0+c1+s1+c2+s2+c3+s3+c4+s4+c5+s5+c6+s6)
summary(fit1)
fit11 = lm(r.fit~s4+c5)
fit12 = lm(r.fit~s0+s4+c5)
summary(fit11)
summary(fit12)



max1 <- max((specvals %>% filter(V1<0.1))$V2)
max2 <- max((specvals %>% filter(V1<0.13))$V2)
max3 <- max((specvals %>% filter(V1<0.2))$V2)
max4 <- max((specvals %>% filter(V1<0.2&V1>0.175))$V2)
max5 <- max((specvals %>% filter(V1<0.25))$V2)
max6 <- max((specvals %>% filter(V1<0.29))$V2)
max7 <- max((specvals %>% filter(V1<0.4))$V2)
max8 <- max((specvals %>% filter(V1<0.35&V1>0.32))$V2)
max9 <- max((specvals %>% filter(V1<0.4&V1>0.35))$V2)
max10 <- max((specvals %>% filter(V1<0.45&V1>0.4))$V2)

ggplot(data=specvals, aes(x=V1,y=V2))+geom_line(color="blue")+geom_point(data=specvals[specvals$V2==max(specvals$V2),],pch=21, fill=NA, size=4, colour="red", stroke=1)+
  geom_point(data=specvals[specvals$V2==max1,],pch=21, fill=NA, size=4, colour="red", stroke=1) +
  geom_point(data=specvals[specvals$V2==max2,],pch=21, fill=NA, size=4, colour="red", stroke=1) +
  geom_point(data=specvals[specvals$V2==max3,],pch=21, fill=NA, size=4, colour="red", stroke=1) +
  geom_point(data=specvals[specvals$V2==max4,],pch=21, fill=NA, size=4, colour="red", stroke=1) +
  geom_point(data=specvals[specvals$V2==max5,],pch=21, fill=NA, size=4, colour="red", stroke=1) +
  geom_point(data=specvals[specvals$V2==max6,],pch=21, fill=NA, size=4, colour="red", stroke=1) +
  geom_point(data=specvals[specvals$V2==max7,],pch=21, fill=NA, size=4, colour="red", stroke=1) +
```

```
  geom_point(data=specvals[specvals$V2==max8,],pch=21, fill=NA, size=4, colour="red", stroke=1) +
  geom_point(data=specvals[specvals$V2==max9,],pch=21, fill=NA, size=4, colour="red", stroke=1) +
  geom_point(data=specvals[specvals$V2==max10,],pch=21, fill=NA, size=4, colour="red", stroke=1) +
  labs(x="Frequency",y="Spectrum",title="AR(27) spectrum estimation")
png("ARSmoothing.png",width=480)
dev.off()

f1 <- specvals[specvals$V2==max1,]$V1
f2 <- specvals[specvals$V2==max2,]$V1
f3 <- specvals[specvals$V2==max3,]$V1
f4 <- specvals[specvals$V2==max4,]$V1
f5 <- specvals[specvals$V2==max5,]$V1
f6 <- specvals[specvals$V2==max6,]$V1
f7 <- specvals[specvals$V2==max7,]$V1
f8 <- specvals[specvals$V2==max8,]$V1
f9 <- specvals[specvals$V2==max9,]$V1
f10 <- specvals[specvals$V2==max10,]$V1
f0 <- specvals[specvals$V2==max(specvals$V2),]$V1

c10 = cos(2*pi*t*f10); s10 = sin(2*pi*t*f10)
c9 = cos(2*pi*t*f9); s9 = sin(2*pi*t*f9)
c8 = cos(2*pi*t*f8); s8 = sin(2*pi*t*f8)
c7 = cos(2*pi*t*f7); s7 = sin(2*pi*t*f7)
c6 = cos(2*pi*t*f6); s6 = sin(2*pi*t*f6)
c5 = cos(2*pi*t*f5); s5 = sin(2*pi*t*f5)
c4 = cos(2*pi*t*f4); s4 = sin(2*pi*t*f4)
c3 = cos(2*pi*t*f3); s3 = sin(2*pi*t*f3)
c2 = cos(2*pi*t*f2); s2 = sin(2*pi*t*f2)
c1 = cos(2*pi*t*f1); s1 = sin(2*pi*t*f1)
c0 = cos(2*pi*t*f0); s0 = sin(2*pi*t*f0)

fit2 = lm(r.fit~c0+s0+c1+s1+c2+s2+c3+s3+c4+s4+c5+s5+c6+s6+c7+s7+c8+s8+c9+s9+c10+s10)
summary(fit2)
fit21 = lm(r.fit~c0+s3+s6+c7+s8+s9)
fit22 = lm(r.fit~c0+s3+s5+c6+s6+c7+s8+s9)
summary(fit21)
summary(fit22)

AIC(fit11,fit12,fit21,fit22)
BIC(fit11,fit12,fit21,fit22)

plotperiod <- setDT(as.data.frame(cbind(r.fit,fit11$fitted)))
plotperiod <- cbind(alpha_weekly$V1[which(!is.na(alpha_weekly$alpha))][2:1246],plotperiod)
plotperiod <- cbind(plotperiod,fit11$residuals)
colnames(plotperiod)[1]="date"
ggplot(data=plotperiod,aes(x=date))+geom_line(aes(y=r.fit),color="blue")+geom_line(aes(y=V2),color="red")



#Test for which arima model is most accurate
ar1 <- Arima(1/alpha_weekly$alpha_weekly, order=c(3,1,0))
ar4 <- Arima(1/alpha_weekly$alpha_weekly, order=c(7,1,0))
ar2 <- Arima(1/alpha_weekly$alpha_weekly, order=c(13,1,0))
Arima(1/alpha_weekly$alpha_weekly, order=c(13,1,1))
Arima(1/alpha_weekly$alpha_weekly, order=c(7,1,1))
Arima(1/alpha_weekly$alpha_weekly, order=c(6,1,1))
Arima(1/alpha_weekly$alpha_weekly, order=c(5,1,1))
Arima(1/alpha_weekly$alpha_weekly, order=c(4,1,1))
Arima(1/alpha_weekly$alpha_weekly, order=c(3,1,1))
Arima(1/alpha_weekly$alpha_weekly, order=c(0,1,1))
ar3 <- Arima(1/alpha_weekly$alpha_weekly, order=c(25,1,0))
auto.arima(1/alpha_weekly$alpha_weekly)
plot(forecast(ar2,20), include = 40)
plot(forecast(auto.arima(1/alpha_weekly$alpha_weekly),20), include = 40)


par(mfrow=c(2,2))
png("acfpacfplot.png",width=480*3)
```

```
par(mfrow=c(2,2))
acf(fit11$residuals,10000,main="")
pacf(fit11$residuals,10000,main="")
acf(fit11$residuals,main="")
pacf(fit11$residuals,main="")
dev.off()
par(mfrow=c(1,1))

Arima(1/alpha_weekly$alpha_weekly, order=c(0,1,1))
Arima(1/alpha_weekly$alpha_weekly, order=c(1,1,1))
Arima(1/alpha_weekly$alpha_weekly, order=c(2,1,1))
Arima(1/alpha_weekly$alpha_weekly, order=c(3,1,1))
Arima(1/alpha_weekly$alpha_weekly, order=c(4,1,1))
Arima(1/alpha_weekly$alpha_weekly, order=c(7,1,1))
Arima(1/alpha_weekly$alpha_weekly, order=c(13,1,1))

#Estimate the values of ARIMA/GARCH model
summary(garchFit(formula = ~arma(4,1)+garch(4,4),data=fit11$residuals,cond.dist = "sstd",trace=FALSE))


png("acfpacfplotsquared.png",width=480*3)
par(mfrow=c(2,1))
acf(x1$resid[which(!is.na((x1$resid)))]^2,main="Squared residuals of the ARMA(4,1,1) model")
pacf(x1$resid[which(!is.na((x1$resid)))]^2,main="Squared residuals of the ARMA(4,1,1) model")
dev.off()
par(mfrow=c(1,1))


