Abstract

This paper investigates the performance of different parametric models, stable and tempered stable distributions, for capturing the tail behaviour of the log-returns. We first define and discuss the properties of the stable and tempered stable random variables. We then show how to estimate their parameters and simulate them from their characteristic functions. Finally, as an illustration, we perform an empirical experiment to explore the performance of different models representing the distribution of log-returns for the S&P500 and DAX indexes.
Modeling Tail Risk with Tempered Stable Distributions: an Overview

Hasan Fallahgoul\textsuperscript{1,2} and Gregoire Loeper\textsuperscript{1,2}

\textsuperscript{1}School of Mathematical Sciences, Clayton Campus, Monash University, VIC, 3800, Australia
\textsuperscript{2}Centre for Quantitative Finance and Investment Strategies\textsuperscript{∗}, Monash University, Australia

November 8, 2018

Abstract

This paper investigates the performance of different parametric models, stable and tempered stable distributions, for capturing the tail behaviour of the log-returns. We first define and discuss the properties of the stable and tempered stable random variables. We then show how to estimate their parameters and simulate them from their characteristic functions. Finally, as an illustration, we perform an empirical experiment to explore the performance of different models representing the distribution of log-returns for the S&P500 and DAX indexes.

Keywords: tail risk, stable distribution, tempered stable distribution, Lévy processes

\textsuperscript{∗}Acknowledgments: We are grateful to Frank Fabozzi, Young (Aaron) Kim, and Stoyan Stoyanov for helpful comments. The Centre for Quantitative Finance and Investment Strategies has been supported by BNP Paribas.
1 Introduction

It is today commonly accepted that returns of financial assets are not normally distributed. One key feature is the existence of a so-called heavy tail in the distribution of those returns, i.e. extreme events occur with a much higher frequency than predicted by a normal distribution, hence the need to use other probability distributions that reflect properly this feature crucial to proper risk management. Heavy tail random variables play an important role for modelling extreme events not only in finance and risk management but also in other fields. Some examples are the distribution of the size of web pages, file size in a computer system, and modelling the distribution of loss in the context of insurance. One wants to model a phenomenon with a heavy tail random variable to account properly for the occurrence of outliers (i.e., very unexpected events) that have an extreme impact, and whose prediction is by essence difficult.

There are two main approaches for modelling tail risk: (i) one can use a parametric distribution that works for both the body and tail of data, or (ii) identifying a threshold and representing only the tail of the distribution. Each approach has some advantages and drawbacks. Since the first approach uses a parametric distribution, calculating a risk measure such as Value-at-Risk (VaR) or Expected Shortfall (ES) is straightforward. On the other hand, by imposing a structure on the data, it induces some model risk. A key step in implementing the second approach is to specify the threshold and to verify the assumption that the tail follows a generalized Pareto distribution. In this paper, we focus on the first approach, which means assume a parametric distribution to model the whole dataset.

There are several definitions of heavy tail or fat tail distributions in the literature, and not everyone agrees on them. For some authors, heavy tail distribution means a distribution with a tail that is heavier than the exponential distribution, which formally

---

1This approach has been investigated under the extreme value theory (EVT). In the context of the EVT, there are other approaches such as block-maxima, peak-over-threshold, and limiting the distribution of the threshold, e.g., generalized Pareto distribution. The generalized Pareto distribution is canonical in the application of the EVT. More information about the EVT can be found in McNeil et al. (2005).
2By using a flexible distribution, which captures all empirical regularities of data, one can overcome this problem. Fallahgoul et al. (2018b) provide an approach to mitigate the model risk and they apply it to an asset allocation problem.
reads as follows: let $X$ be a random variable with cumulative distribution function $F$ and $\bar{F}(x) = 1 - F(x)$ is its survival function, then $\forall \lambda > 0, \lim_{x \to +\infty} \bar{F}(x)e^{\lambda x} = +\infty$. Classic examples of heavy tail distributions are the Pareto, Student’s $t$, Frechet, stable and tempered stable distributions as well as the log-normal distribution. Distributions whose probability density function at infinity behaves like a power law are sometimes called fat-tail distributions. A heavy tail distribution might have moments of any order, while a fat-tail distribution will have infinite moments at some point.