##Summaries for varying ARIMA/GARCH models dependent on parameter values and conditional distribution
summary(garchFit(formula = ~arma(4,1)+garch(1,1),data=fit11$residuals,cond.dist = "sstd",trace=FALSE))
summary(garchFit(formula = ~arma(3,1)+garch(1,1),data=fit11$residuals,cond.dist = "sstd",trace=FALSE))
summary(garchFit(formula = ~arma(1,1)+garch(1,1),data=fit11$residuals,cond.dist = "sstd",trace=FALSE))
summary(garchFit(formula = ~arma(1,1)+garch(1,1),data=fit11$residuals,cond.dist = "std",trace=FALSE))
summary(garchFit(formula = ~arma(4,1)+garch(1,1),data=fit11$residuals,trace=FALSE))
summary(garchFit(formula = ~arma(4,1)+garch(4,4),data=fit11$residuals,trace=FALSE))
gf44 <- garchFit(formula = ~arma(4,1)+garch(4,4),data=fit11$residuals,trace=FALSE)
gf11 <- garchFit(formula = ~arma(4,1)+garch(1,1),data=fit11$residuals,trace=FALSE)



used <- garchFit(formula = ~arma(4,1)+garch(1,1),data=fit11$residuals,cond.dist = "sstd",trace=FALSE,include.mean = FALSE)
simmed <- garchSim(spec = garchSpec(used), n = 100, n.start = 1000, extended = FALSE)


##Rolling forecasts
model<-ugarchspec(variance.model = list(model = "sGARCH", garchOrder = c(1, 1)),
                  mean.model = list(armaOrder = c(4, 1), include.mean = FALSE),
                  distribution.model = "sstd")
modelfit2<-ugarchfit(spec=model,data=fitres)
mydata <- fitres
spec = getspec(modelfit);
setfixed(spec) <- as.list(coef(modelfit));
forecast = ugarchforecast(spec, data = mydata[1:1245,],n.ahead = 1,n.roll=1200,out.sample = 1200)
forecast2 = ugarchforecast(spec, data = mydata[1:1245,],n.ahead = 1,n.roll=104,out.sample = 104)
forecast3 = ugarchforecast(spec, data = mydata[1:1245,],n.ahead = 50)
head(sigma(forecast));
resfit <- mydata$V2[45:1245]-fitted(forecast)


modelfit11<-ugarchfit(spec=ugarchspec(variance.model = list(model = "sGARCH", garchOrder = c(1, 1)),
                                    mean.model = list(armaOrder = c(4, 1),
```

```
                                                    include.mean = FALSE),
                                                    distribution.model = "sstd"),data=fit11$residuals)

##Calculations of key values
2*10-2*likelihood(modelfit)
log(1245)*10-2*likelihood(modelfit)
model <- ugarchspec(variance.model = list(model = "sGARCH", garchOrder = c(1, 1)),
               mean.model = list(armaOrder = c(4, 1), include.mean = FALSE),
               distribution.model = "sstd",fixed.pars = list(ar1=0.1725075492,ar2=-0.0050270796,ar3=0.1141669267,ar4=0.0476463948,ma1=-0.9054647368,omega=0.0001406652,alpha1=0.2569652889,beta1=0.7827887869,skew=1.6032409401,shape=3.0220608649))
png("Forecast.png",width = 480*3,height = 480*2)
par(mfrow=c(2,1))
plot(forecast)
2
0
plot(forecast2)
2
0
dev.off()

##Constructing weekly risk free
rf <- fread(file.path(data_dir,"FamaFrench.csv"), sep = ",")
rf <- rf %>% mutate(date=ymd(date))
testdate <- setDT(as.data.frame(interpolation_scheme))
testdate <- testdate %>% add_column(new_col = NA) %>% mutate(new_col=as.numeric(new_col))
colnames(testdate)=c("date","rf")
for(i in 1:length(testdate$date)){
  if(testdate[i,1] %in% rf$date){
    testdate[i,2] = (rf %>% filter(date %in% testdate[i,1]))$rf
  }
  else if(dim((rf %>% filter(date %in% testdate[i-1,1])))[1]>0 & dim((rf %>% filter(date %in% testdate[i+1,1])))[1]>0){
    testdate[i,2] <- mean((rf %>% filter(date %in% testdate[i-1,1]))$rf,(rf %>% filter(date %in% testdate[i+1,1]))$rf)
  }
  else if(dim((rf %>% filter(date %in% testdate[i-1,1])))[1]==0 & dim((rf %>% filter(date %in% testdate[i+1,1])))[1]>0){
    testdate[i,2] <- (rf %>% filter(date %in% testdate[i+1,1]))$rf
  }
  else if(dim((rf %>% filter(date %in% testdate[i-1,1])))[1]>0 & dim((rf %>% filter(date %in% testdate[i+1,1])))[1]==0){
    testdate[i,2] <- (rf %>% filter(date %in% testdate[i-1,1]))$rf
  }

  else if(dim((rf %>% filter(date %in% (as.Date(testdate[i,]$date)-1))))[1]>0 & dim((rf %>% filter(date %in% (as.Date(testdate[i,]$date)+1))))[1]>0){
    testdate[i,2] <- mean((rf %>% filter(date %in% (as.Date(testdate[i,]$date)-1)))$rf,(rf %>% filter(date %in% (as.Date(testdate[i,]$date)+1)))$rf)
  }
  else if(dim((rf %>% filter(date %in% (as.Date(testdate[i,]$date)-1))))[1]==0 & dim((rf %>% filter(date %in% (as.Date(testdate[i,]$date)+1))))[1]>0){
    testdate[i,2] <- (rf %>% filter(date %in% (as.Date(testdate[i,]$date)+1)))$rf
  }
  else if(dim((rf %>% filter(date %in% (as.Date(testdate[i,]$date)-1))))[1]>0 & dim((rf %>% filter(date %in% (as.Date(testdate[i,]$date)+1))))[1]==0){
    testdate[i,2] <- (rf %>% filter(date %in% (as.Date(testdate[i,]$date)-1)))$rf
  }


  else if(dim((rf %>% filter(date %in% (as.Date(testdate[i,]$date)-2))))[1]>0 & dim((rf %>% filter(date %in% (as.Date(testdate[i,]$date)+2))))[1]>0){
    testdate[i,2] <- mean((rf %>% filter(date %in% (as.Date(testdate[i,]$date)-2)))$rf,(rf %>% filter(date %in% (as.Date(testdate[i,]$date)+2)))$rf)
  }
  else if(dim((rf %>% filter(date %in% (as.Date(testdate[i,]$date)-2))))[1]==0 & dim((rf %>% filter(date %in% (as.Date(testdate[i,]$date)+2))))[1]>0){
    testdate[i,2] <- (rf %>% filter(date %in% (as.Date(testdate[i,]$date)+2)))$rf
  }
```

```
    else if(dim((rf %>% filter(date %in% (as.Date(testdate[i,]$date)-2))))[1]>0 & dim((rf %>% filter(date %in% (as.Date(testdate[i,]$date)+2))))[1]==0){
      testdate[i,2] <- (rf %>% filter(date %in% (as.Date(testdate[i,]$date)-2)))$rf
    }

    else{print(i)}
  }

  ##LJI
  ATM_vol <- NULL
  for(i in 1:(length(interpolation_scheme)-1)){
    Option_sel <- Option_P %>% filter(date>=interpolation_scheme[i]&date<interpolation_scheme[i+1])
    ATM_vol[i] <- mean(Option_sel$atmvol)
    print(i)
  }
  psi_we2 <- psi$psi
  for(i in 1:(length(interpolation_scheme)-1)){
    if(!is.na(psi_we2[i])&psi_we2[i]>10){
      psi_we2[i]=psi_we2[i-1]
    }
    print(i)
  }
  for(i in 1:(length(interpolation_scheme)-1)){
    if(!is.na(LJV3[i])&LJV3[i]>0.025){
      LJV3[i]=LJV3[i-1]
    }
    print(i)
  }

  #Determination of limit for big jump
  kt <- NULL
  kt1 <- NULL
  fw <- NULL
  for(i in 1:(length(interpolation_scheme)-1)){
    Option_sel <- Option_P %>% filter(date>=interpolation_scheme[i]&date<interpolation_scheme[i+1])
    kt1 <- NULL
    if(dim(Option_sel)[1]>0){
      for(j in unique(Option_sel$comdate)){
        Option_sell <- Option_sel %>% filter(comdate==j)
        kt1<-rbind(kt1,Option_sel[dim(Option_sell)[1]])
      }
      kt[i] <- median(kt1$strike)
      fw[i] <- median(kt1$ForwardPrice)
    }
    print(i)
  }

  #Different cut-off limit
  plot(psi$psi*exp(-alpha_we$alpha_weekly*log(kt)/log(925)*6.868*ATM_vol)/alpha_we$alpha_weekly,type="l")
  plot(psi$psi*exp(-alpha_we$alpha_weekly*fw/kt*6.868*ATM_vol)/alpha_we$alpha_weekly,type="l")
  plot(psi$psi*exp(-alpha_we$alpha_weekly*6.868*ATM_vol)/alpha_we$alpha_weekly,type="l")

  ##Melt for plots
  shap <- setDT(as.data.frame(alpha_we$V1))
  shap <- cbind(shap,psi*exp(-alpha_we$alpha_weekly*3.5*ATM_vol)/alpha_we$alpha_weekly,
                psi*exp(-alpha_we$alpha_weekly*6.868*ATM_vol)/alpha_we$alpha_weekly,

                psi*exp(-alpha_we$alpha_weekly*1500/925*6.868*ATM_vol)/alpha_we$alpha_weekly,

                psi*exp(-alpha_we$alpha_weekly*log(kt)/log(925)*6.868*ATM_vol)/alpha_we$alpha_weekly)
  colnames(shap)=c("date","3.5","6.868","1500/925*6.868","log(xt)/log(925)*6.868")
  dd = melt(shap,id=c("date"))
  ggplot(dd) + geom_line(aes(x=date, y=value, colour=variable)) +
    scale_colour_manual(values=c("orange","red","green","blue"))
  LJI <- psi*exp(-alpha_we$alpha_weekly*log(kt)/log(925)*6.868*ATM_vol)/alpha_we$alpha_weekly
  shap2 <- setDT(as.data.frame(interpolation_scheme[2:1304]))
```

```
shap2 <- cbind(shap2,psi*exp(-alpha_we$alpha_weekly*6.868*ATM_vol)/alpha_we$alpha_weekly,
               psi*exp(-alpha_we$alpha_weekly*1500/925*6.868*ATM_vol)/alpha_we$alpha_weekly,
               psi*exp(-alpha_we$alpha_weekly*log(kt)/log(925)*6.868*ATM_vol)/alpha_we$alpha_weekly)
colnames(shap2)=c("date","6.868","1500/925*6.868","log(xt)/log(925)*6.868")
dd2 = melt(shap2,id=c("date"))
ggplot(dd2) + geom_line(aes(x=date, y=value, colour=variable)) +
  scale_colour_manual(values=c("red","green","blue"))


####LJV
cor(shap2[,2:4],use="complete.obs")
k <- log(kt)/log(925)*6.868
alpha_const <- mean(alpha_we$alpha_weekly,na.rm=TRUE)
LJV1<-as.numeric(interpolation_scheme[2:1304]-interpolation_scheme[1:1303])*psi*exp(-alpha_we$alpha_weekly*k*ATM_vol)*(alpha_we$alpha_weekly*k*ATM_vol*(alpha_we$alpha_weekly*k*ATM_vol+2)+2)/(alpha_we$alpha_weekly^3)
LJV2<-as.numeric(interpolation_scheme[2:1304]-interpolation_scheme[1:1303])*psi*exp(-alpha_const*k*ATM_vol)*(alpha_const*k*ATM_vol*(alpha_const*k*ATM_vol+2)+2)/(alpha_const^3)
LJV3<-as.numeric(interpolation_scheme[2:1304]-interpolation_scheme[1:1303])*exp(-alpha_we$alpha_weekly*k*ATM_vol)*(alpha_we$alpha_weekly*k*ATM_vol*(alpha_we$alpha_weekly*k*ATM_vol+2)+2)/(alpha_we$alpha_weekly^3)
plot(x=alpha_we$V1,y=LJV1,type="l")
plot(x=alpha_we$V1,y=LJV2,type="l")
plot(x=alpha_we$V1,y=LJV3,type="l")
LJVplot <- setDT(as.data.frame(alpha_we$V1))
LJVplot <- cbind(LJVplot,LJV1,LJV2,LJV3)
LJV1p <- ggplot(data=LJVplot,aes(x=LJVplot$`alpha_we$V1`))+geom_line(aes(y=LJV1),color="blue")
LJV2p <- ggplot(data=LJVplot,aes(x=LJVplot$`alpha_we$V1`))+geom_line(aes(y=LJV2),color="blue")
LJV3p <- ggplot(data=LJVplot,aes(x=LJVplot$`alpha_we$V1`))+geom_line(aes(y=LJV3),color="blue")
grid.arrange(LJV1p,LJV2p,LJV3p,nrow=3)



##Prediction regression
BAB <- fread(file.path(data_dir,"BAB.csv"), sep = ";")
QMJ <- fread(file.path(data_dir,"QMJ.csv"), sep = ";")
UMD <- fread(file.path(data_dir,"UMD.csv"), sep = ";")
HML <- fread(file.path(data_dir,"HML.csv"), sep = ";")
SMB <- fread(file.path(data_dir,"SMB.csv"), sep = ";")
MKT <- fread(file.path(data_dir,"MKT.csv"), sep = ";")
RF <- fread(file.path(data_dir,"RF.csv"), sep = ";")
FAF <- fread(file.path(data_dir,"FFFull.csv"), sep = ",")

BAB <- BAB %>% mutate(DATE=as.Date(DATE,"%d-%m-%Y"),USA=as.numeric(sub("%","",USA))/100) %>% select(c("DATE","USA"))
QMJ <- QMJ %>% mutate(DATE=as.Date(DATE,"%d-%m-%Y"),USA=as.numeric(sub("%","",USA))/100) %>% select(c("DATE","USA"))
UMD <- UMD %>% mutate(DATE=as.Date(DATE,"%d-%m-%Y"),USA=as.numeric(sub("%","",USA))/100) %>% select(c("DATE","USA"))
HML <- HML %>% mutate(DATE=as.Date(DATE,"%d-%m-%Y"),USA=as.numeric(sub("%","",USA))/100) %>% select(c("DATE","USA"))
SMB <- SMB %>% mutate(DATE=as.Date(DATE,"%d-%m-%Y"),USA=as.numeric(sub("%","",USA))/100) %>% select(c("DATE","USA"))
MKT <- MKT %>% mutate(DATE=as.Date(DATE,"%d-%m-%Y"),USA=as.numeric(sub("%","",USA))/100) %>% select(c("DATE","USA"))
RF <- RF %>% mutate(DATE=as.Date(RF$DATE,"%d-%m-%Y"),`Risk Free Rate`=as.numeric(sub("%","",`Risk Free Rate`))/100)
FAF <- FAF %>% mutate(Date=ymd(Date))
```