In this paper, we consider here a broader definition of the heavy-tail property, by considering the distributions that exhibit large skewness and kurtosis compared to the normal distribution.

There are two main approaches for identifying the heavy tail property: (i) statistical tests (like the Kolmogorov-Smirnov test) and; (ii) graphical methods (like QQ-plots). Both approaches are based on the fact that a heavy tail variable exhibits more extreme values than a normally distributed random variable with the same location, i.e. average, and dispersion parameter, i.e., variance. Applied to financial time series, it means that large (positive or negative) price variations are observed more often than what would be expected if they were drawn from a normal distribution. Under the normality assumption, the distribution is entirely characterised by its first two moments, while taking into account the heavy tail property requires the knowledge of higher order moments, in particular the third moment i.e., skewness, which is a measure of the distribution’s asymmetry and the fourth moment, i.e., kurtosis which accounts for fatness of the tail.

The tempered stable distributions can be chosen to have both heavy tail property, i.e. high kurtosis, as well as asymmetric property, i.e. high skewness. They have been used in numerous financial and risk management applications. The variance gamma (see, Carr et al. (2003)) and normal inverse Gaussian (see, Carr et al. (2003)) distribution are special cases of tempered stable distributions. The useful properties of such distributions for application in financial and risk management are discussed in several papers (see, for example Fallahgoul et al. (2016), Fallahgoul et al. (2018c), Fallahgoul et al. (2017), and Kim et al. (2009), among others).

---

3This is an informal definition for the heavy tail property among several others. See McNeil et al. (2005).
The first step when using a tempered stable distribution is to estimate its parameters. This task can be done in three ways: parametric, non-parametric, or through simulation. In order to estimate the parameters of a distribution based on standard parametric methods we need either: (1) a closed-form formula for the probability density function (pdf) or cumulative distribution function (cdf) or (2) finite moments of some orders. When estimating the parameters of stable and tempered stable distributions, there are some difficulties. First a closed-form formula for the pdf and cdf of the stable and tempered stable distribution do not exist, deriving the maximum likelihood estimation (MLE) function is not an easy task. Moreover, at least for stable distributions, due to non-existence of moments at all orders, the general moments method (GMM), e.g. Hansen (1982), is not applicable in general. Some important papers addressing these problems are Fama and Roll (1968, 1971), McCulloch (1986) and DuMouchel (1973, 1983). More recently, an efficient method based on simulation for estimating the different classes of tempered stable distributions has been proposed in Fallahgoul et al. (2017).

In this paper, we first review the definitions and some theoretical properties of the stable and tempered stable distributions in Section 2. In Section 3, we explain how to simulate a heavy tail random variable, and also how to compute the pdf and cdf from the characteristic function. Finally, we actually carry out these computations on two empirical examples of financial times series (daily returns of the DAX and SPX index).

2 Stable and Tempered Stable Distributions

In this section we discuss the definition as well as some properties of stable and tempered stable random variables. Since none of these random variables has a closed-form formula for their pdf and cdf, we are only interested in the representation of their characteristic function. Furthermore, in order to calculate their moments, we discuss their cumulant generating function.
2.1 Lévy processes and the Lévy Khintchine formula

Both stable and tempered stable random variables belong to class of the Lévy processes. The characteristic function of a random variable $X$, that we denote $\Phi_X(u; X)$, is the Fourier transform of its distribution. Therefore, the related pdf or cdf can be recovered with only one Fourier inversion. The Lévy Khintchine formula says that characteristic function of any Lévy process $X = (X_t)_{t \geq 0}$, defined by

$$\Phi_X(u; X_t) = \mathbb{E}[e^{iuX_t}],$$

can be written under the following form:

$$\mathbb{E}[e^{iuX_t}] = \exp \left( t \left( aiu - \frac{1}{2} \sigma^2 u^2 + \int_{\mathbb{R}\setminus\{0\}} \left( e^{iux} - 1 - iux I_{|x|<1} \right) \nu(dx) \right) \right)$$

(1)

where $a \in \mathbb{R}$, $\sigma \geq 0$, $I$ is the indicator function and $\nu$ is a Lévy measure of $X$, satisfying the property

$$\int_{\mathbb{R}\setminus\{0\}} (1 \wedge x^2) \nu(dx) < \infty.$$

The measure $\nu$ describes the frequency of jumps of size $x$. A complete reference on Lévy processes can be found in Cont and Tankov (2003) or Rachev et al. (2011). A well know particular case is when $\nu$ vanishes, and we recover that $X$ follows a Brownian motion with drift $a$ and volatility $\sigma$.