```
FAF[,2:7] = FAF[,2:7]/100

colnames(BAB)[1]="date"
colnames(QMJ)[1]="date"
colnames(FAF)[1]="date"
colnames(UMD)[1]="date"
colnames(HML)[1]="date"
colnames(SMB)[1]="date"
colnames(MKT)[1]="date"
colnames(RF)=c("date","rf")

BAB <- BAB %>% filter(date>="1996-01-01" & date<="2020-12-31")
QMJ <- QMJ %>% filter(date>="1996-01-01" & date<="2020-12-31")
FAF <- FAF %>% filter(date>="1996-01-01" & date<="2020-12-31")
UMD <- UMD %>% filter(date>="1996-01-01" & date<="2020-12-31")
HML <- HML %>% filter(date>="1996-01-01" & date<="2020-12-31")
SMB <- SMB %>% filter(date>="1996-01-01" & date<="2020-12-31")
MKT <- MKT %>% filter(date>="1996-01-01" & date<="2020-12-31")
RF <- RF %>% filter(date>="1996-01-01" & date<="2020-12-31")

##Convert daily data to weekly return
BAB_we <- QMJ_we <- UMD_we <- HML_we <- SMB_we <- MKT_we <- RF_we <- NULL
for(i in 1:(length(interpolation_scheme)-1)){
  BAB_sel <- BAB %>% filter(date>=interpolation_scheme[i]&date<interpolation_scheme[i+1])
  QMJ_sel <- QMJ %>% filter(date>=interpolation_scheme[i]&date<interpolation_scheme[i+1])
  UMD_sel <- UMD %>% filter(date>=interpolation_scheme[i]&date<interpolation_scheme[i+1])
  HML_sel <- HML %>% filter(date>=interpolation_scheme[i]&date<interpolation_scheme[i+1])
  SMB_sel <- SMB %>% filter(date>=interpolation_scheme[i]&date<interpolation_scheme[i+1])
  MKT_sel <- MKT %>% filter(date>=interpolation_scheme[i]&date<interpolation_scheme[i+1])
  RF_sel <- RF %>% filter(date>=interpolation_scheme[i]&date<interpolation_scheme[i+1])
  return1 <- 1
  return2 <- 1
  return3 <- 1
  return4 <- 1
  return5 <- 1
  return6 <- 1
  return7 <- 1
  for(j in 1:dim(BAB_sel)[1]){
    return1 <- return1*(1+BAB_sel[j,2])
    return2 <- return2*(1+QMJ_sel[j,2])
    return3 <- return3*(1+RF_sel[j,2])
    return4 <- return4*(1+UMD_sel[j,2])
    return5 <- return5*(1+HML_sel[j,2])
    return6 <- return6*(1+SMB_sel[j,2])
    return7 <- return7*(1+MKT_sel[j,2])
  }
  BAB_we[i] <- as.numeric(return1-1)
  QMJ_we[i] <- as.numeric(return2-1)
  RF_we[i]  <- as.numeric(return3-1)
  UMD_we[i] <- as.numeric(return4-1)
  HML_we[i] <- as.numeric(return5-1)
  SMB_we[i] <- as.numeric(return6-1)
  MKT_we[i] <- as.numeric(return7-1)
  print(i)
}

##Summary statistics for key measures and returns
Whole <- setDT(as.data.frame(alpha_we$V1))
Whole <- cbind(Whole,MKT_we,SMB_we,HML_we,UMD_we,QMJ_we,BAB_we,alpha_we$alpha_weekly,LJI,LJV1,LJV2,LJV3)
summary(Whole)
round(as.numeric(acf(MKT_we*100,lag.max=1,plot=FALSE)$acf),2)
round(as.numeric(acf(SMB_we*100,lag.max=1,plot=FALSE)$acf),2)
round(as.numeric(acf(HML_we*100,lag.max=1,plot=FALSE)$acf),2)
round(as.numeric(acf(UMD_we*100,lag.max=1,plot=FALSE)$acf),2)
```

```r
round(as.numeric(acf(QMJ_we*100,lag.max=1,plot=FALSE)$acf),2)
round(as.numeric(acf(BAB_we*100,lag.max=1,plot=FALSE)$acf),2)
round(as.numeric(acf(alpha_we[which(!is.na(alpha_we$alpha_weekly))]$alpha_weekly,lag.max=1,plot=FALSE)$acf),2)
round(as.numeric(acf(dtLJI[which(!is.na(dtLJI$LJI))]$LJI*52*100,lag.max=1,plot=FALSE)$acf),2)
round(as.numeric(acf(LJVplot[which(!is.na(LJVplot$LJV1))]$LJV1*52*100,lag.max=1,plot=FALSE)$acf),2)
round(as.numeric(acf(LJVplot[which(!is.na(LJVplot$LJV2))]$LJV2*52*100,lag.max=1,plot=FALSE)$acf),2)
round(as.numeric(acf(LJVplot[which(!is.na(LJVplot$LJV3))]$LJV3*52*100,lag.max=1,plot=FALSE)$acf),2)

#One-week lagged
Whole2 <- setDT(as.data.frame(alpha_we$V1))
Whole2 <- cbind(Whole2[1:1302],MKT_we[2:1303],SMB_we[2:1303],HML_we[2:1303],UMD_we[2:1303],QMJ_we[2:1303],BAB_we[2:1303],alpha_we$alpha_weekly[1:1302],LJI[1:1302],LJV1[1:1302],LJV2[1:1302],LJV3[1:1302])
colnames(Whole2) = c("date","MKT","SMB","HML","UMD","QMJ","BAB","alp","LJI","LJV1","LJV2","LJV3")
round(cor(Whole2[complete.cases(Whole2),2:12]),2)[1:6,7:11]



##Regressions based on a monthly basis
BAB_MO <- fread(file.path(data_dir,"BAB_MO.csv"), sep = ";")
QMJ_MO <- fread(file.path(data_dir,"QMJ_MO.csv"), sep = ";")
MKT_MO <- fread(file.path(data_dir,"MKT_MO.csv"), sep = ";")
VIX_MO <- fread(file.path(data_dir,"VIX_MO.csv"), sep = ";")
VIX_WE <- fread(file.path(data_dir,"VIX_WE.csv"), sep = ";")

BAB_MO <- BAB_MO %>% mutate(DATE=as.Date(DATE,"%d-%m-%Y"),USA=as.numeric(sub("%","",USA))/100) %>% select(c("DATE","USA"))
QMJ_MO <- QMJ_MO %>% mutate(DATE=as.Date(DATE,"%d-%m-%Y"),USA=as.numeric(sub("%","",USA))/100) %>% select(c("DATE","USA"))
MKT_MO <- MKT_MO %>% mutate(DATE=as.Date(DATE,"%d-%m-%Y"),USA=as.numeric(sub("%","",USA))/100) %>% select(c("DATE","USA"))
VIX_MO <- VIX_MO %>% mutate(Date=as.Date(Date,"%d-%m-%Y"))
VIX_WE <- VIX_WE %>% mutate(date=as.Date(date,"%d-%m-%Y"))
colnames(BAB_MO)[1]="date"
colnames(QMJ_MO)[1]="date"
colnames(MKT_MO)[1]="date"
colnames(VIX_MO)[1]="date"

BAB_MO <- BAB_MO %>% filter(date>="1996-01-01" & date<="2020-12-31")
QMJ_MO <- QMJ_MO %>% filter(date>="1996-01-01" & date<="2020-12-31")
MKT_MO <- MKT_MO %>% filter(date>="1996-01-01" & date<="2020-12-31")
#Convert measures to monthly
LJV1_MO <- NULL
LJV1_MO[1] <- mean((LJVplot %>% filter(LJVplot$`alpha_we$V1`<=BAB_MO[1]$date))$LJV1)
for(i in 2:dim(BAB_MO)[1]){
  LJV_sel <- LJVplot %>% filter(LJVplot$`alpha_we$V1`<=BAB_MO[i]$date & LJVplot$`alpha_we$V1`>BAB_MO[i-1]$date)
  LJV1_MO[i] <- mean(LJV_sel$LJV1,na.rm=TRUE)
}
summary(lm(MKT_MO[5:300]$USA~LJV1_MO[1:296]))


first <- lm(MKT_MO[7:300]$USA~LJV1_MO[1:294])
summary(lm(MKT_MO[2:300]$USA~VIX_MO$`Adj Close`[1:299]^2))
summary(lm(MKT_MO[2:300]$USA~LJV1_MO[1:299]+VIX_MO$`Adj Close`[1:299]^2))

lag1=6
first <- lm(BAB_MO[(2+lag1):300]$USA~LJV1_MO[2:(300-lag1)])
NW <- NeweyWest(first,lag=2*lag1,prewhite = F,adjust=T)
round(coeftest(first,vcov=NW),3)
summary(first)
sec <- lm(QMJ_MO[(1+lag1):300]$USA~VIX_MO2[1:(300-lag1)]$C^2)
NW <- NeweyWest(sec,lag=2*lag1,prewhite = F,adjust=T)
```

```
round(coeftest(sec,vcov=NW),3)
summary(sec)
third <- lm(QMJ_MO[(2+lag1):300]$USA~alpha_mo2[1:(299-lag1)])
NW <- NeweyWest(third,lag=2*lag1,prewhite = F,adjust=T)
round(coeftest(third,vcov=NW),3)
summary(third)
fourth <- lm(BAB_MO[(2+lag1):300]$USA~LJV1_MO[2:(300-lag1)]+VIX_MO2[2:(300-lag1)]$C^2+alpha_mo2[1:(299-lag1)])
NW <- NeweyWest(fourth,lag=2*lag1,prewhite = F,adjust=T)
round(coeftest(fourth,vcov=NW),3)
summary(fourth)
fifth <- lm(QMJ_MO[(2+lag1):300]$USA~LJV1_MO[2:(300-lag1)]+alpha_mo2[1:(299-lag1)])
vif(fifth)
waldtest(sec,fourth)


alpha3 <- VIX_MO2[1:(300-lag1)]$C^2
alpha3[16] <- LJV1_MO[16]
alpha3[58] <- LJV1_MO[58]



#Times of distress
alpha_we <- fread(file.path(data_dir,"alpha_weekly2904.csv"), sep = ",")
alpha_we <- cbind(interpolation_scheme[2:1304],alpha_we)
alpha_we <- alpha_we[,c(1,3)]
alpha_we1 <- alpha_we %>% filter(V1>="2020-01-01" & V1<"2021-01-01")
alpha_we2 <- alpha_we1$alpha_weekly/100
alpha_we2 <- alpha_we2[2:53]
LJV_WE1 <- LJVplot %>% filter(alpha_we$V1>="2020-01-01" & alpha_we$V1<"2021-01-01")
VIX_WE <- VIX_WE %>% filter(date>="2020-01-01" & date<"2021-01-01")
VIX_WE4 <- VIX_WE4 %>% filter(date>="2020-01-01" & date<"2021-01-01")
plot(LJV_WE1$LJV1,type="l")
plot(1/alpha_we1$alpha_weekly,type="l")
plot(VIX_WE$Close^2,type="l")
LJV_WE1 <- LJV_WE1[2:53]
#Market returns
mkt_return <- setDT(as.data.frame(interpolation_scheme[2:1304]))
colnames(mkt_return)="date"
mkt_return <- cbind(mkt_return,MKT_we,BAB_we,QMJ_we,UMD_we,SMB_we,HML_we)
mkt_return <- mkt_return %>% filter(date>="2020-01-01" & date<"2021-01-01")

lag1=13
first <- lm(mkt_return[(1+lag1):52]$BAB_we~LJV_WE1[1:(52-lag1)]$LJV1)
NW <- NeweyWest(first,lag=2*lag1,prewhite = F,adjust=T)
round(coeftest(first,vcov=NW),3)
summary(first)
sec <- lm(mkt_return[(1+lag1):52]$HML_we~VIX_WE2[1:(52-lag1)]$Close^2)
NW <- NeweyWest(sec,lag=2*lag1,prewhite = F,adjust=T)
round(coeftest(sec,vcov=NW),3)
summary(sec)
third <- lm(mkt_return[(1+lag1):52]$HML_we~alpha_we2[1:(52-lag1)])
NW <- NeweyWest(third,lag=2*lag1,prewhite = F,adjust=T)
round(coeftest(third,vcov=NW),3)
summary(third)
fourth <- lm(mkt_return[(1+lag1):52]$HML_we~LJV_WE1[1:(52-lag1)]$LJV1+VIX_WE2[1:(52-lag1)]$Close^2+alpha_we2[1:(52-lag1)])
NW <- NeweyWest(fourth,lag=2*lag1,prewhite = F,adjust=T)
round(coeftest(fourth,vcov=NW),3)
summary(fourth)
fifth <- lm(mkt_return[(1+lag1):52]$BAB_we~LJV_WE1[1:(52-lag1)]$LJV1+alpha_we2[1:(52-lag1)])
NW <- NeweyWest(fifth,lag=2*lag1,prewhite = F,adjust=T)
round(coeftest(fifth,vcov=NW),3)
summary(fifth)

VIX_WE4[11] <-LJV_WE1[11]$LJV1
waldtest(fifth,first)
```

```
###Attempt at High-Frequency based calculations
HFdata <- fread(file.path(data_dir,"HFdata.csv"), sep = ";")
HFdata <- HFdata %>% mutate(Hour=substr(Date,11,16),Date=as.Date(substr(Date,1,10),"%d-%m-%Y"),Price = as.numeric(gsub(",",".",HFdata$`ES1 Index - Last Price`)))
HFdata <- HFdata[,c(1,3,4)]
#Split each day into 30 min increments
#First the RV
RV <- NULL
j=1
#Remove missing days
lf <- NULL
j=1
for(i in unique(HFdata$Date)){
  HFdata_sel <- HFdata %>% filter(Date==i)
  if(dim(HFdata_sel)[1]<46){
    lf[j]=as.Date(i)
  }
  j=j+1
}
HFdata <- HFdata %>% filter(!(Date %in% as.Date(lf)))
for(i in unique(HFdata$Date)){
  HFdata_sel <- HFdata %>% filter(Date==i)
  RV[j] <- sum((HFdata_sel[2:dim(HFdata_sel)[1]]$Price-HFdata_sel[1:(dim(HFdata_sel)[1]-1)]$Price)^2)
  j = j + 1
}

#abar
abar <- NULL
j=1
for(i in unique(HFdata$Date)){
  HFdata_sel <- HFdata %>% filter(Date==i)
  abar[j] <- sum(abs(HFdata_sel[3:(dim(HFdata_sel)[1])]$Price-HFdata_sel[2:(dim(HFdata_sel)[1]-1)]$Price)*
                 abs(HFdata_sel[2:(dim(HFdata_sel)[1]-1)]$Price-HFdata_sel[1:(dim(HFdata_sel)[1]-2)]$Price))
  j = j + 1
}
abar = 3*sqrt(pi/2)*
  sqrt(1/length(unique(HFdata$Date))*sum(abar))
##NOI
NOI <- NULL
NOI_t <- 0
##NOI_t
for(t in unique(HFdata$Date)){
  HFdata_sel <- HFdata %>% filter(Date==t)
  for(i in 2:dim(HFdata_sel)[1]){
    if(abs(HFdata_sel$Price[i]-HFdata_sel$Price[i-1])<=abar*((1/dim(HFdata_sel)[1])^0.49)){
      NOI_t = NOI_t + 1
    }
  }
}
#NOI_n
NOI_n <- rep(0,46)
for(i in 2:46){
  for(t in unique(HFdata$Date)){
    HFdata_sel <- HFdata %>% filter(Date==t)
    if(abs(HFdata_sel$Price[i]-HFdata_sel$Price[i-1])<=abar*((1/dim(HFdata_sel)[1])^0.49)){
      NOI_n[i] = NOI_n[i] + 1
    }
  }
}
#NOI
NOI = NOI_t/NOI_n
```