2.2 Stable Distribution

Stable random variables were first introduced by Gnedenko and Kolmogorov (1954) while they were studying the sum of random variables. However, the first formal definition of a stable random variable was given by Feller (1971). There are several ways to define a

---

4There is a slight abuse of notation here, a random variable is a snapshot at a given time of a random process.

5Detailed information about this approach can be found in Fallahgoul et al. (2016) and Fallahgoul et al. (2017).
stable random variable through its characteristic function\(^6\).

Stable variables can first be defined via their Lévy measure. \(^7\) For any \(\alpha < 2\), the Lévy measure of a stable process is given by

\[
\nu(dx) = \left( \frac{C^+}{x^{1+\alpha}} 1_{x>0} + \frac{C^-}{x^{1+\alpha}} 1_{x<0} \right) dx,
\]

where \(1_{x>0}\) is an indicator function. The calculation of the characteristic function of the stable distribution is based on the Lévy-Khintchine formula. By replacing the Lévy measure in (1) by (2), one can recover the characteristic function of \(S\) is given by

\[
\Phi(u; S) = \mathbb{E}[e^{iuS}] = \begin{cases} 
\exp(i\mu u - |\sigma u|^\alpha(1 - i\beta \text{sign}(u) \tan(\frac{\pi\alpha}{2}))), & \alpha \neq 1 \\
\exp(i\mu u - |\sigma u|(1 + i\beta^2 \frac{2}{\pi} \text{sign}(u) \ln|u|)), & \alpha = 1,
\end{cases}
\]

where

\[
\text{sign } u = \begin{cases} 
1, & u > 0 \\
0, & u = 0 \\
-1, & u < 0,
\end{cases}
\]

\(0 < \alpha \leq 2\), \(\sigma \geq 0\), \(-1 \geq \beta \geq 1\), and \(\mu \in \mathbb{R}\). The parameter \(\alpha\) is the index of stability. It controls the behavior of the left and right tails. When \(\alpha\) is close to 2, the tail becomes thin.\(^8\)

When moving down from 2 to 1.5, a stable random variable exhibits a heavier tail. For large values of \(x\) the pdf of a stable variable behaves like a power law (sometimes called Pareto tail)

\[
f(x) \simeq \frac{C_{\pm}^{\alpha, \beta}}{|x|^{1+\alpha}}
\]

where \(C_{\pm}^{\alpha, \beta}\) are the parameters for large positive or negative values. In empirical appli-

\(^6\)It should be noted that some random variables are a special case of a stable random variable such as Lévy and Gamma random variable. And they have a closed-formula for their pdf and cdf.

\(^7\)the arrival rate of jumps of size \(x \in \mathbb{R}\setminus\{0\}\)

\(^8\)A Gaussian random variable is a special case of stable variable, when \(\alpha = 2\) and \(\beta = 0\).
cations, the estimated or calibrated value of $\alpha$ is usually between 1.3 and 1.9. The $\beta$, $\sigma$, and $\mu$ parameters are respectively the skewness, scale and location parameters. When a random variable $S$ follows the stable distribution, we denote it by $S_\alpha(\sigma, \beta, \mu)$.\(^9\)

Special cases of stable distribution where an analytic expression exists for the pdf are the Cauchy distribution ($\alpha = 1$):

$$f(x) = \frac{C}{x^2 + \pi^2C^2}$$

and the Lévy distribution ($\alpha = 1/2$)

$$\sqrt{\frac{c}{2\pi x}} e^{-\frac{c^2}{x}}.$$

Detailed informations about the characteristic function of the stable distribution can be found in Samorodnitsky and Taqqu (1994).

Historically, the initial definition of a stable random variable was based on sums of random variables. A random variable $X$ is said to have a stable distribution if for any $n \geq 2$, there is a positive number $C_n$ and real number $D_n$ such that

$$X_1 + X_2 + \cdots + X_n = C_nX + D_n,$$

where $X_1, X_2, \cdots, X_n$ are independent copies of $X$.\(^{10}\) One can show that $C_n = n^{1/\alpha}$, where $\alpha \in (0, 2]$, in other words, there is a generalised central limit theorem (CLT) which holds for stable variables: a sum of i.i.d variables with paretian tails behaving like $\frac{1}{|x|^{1+\alpha}}$ normalised by $n^{1/\alpha}$ will converge to a stable distribution (one recognises the usual CLT when $\alpha = 2$).

---

\(^9\)Some people call a stable random variable $\alpha$—stable. And also, there are different parameterizations for a stable random variable. Detailed information about all parametrization of a stable random variable can be found in Nolan (2003) and Rachev et al. (2011).

\(^{10}\)For more information about the connection of the different definitions of a stable random variable see Chapter 1 of Samorodnitsky and Taqqu (1994).
2.3 Variation and Existence of Moments

Lévy processes are said to have finite activity if

\[ \int_{\mathbb{R}\setminus\{0\}} v(dx) < \infty \]

and infinite activity otherwise. Stable process always have infinite activity, which means there is an infinite number of jumps within any finite time interval (except in the normal case where there are no jumps at all).

The sample paths of Lévy process have finite variation when

\[ \int_{\mathbb{R}\setminus\{0\}} (1 \wedge |x|)v(dx) < \infty, \]

and infinite variation otherwise.\(^{11}\) Therefore, by (2), a stable process has finite variation if and only if \( \alpha < 1 \) (Note that the Brownian motion, i.e. \( \alpha = 2 \), does not have finite variation).

Furthermore, the second moment of a stable process do not exist (except for \( \alpha = 2 \)), and actually the moment of order \( p \) for \( p \geq \alpha \) do not exist. To overcome to this problem, one will use a tempering function, hereafter \( t(x) \), to ensure that the second or higher moments do exist. The new process is then called a tempered stable process. The Lévy measure of the tempered stable process is given by \( \nu_{ts}(dx) = t(x)\nu_s(dx) \), where \( \nu_s(dx) \) is the Lévy measure of the stable process. With different choices of the tempering function come different classes of tempered stable process.\(^{12}\)

2.4 Tempered Stable Distributions

In this section, we discuss the properties of different tempered stable distributions.\(^{13}\) The characteristic function and Lévy measures of all classes of tempered stable random variables are summarised in Appendix A.

---

\(^{11}\)Loosely speaking, a process has finite variation when the length of its sample paths is locally finite almost surely

\(^{12}\)Detailed information about the tempered stable process and their properties can be found in Rachev et al. (2011), Fallahgoul et al. (2016), Fallahgoul et al. (2017), and the references therein.

\(^{13}\)We refer readers to Rachev et al. (2011) for detailed information about these classes of distributions and processes.
2.4.1 Classical Tempered Stable

A classical tempered stable (CTS) random variable, also known as a CGMY random variable after the work of Carr et al. (2003) has a tempering function given by

\[
t(x) = \begin{cases} 
  e^{-\lambda_+ x}, & x > 0 \\
  e^{\lambda_- x}, & x < 0
\end{cases}
\]

where \( \lambda_+ \) and \( \lambda_- \) are non-negative and are the tempering parameter for right and left tails.\(^\text{14}\) Therefore, the Lévy measure of a CTS random variable is given by

\[
v(dx) = \left( C e^{-\lambda_+ x} 1_{x > 0} + C e^{\lambda_- x} 1_{x < 0} \right) dx.
\]

The role of parameter \( \alpha \) is the same as for a stable random variable, which controls the behavior of the left and right tails. The parameter \( C \) is the scale parameter which controls the kurtosis of the distribution. The parameters \( \lambda_+ \) and \( \lambda_- \) control the rate of decay on the positive (right) and negative (left) tails. By contraction when \( \lambda_+ = \lambda_- \), a CTS random variable is a symmetric random variable. We denote a CTS random variable by \( X_{\text{CTS}} \).

I think we should move this to an appendix, as this is general knowledge.