```
#TOD
TOD <- TOD_n <- NULL
#TOD_t
TOD_t <- rep(0,46)
for(i in 2:46){
  for(t in unique(HFdata$Date)){
    HFdata_sel <- HFdata %>% filter(Date==t)

    if(abs(HFdata_sel$Price[i]-HFdata_sel$Price[i-1])<=abar*((1/dim(HFdata_sel)[1])^0.49)){
      TOD_t[i] = TOD_t[i] + (HFdata_sel$Price[i]-HFdata_sel$Price[i-1])^2
    }
  }
}
#TOD_n
TOD_n <- 0
for(i in unique(HFdata$Date)){
  HFdata_sel <- HFdata %>% filter(Date==i)
  TOD_n = TOD_n + sum((HFdata_sel[2:dim(HFdata_sel)[1]]$Price-HFdata_sel[1:(dim(HFdata_sel)[1]-1)]$Price)^2)
}
#TOD
TOD <- NOI*TOD_t/TOD_n

#Now, back to the estimates
CV <- RJV <- LJV <- CV2 <- RJV2 <- LJV2 <- rep(0,length(unique(HFdata$Date)))
HFdata_sel <- HFdata %>% filter(Date==unique(HFdata$Date)[1])
for(i in 2:dim(HFdata_sel)[1]){

  if(abs(HFdata_sel$Price[i]-HFdata_sel$Price[i-1])<=abar*((1/dim(HFdata_sel)[1])^0.49)){
    CV[1] = CV[1] + (HFdata_sel$Price[i]-HFdata_sel$Price[i-1])^2
  }

  if(abs(HFdata_sel$Price[i]-HFdata_sel$Price[i-1])>abar*((1/dim(HFdata_sel)[1])^0.49)){
    RJV[1] = RJV[1] + (HFdata_sel$Price[i]-HFdata_sel$Price[i-1])^2
  }
  if(HFdata_sel$Price[i]-HFdata_sel$Price[i-1]< -abar*((1/dim(HFdata_sel)[1])^0.49)){
    LJV[1] = LJV[1] + (HFdata_sel$Price[i]-HFdata_sel$Price[i-1])^2
  }
  RJV2[1]=RJV[1]
  LJV2[1]=LJV[1]
  CV2[1]=CV[1]
}
a <- NULL
for(j in 2:length(unique(HFdata$Date))){
  HFdata_sel <- HFdata %>% filter(Date==unique(HFdata$Date)[j])
  for(i in 2:dim(HFdata_sel)[1]){
    a[i] = 3*sqrt(CV[j-1])*TOD[i]*((1/dim(HFdata_sel)[1])^0.49)

    if(abs(HFdata_sel$Price[i]-HFdata_sel$Price[i-1])<=abar*((1/dim(HFdata_sel)[1])^0.49)){
      CV[j] = CV[j] + (HFdata_sel$Price[i]-HFdata_sel$Price[i-1])^2
    }

    if(abs(HFdata_sel$Price[i]-HFdata_sel$Price[i-1])<=a[i]*((1/dim(HFdata_sel)[1])^0.49)){
      CV2[j] = CV2[j] + (HFdata_sel$Price[i]-HFdata_sel$Price[i-1])^2
    }

    if(abs(HFdata_sel$Price[i]-HFdata_sel$Price[i-1])>abar*((1/dim(HFdata_sel)[1])^0.49)){
      RJV[j] = RJV[j] + (HFdata_sel$Price[i]-HFdata_sel$Price[i-1])^2
    }

    if(abs(HFdata_sel$Price[i]-HFdata_sel$Price[i-1])>a[i]*((1/dim(HFdata_sel)[1])^0.49)){
      RJV2[j] = RJV2[j] + (HFdata_sel$Price[i]-HFdata_sel$Price[i-1])^2
    }
```

```
    if(HFdata_sel$Price[i]-HFdata_sel$Price[i-1]<-abar*((1/dim(HFdata_sel)[1])^0.49)){
      LJV[j] = LJV[j] + (HFdata_sel$Price[i]-HFdata_sel$Price[i-1])^2
    }
    if(HFdata_sel$Price[i]-HFdata_sel$Price[i-1]<-a[i]*((1/dim(HFdata_sel)[1])^0.49)){
      LJV2[j] = LJV2[j] + (HFdata_sel$Price[i]-HFdata_sel$Price[i-1])^2
    }
  }
}


## MARIMA attempt based on HF-data
TSVAR <- setDT(as.data.frame(cbind(CV,RJV,LJV)))
VARmodel <- define.model(kvar=3,ar=c(1))
short.form(VARmodel$ar.pattern)
Model <- marima(TSVAR,ar.pattern = VARmodel$ar.pattern)
plot(RJV,type="l")
plot(diff(RJV,1),type="l")
plot(diff(RJV,2),type="l")

TSVAR.dif <- define.dif(TSVAR,difference=c(1,1,1,1,1,1))
TSVAR.dif.analysis <- TSVAR.dif$y.dif
TSVAR.dif.analysis_F <- define.dif(TSVAR[1:93],difference=c(1,1,1,1,1,1))$y.dif
Model2 <- marima(TSVAR.dif.analysis_F,ar.pattern = VARmodel$ar.pattern)
short.form(Model$ar.estimates)
Model5 <- define.model(kvar=3, ar=c(1,5,22), ma=0, rem.var=0, reg.var=0)
Marima5 <- marima(ts(TSVAR[1:90, ]), Model5$ar.pattern, Model5$ma.pattern,
                  penalty=1)
nstart <- 261
nstep  <- 10
Forecasts <- arma.forecast(series=TSVAR, marima=Model,
                           nstart=nstart, nstep=nstep )

One.step <- Forecasts$forecasts[, (nstart+1)]
One.step
One.step
Predict <- Forecasts$forecasts[ 2, 91:100]
Predict
stdv<-sqrt(Forecasts$pred.var[2, 2, ])
upper.lim=Predict+stdv*1.645
lower.lim=Predict-stdv*1.645
Out<-rbind(Predict, upper.lim, lower.lim)
print(Out)
# plot results:
plot(x=Forecasts$forecasts[2, ],type="l")
lines(271:280,Predict, type='l')
lines(271:280,upper.lim, type='l')
lines(271:280,lower.lim, type='l')


######
###### Below are calculations that are of interest in the plots
######
###The witdh of the cross section
x<-NULL
Option_p_sort <-  Option_P[order(date,strike_price)]
k=0
for(i in unique(Option_P$comdate)){
  Option_sel <- Option_P %>% filter(comdate==i)
  witdth <- log(Option_sel[dim(Option_sel)[1]]$strike_price/Option_sel[dim(Option_sel)[1]]$ForwardPrice)
  x <- append(x,witdth)
  k=k+1
  print(k)
}
for(i in unique(x4$date)){
  x_sel <- x4 %>% filter(date==i)
  x_max <- append(x_max,max(x_sel$x))
  x_min <- append(x_min,min(x_sel$x))
}
```