Tempering the tail of a stable random variable allows the existence of higher moments. Moments of a random variable can be obtained from its cumulant generating function (CGF). Let \( \Phi(u, X) \) be as in (2.1), we let \( \psi(u, X_{\text{CTS}}) = \log \Phi(u, X_{\text{CTS}}) \), then a cumulant of order \( n \) is given by

\[
C_m(X_{\text{CTS}}) = \frac{1}{i^n} \frac{\partial^n \psi(u, X_{\text{CTS}})}{\partial u^n} \quad n = 1, 2, \ldots
\]

where \( \psi \) is the CGF. The cumulants for a CTS random variable are equal to

\[
c_1(X_{\text{CTS}}) = 0 \\
c_n(X_{\text{CTS}}) = \Gamma(n - \alpha) \left( \lambda_+^{\alpha-n} + (-1)^n \lambda_-^{\alpha-n} \right), \quad n = 2, 3, \ldots
\]

\(^\text{14}\)In Carr et al. (2003), they represent the tail parameter by \( Y \) instead of \( \alpha \), and also, the \( \lambda_+ \) and \( \lambda_- \) are given by \( G \) and \( M \), respectively.
where $\Gamma(n - \alpha) = (n - \alpha - 1)!$.

There is a one-to-one relation between the CGF and a moment generating function (MGF). The moment generating function of a random variable $X$ is given by

$$M(u) = \mathbb{E} \left[ e^{uX} \right].$$

The CGF is the logarithm of the MGF, i.e. $\psi(u) = \log(M(u))$. The first cumulant is the mean. The second and third cumulants are the second and third central moments, respectively. However, the higher cumulants are neither moments nor central moments.

### 2.4.2 Generalized Tempered Stable

A CTS random variable has the same scale parameter, i.e., $C$, for the left and right tail. An extension of the CTS random variable can be obtained by allowing more flexibility to the scale parameter. The new random variable is called generalized CTS (GCTS) random variable. Several well-known tempered stable random variables, such as CGMY (Carr et al. (2003)), KoBoL (Boyarchenko and Levendorskiı̆ (2002)), and Lévy flight, are special case of a GCTS random variable. We denote a GCTS random variable by $X_{\text{GCTS}}$.

The cumulants of $X_{\text{GCTS}}$ are given by

$$c_1(X_{\text{GCTS}}) = 0$$

$$c_n(X_{\text{GCTS}}) = C_+ \Gamma(n - \alpha_+) \lambda_+^{\alpha_+ - n} + (-1)^n C_- \Gamma(n - \alpha_-) \lambda_-^{\alpha_- - n}, \quad n = 1, 2, \ldots.$$  

The tempering functions of the next two tempered stable random variables are based on some hypergeometric functions.\(^{15}\) This choice allows the finiteness of not only higher order moments, but of some exponential moments as well, which it might be useful for in some application to option pricing.

\(^{15}\)Detailed information about a hypergeometric function can be found in Andrews and Andrews (1992).
2.4.3 Modified Tempered Stable

A modified tempered stable (MTS) random variable is obtained by taking the Lévy measure of a symmetric stable random variable and multiplying it by the following tempering function

\[ t(x) = (\lambda |x|)^{\frac{\alpha+1}{2}} K_{\frac{\alpha+1}{2}}(\lambda |x|) \]

on each half of the real axis, where \( K_p(x) \) is the modified Bessel function of the second kind (see Appendix B).\(^{16}\)

The Lévy measure of a MTS random variable is given by

\[
\nu(dx) = C \left( \frac{\lambda_+^{\frac{\alpha+1}{2}} K_{\frac{\alpha+1}{2}}(\lambda_+ x)}{x^{\frac{\alpha+1}{2}}} 1_{x>0} + \frac{\lambda_-^{\frac{\alpha+1}{2}} K_{\frac{\alpha+1}{2}}(\lambda_- x)}{x^{\frac{\alpha+1}{2}}} 1_{x<0} \right) dx
\]

where \( C, \lambda_+, \lambda_- > 0 \), and \( \alpha \in (0,1) \cup (1,2) \). We denote a MTS random variable via \( X_{MTS} \).