```
Picture_P2 <- setDT(Option_P)[,.N, by = c("date")]
Differenced_data <- cbind((alpha_weekly_3 %>% select(V1))[2:1303],diff(1/alpha_weekly_3$alpha,1))



#Values for plots
average_forward <- NULL
for(i in unique(future$date)){
  future_sel <- future %>% filter(date==i)
  average_forward <- append(average_forward,mean(future_sel$ForwardPrice))
}
average_strike <- NULL
k=0
for(i in unique(Option_P$date)){
  option_sel <- Option_P %>% filter(date==i)
  average_strike <- append(average_strike,mean(option_sel$strike_price))
  k=k+1
  print(k)
}

forward_return <- log(average_forward[2:length(average_forward)]/average_forward[1:(length(average_forward)-1)])

##Average number of weekly options
Option_Smooth <- Option_Smooth  %>% mutate(date = interpolation_scheme[findInterval(Option_Smooth $date,interpolation_scheme)] %>% floor_date %>% ymd)
Option_Smooth <- Option_Smooth %>% mutate(atmvol=atmvol*sqrt(timetoexp/365))
limit = 10
for(j in 1:1){
  limit = limit + 0.1
  k=0
  for(i in unique(Option_Smooth$date)){
    k=k+1
    option_sel <- Option_Smooth %>% filter(date==i&log_moneyness < -limit*atmvol)
    x[k,j] <- dim(option_sel)[1]
  }
  print(j)
}
average_early <- averagex %>% filter(date<"2015-01-01")
average_late <- averagex %>% filter(date>="2015-01-01")
average_2020 <- averagex %>% filter(date>="2020-01-01")
average_e <- averagex %>% filter(date<"2010-01-01")
for(j in 1:50){
  #av_ea <- append(av_ea,mean(average_early[[paste0("V",j)]]))
  #av_la <- append(av_la,mean(average_late[[paste0("V",j)]]))
  #av_20 <- append(av_20,mean(average_2020[[paste0("V",j)]]))
  #av <- append(av,mean(averagex[[paste0("V",j)]]))
  av_e <- append(av_e,mean(average_e[[paste0("V",j)]]))
}
limit <- seq(1.1,6,by=0.1)
colnames(av_comb)[2:5] <- c("Average before 2015", "Average after 2015", "Average 2020", "Average overall")
dd = melt(av_comb,id=c("limit"))
ggplot(dd) + geom_line(aes(x=limit, y=value, colour=variable)) +
  scale_colour_manual(values=c("red","green","blue","pink"))


for(i in unique(Option_Smooth$date)){
  option_sel <- Option_Smooth %>% filter(date==i&log_moneyness < -6*atmvol)
  x <- append(x,dim(option_sel)[1])
}

####PLOTS
png("TOD.png",width=480*3)

a<- ggplot(data=Picture_all,aes(x=date,y=N))+geom_line(color="blue")+
  scale_x_date(limits = c(min(Picture_all$date), max(Picture_all$date)),breaks = scales::pretty_breaks(n = 18))+
  labs(x = "Date",y="Number of outstanding options")
b <- ggplot(data=Picture_P,aes(x=date,y=N))+geom_line(color="blue")+
```

```r
  scale_x_date(limits = c(min(Picture_all$date), max(Picture_all$date)),breaks = scales::pretty_breaks(n = 18))+
  labs(x = "Date",y="Number of outstanding puts")
c <- ggplot(data=Picture_C,aes(x=date,y=N))+geom_line(color="blue")+
  scale_x_date(limits = c(min(Picture_all$date), max(Picture_all$date)),breaks = scales::pretty_breaks(n = 18))+
  labs(x = "Date",y="Number of outstanding calls")
ggplot(data=Picture_P2,aes(x=date,y=N))+geom_line(color="blue")+
  scale_x_date(limits = c(min(Picture_all$date), max(Picture_all$date)),breaks = scales::pretty_breaks(n = 18))+
  labs(x = "Date",y="Number of outstanding puts")
ggplot(data=x3,aes(x=date))+geom_line(aes(y=x_max),colour="blue")+geom_line(aes(y=x_min),colour="red")+
  scale_x_date(limits = c(min(Picture_all$date), max(Picture_all$date)),breaks = scales::pretty_breaks(n = 18))+
  labs(x = "Date",y="How out-the-money puts are")

ratioplot <- ggplot()+geom_line(aes(x=Picture_all$date,y=average_forward/average_strike),colour="blue")+
  scale_x_date(limits = c(min(Picture_all$date), max(Picture_all$date)),breaks = scales::pretty_breaks(n = 18))+
  labs(x = "Date",y="Average Forward Price / Average Strike Price")
log_forward <- ggplot()+geom_line(aes(x=Picture_all$date[2:length(Picture_all$date)],y=forward_return),colour="blue")+
  scale_x_date(limits = c(min(Picture_all$date), max(Picture_all$date)),breaks = scales::pretty_breaks(n = 18))+
  labs(x = "Date",y="Daily log-returns on Forward Price")
averagelogmon <- ggplot(data = logmoneyness,aes(x=date,y=V2))+geom_line(color="darkblue")+geom_smooth(method="lm",color="red")+
  scale_x_date(limits = c(min(Picture_all$date), max(Picture_all$date)),breaks = scales::pretty_breaks(n = 18))+
  labs(x = "Date",y="Average log-moneyness")
maxminlogmon <- ggplot(data = logmoneynessmax ,aes(x=date))+geom_line(color="blue",aes(y=x_max))+geom_line(color="red",aes(y=x_min))+

  geom_smooth(method="lm",formula=y~x,color="blue",aes(y=x_max))+geom_smooth(method="lm",formula=y~x,color="red",aes(y=x_min))+
  scale_x_date(limits = c(min(Picture_all$date), max(Picture_all$date)),breaks = scales::pretty_breaks(n = 18))+
  labs(x = "Date",y="Log-moneyness")
ggplot(data=av_comb,aes(x=limit))+geom_line(color="darkblue",aes(y=av_ea))+geom_line(color="green",aes(y=av))+
  geom_line(color="red",aes(y=av_la))+geom_line(color="pink",aes(y=av_20))+labs(x = "Ratio",y="Average weekly options")+
  theme(legend.position="right")

ggplot(dd) + geom_line(aes(x=limit, y=value, colour=variable)) +
  scale_colour_manual(values=c("red","green","blue","pink"))+labs(x = "Ratio",y="Average weekly puts")
a <- ggplot(data=alpha_weekly,aes(x=V1,y=1/alpha))+geom_line(color="blue")+
  scale_x_date(limits = c(min(Picture_all$date), max(Picture_all$date)),breaks = scales::pretty_breaks(n = 18))+
  labs(x = "Year",y="")
b <- ggplot(data=alpha_6_weekly,aes(x=date,y=1/alpha))+geom_line(color="red")+
  scale_x_date(limits = c(min(Picture_all$date), max(Picture_all$date)),breaks = scales::pretty_breaks(n = 18))+
  labs(x = "Year",y="")
c <- ggplot(data=alpha_diff,aes(x=V1,y=V2))+geom_line(color="blue")+
  scale_x_date(limits = c(min(Picture_all$date), max(Picture_all$date)),breaks = scales::pretty_breaks(n = 18))+
  labs(x = "Year",y="")
plot1 <- ggplot(data=plotperiod,aes(x=date))+geom_line(aes(y=r.fit),color="blue")+geom_line(aes(y=V2),color="red")+
  scale_x_date(limits = c(min(Picture_all$date), max(Picture_all$date)),breaks = scales::pretty_breaks(n = 18))+
  labs(x = "Year",y="")
plot2 <- ggplot(data=plotperiod,aes(x=date))+geom_line(aes(y=V3),color="blue")+
  scale_x_date(limits = c(min(Picture_all$date), max(Picture_all$date)),breaks =
```