The cumulants of \( X_{MTS} \) are equal to

\[
c_n(X_{MTS}) = 2^{n-\frac{\alpha+3}{2}} C \Gamma \left( \frac{n+1}{2} \right) \Gamma \left( \frac{n-\alpha}{2} \right) \left( \lambda_+^{\alpha-n} + (-1)^n \lambda_-^{\alpha-n} \right)
\]

for \( n = 2, 3, 4, \ldots \).

There are two important points about the MTS random variable. Firstly, the reason for using the modified Bessel function of the second kind, i.e., \( K_p(x) \), as a tempering function, is related to the characteristic function of the MTS random variable. \( K_p(x) \) allows to have a closed-form formula for the characteristic function. Secondly, because of complicated form of the tempering function, ex-ante it is not obvious what is the asymptotic behavior for the Lévy measure of a MTS random variable. The asymptotic behavior for the Lévy measure of the MTS random variable is explored in Kim et al. (2008). For \( \alpha \in (0,2) \setminus \{1\} \), the Lévy measure of the stable and MTS, CTS, and GCTS random variable have the same asymptotic behavior at the zero neighborhood, while, the tails of the MTS

\(^{16}\)Detailed information about the Bessel function can be found in Andrews and Andrews (1992). This random variable was introduced by Kim et al. (2008).
random variable are thinner (heavier) than those of the stable (CTS and/or GCTS) random variable.

2.4.4 Rapidly Decreasing Tempered Stable

Rapidly decreasing tempered stable (RDTS) random variables were introduced by Kim et al. (2010b). The tempering function of the RDTS is given by

$$t(x) = e^{-\frac{\lambda_+^2 x^2}{2}1_{x>0}} + e^{-\frac{\lambda_-^2 x^2}{2}1_{|x|<0}}$$

where $\lambda_+, \lambda_- > 0$, and $\alpha \in (0, 1) \cup (1, 2)$. Therefore, the Lévy measure of the RDTS is given by

$$\nu(dx) = \left(C e^{-\frac{\lambda_+^2 x^2}{2}1_{x>0}} + e^{-\frac{\lambda_-^2 x^2}{2}1_{|x|<0}}\right) dx$$

where $C > 0$. We denote a RDTS random variable via $X_{\text{RDTS}}$.

2.5 Tempering with a subordinator

The next two tempered stable random variables are obtained by replacing the physical time of the Brownian motion with a subordinator.\(^\dagger\) Constructing a tempered stable random variable in this way has several advantages. Firstly, by choosing the appropriate subordinator, one can control the rate of decay of tails. Secondly, the extension from a univariate tempered stable random variable to a multivariate one is straightforward, see Fallahgoul et al. (2016) and Fallahgoul et al. (2018c).

2.5.1 Normal Tempered Stable

A normal tempered stable (NTS) random variable is obtained by replacing the physical time, i.e., $t$, by the CTS subordinator. The CTS subordinator is a Lévy process whose Lévy measure is the positive part of the Lévy measure of a CTS random variable: If $T_t$ is

\(^\dagger\)Detailed information about a subordinator can be found in Cont and Tankov (2003), Fallahgoul et al. (2016), and Fallahgoul et al. (2017).
the CTS subordinator, then the Lévy measure of $T_t$ is given by

$$v(dx) = \frac{Ce^{-\theta x}}{x^{\frac{\alpha}{2}+1}}1_{x>0}dx$$

where $C, \theta > 0$, and $0 < \alpha < 2$. The unconditional expectation of $T_t$ should be equal to $t$, i.e. $\mathbb{E}[T_t] = t$. To achieve it, by setting $C = \frac{1}{\Gamma(1-\frac{\alpha}{2})\theta^{\frac{\alpha}{2}-1}}$, we have

$$\mathbb{E}[T_t] = tC\Gamma \left( 1 - \frac{\alpha}{2} \right) \theta^{\frac{\alpha}{2}-1} = t.$$  

The characteristic function of the subordinator $T_t$ is given by

$$\Phi(u; T_t) = \exp \left( \left( -\frac{2\theta^{1-\frac{\alpha}{2}}}{\alpha}((\theta - iu)\frac{\alpha}{2} - \theta^{\frac{\alpha}{2}}) \right) t \right).$$