```
    scales::pretty_breaks(n = 18))+
  labs(x = "Year",y="")

weekly <- ggplot(data=alpha_we,aes(x=V1,y=1/alpha_weekly))+geom_line(color="blue")+
  scale_x_date(limits = c(min(alpha_we$V1), max(alpha_we$V1)),breaks = scales::pretty_breaks(n = 18))+
  
  geom_hline(yintercept=mean(1/alpha_we[which(!is.na(alpha_we$alpha_weekly))]$alpha_weekly),lwd=0.5,color="darkblue")+
  labs(x = "Year",y="",title="Weekly")+theme(plot.title=element_text(hjust=0.5))
monthly <- ggplot(data=alpha_mo,aes(x=V1,y=1/alpha))+geom_line(color="blue")+
  
  geom_hline(yintercept=mean(1/alpha_mo[which(!is.na(alpha_mo$alpha))]$alpha),lwd=0.5, color="darkblue")+
  scale_x_date(limits = c(min(alpha_we$V1), max(alpha_we$V1)),breaks = scales::pretty_breaks(n = 18))+
  labs(x = "Year",y="",title="Monthly")+theme(plot.title=element_text(hjust=0.5))
quarterly <- ggplot(data=alpha_qu,aes(x=V1,y=1/alpha))+geom_line(color="blue")+
  geom_hline(yintercept=mean(1/alpha_qu$alpha),lwd=0.5,color="darkblue")+
  scale_x_date(limits = c(min(alpha_we$V1), max(alpha_we$V1)),breaks = scales::pretty_breaks(n = 18))+
  labs(x = "Year",y="",title="Quarterly")+theme(plot.title=element_text(hjust=0.5))
annual <- ggplot(data=alpha_an,aes(x=V1,y=1/alpha))+geom_line(color="blue")+
  geom_hline(yintercept=mean(1/alpha_an$alpha),lwd=0.5,color="darkblue")+
  scale_x_date(limits = c(min(alpha_we$V1), max(alpha_we$V1)),breaks = scales::pretty_breaks(n = 18))+
  labs(x = "Year",y="",title="Annual")+theme(plot.title=element_text(hjust=0.5))
ddplot1 <- ggplot(dd) + geom_line(aes(x=date, y=value, colour=variable)) +
  scale_colour_manual(values=c("orange","red","green","blue"))+
  scale_x_date(limits = c(min(alpha_we$V1), max(alpha_we$V1)),breaks = scales::pretty_breaks(n = 18))+
  labs(x = "Year",y="")
ddplot2 <- ggplot(dd2) + geom_line(aes(x=date, y=value, colour=variable)) +
  scale_colour_manual(values=c("red","green","blue"))+
  scale_x_date(limits = c(min(alpha_we$V1), max(alpha_we$V1)),breaks = scales::pretty_breaks(n = 18))+
  labs(x = "Year",y="")
LJV1p <- ggplot(data=LJVplot,aes(x=LJVplot$`alpha_we$V1`))+geom_line(aes(y=LJV1),color="blue")+
  scale_x_date(limits = c(min(alpha_we$V1), max(alpha_we$V1)),breaks = scales::pretty_breaks(n = 18))+
  labs(x = "Year",y="",title="LJV")+theme(plot.title=element_text(hjust=0.5))
LJV2p <- ggplot(data=LJVplot,aes(x=LJVplot$`alpha_we$V1`))+geom_line(aes(y=LJV2),color="blue")+
  scale_x_date(limits = c(min(alpha_we$V1), max(alpha_we$V1)),breaks = scales::pretty_breaks(n = 18))+
  labs(x = "Year",y="",title="LJV*")+theme(plot.title=element_text(hjust=0.5))
LJV3p <- ggplot(data=LJVplot,aes(x=LJVplot$`alpha_we$V1`))+geom_line(aes(y=LJV3),color="blue")+
  scale_x_date(limits = c(min(alpha_we$V1), max(alpha_we$V1)),breaks = scales::pretty_breaks(n = 18))+
  labs(x = "Year",y="",title="LJV**")+theme(plot.title=element_text(hjust=0.5))

LJV1pmo <- ggplot(data=MKT_MO,aes(x=date))+geom_line(aes(y=LJV1_MO),color="blue")+
  scale_x_date(limits = c(min(alpha_we$V1), max(alpha_we$V1)),breaks = scales::pretty_breaks(n = 18))+
  labs(x = "Year",y="",title="LJV Monthly")+theme(plot.title=element_text(hjust=0.5))
VIX2 <- ggplot(data=MKT_MO,aes(x=date))+geom_line(aes(y=VIX_MO$C^2),color="blue")+
  scale_x_date(limits = c(min(alpha_we$V1), max(alpha_we$V1)),breaks = scales::pretty_breaks(n = 18))+
  labs(x = "Year",y="",title="VIX Squared")+theme(plot.title=element_text(hjust=0.5))

aw<- ggplot(data=alpha_we1,aes(x=V1))+geom_line(aes(y=alpha_weekly),color="blue")+
  scale_x_date(limits = c(min(alpha_we1$V1), max(alpha_we1$V1)),breaks = scales::pretty_breaks(n = 18))+
  geom_vline(xintercept = as.numeric(key_date),linetype="dotted")+
  geom_vline(xintercept = as.numeric(key_date_fiscal),linetype="dotted",col="green4")+
  geom_vline(xintercept = as.numeric(key_date_health),linetype="dotted",col="red")+
  labs(x = "Month",y="",title="Alpha")+theme(plot.title=element_text(hjust=0.5))+
  scale_y_continuous(labels=scales::scientific)
lw <- ggplot(data=alpha_we1[2:53,],aes(x=V1))+geom_line(aes(y=LJV_WE1$LJV1),color="blue")+
  geom_vline(xintercept = as.numeric(key_date),linetype="dotted")+
```

```r
  geom_vline(xintercept = as.numeric(key_date_fiscal),linetype="dotted",col="green4")+
  geom_vline(xintercept = as.numeric(key_date_health),linetype="dotted",col="red")+
  scale_x_date(limits = c(min(alpha_we1$V1), max(alpha_we1$V1)),breaks = scales::pretty_breaks(n = 18))+
  labs(x = "Month",y="",title="LJV")+theme(plot.title=element_text(hjust=0.5))+
  scale_y_continuous(labels=scales::scientific)
vw <- ggplot(data=alpha_we1[2:53],aes(x=V1))+geom_line(aes(y=VIX_WE3$V2^2),color="blue")+
  geom_vline(xintercept = as.numeric(key_date),linetype="dotted")+
  geom_vline(xintercept = as.numeric(key_date_fiscal),linetype="dotted",col="green4")+
  geom_vline(xintercept = as.numeric(key_date_health),linetype="dotted",col="red")+
  scale_x_date(limits = c(min(alpha_we1$V1), max(alpha_we1$V1)),breaks = scales::pretty_breaks(n = 18))+
  labs(x = "Month",y="",title="VIX Squared")+theme(plot.title=element_text(hjust=0.5))+
  scale_y_continuous(labels=scales::scientific)

ggplot(data=plotTOD,aes(x=V1,y=TOD))+geom_line(colour="blue")+
  scale_x_continuous(limits = c(0,24),breaks = scales::pretty_breaks(n = 12))+
  labs(x = "Time ",y="",title="Time-of-day factor TOD")+theme(plot.title=element_text(hjust=0.5))

grid.arrange(aw,lw,vw,nrow=3)
dev.off()
key_date <- as.Date(c("2020-03-06","2020-03-18","2020-03-27"))
key_date_fiscal <- as.Date(c("2020-03-15","2020-03-23","2020-04-09"))
key_date_health <- as.Date(c("2020-01-30","2020-03-13"))
```