Let $\mu, \beta \in \mathbb{R}$, $\sigma > 0$, $B_t$ be Brownian motion, and $T_t$ be CTS subordinator. Then the NTS process is defined by

$$X_t = \mu t + \beta(T_t - t) + \sigma B_{T_t}.$$  

Given that $T_t$ and $B_t$ are independent, the characteristic function of $X_t$ is given by

$$\Phi(u; X_t) = \exp \left( iu(\mu - \beta)t - \frac{2\theta^{1-\frac{\alpha}{2}}}{\alpha}((\theta - i\beta u + \frac{\sigma^2 u^2}{2})\frac{\alpha}{2} - \theta^{\frac{\alpha}{2}}) t \right).$$

By using the cumulant generating function, the first two moments of $X_t$ are equal to

$$\mathbb{E}[X_t] = \mu t$$

$$\mathbb{V}ar[X_t] = \sigma^2 t + \beta^2 \left( \frac{2 - \alpha}{2\theta} \right) t.$$  

\[18\]So far we have discussed the unconditional tempered random variables. More precisely, we only have assumed the physical time is one.
2.5.2 Exponential Tilting Stable

An exponential tilting stable (ETS) process is obtained by replacing the physical time of a Brownian motion with an exponential tilting (ET) subordinator whose Levy measure is given below in equation (4). The way we construct the ETS process is similar to the way Barndorff-Nielsen and Shephard (2001) use the classical tempered stable subordinator to construct the NTS process, see, Barndorff-Nielsen and Shephard (2001) and Kim et al. (2008), among others. The ETS process, i.e., $X_t$, is given by

$$X_t = \mu t + \beta (s_t - t) + \gamma B_s$$

where $\mu, \beta \in \mathbb{R}$, $\gamma \in \mathbb{R}^+$, $B_s$ is the Brownian motion, and $s_t$ is the ET subordinator and independent of $B_t$. By rearranging (3), we can use the properties of the arithmetic Brownian motion

$$X_t = (\mu - \beta) t + \beta s_t + B_s$$

$$= (\mu - \beta) t + Y_t,$$

where the $(Y_t)_{t \geq 0}$ is a subordinated arithmetic Bm with drift $\beta$ and volatility $\gamma$. It should be noted that the characteristic exponent for the arithmetic Bm is $\varphi_{B_t}(u) = i\beta u - \frac{1}{2} \sigma^2 u^2$.

The Lévy measure of an ET subordinator is given by

$$v(dx) = \left( e^{-\frac{\lambda^2 x^2}{2 \alpha}} \frac{1}{\Gamma(\alpha+1)} \right) dx,$$

where $\alpha \in (0,1)$ and $\lambda > 0$.$^{19}$ The characteristic function of the ET subordinator $s_t$ is equal to

$$\Phi(u; s_t) = \exp \left[ -\frac{\lambda}{2} \frac{1}{\Gamma\left(\frac{1-\alpha}{2}\right)} G \left( \alpha, \lambda; \frac{iu}{\lambda} \right) t \right].$$

$^{19}$This Lévy measure is the positive side of the Lévy measure of a rapidly decreasing tempered stable random variable.
The function $G$ is based on the confluent hyper-geometric function.\(^{20}\) By using the characteristic function of $s_t$, one can derive all cumulants for the ET subordinator. They are given by

$$c_n(s_t) = \frac{2^{n-1} \Gamma \left(\frac{n-\alpha}{2}\right) t}{\lambda^{n-1} \Gamma \left(\frac{1-\alpha}{2}\right)}$$

where $n = 1, 2, 3, \ldots$.

Let $(B_t)_{t \geq 0}$ be the Bm on $\mathbb{R}$ with the characteristic exponent $i\beta u - \frac{1}{2} \sigma^2 u^2$ and let $s_t$ be the ET subordinator with the parameters $(\alpha, \lambda)$. Then the characteristic function for the ETS process $X_t$, i.e., equation (3), is

$$\Phi(u; X_t) = \exp \left[ \frac{\lambda}{2\Gamma \left(\frac{1-\alpha}{2}\right)} G \left(\alpha, \lambda; \frac{i\beta u - \frac{1}{2} \sigma^2 u^2}{\lambda}\right) t + iu(\mu - \beta) t \right].$$

The mean and variance for the stochastic process $X_t$ are equal to

$$E[X_t] = \mu t$$

$$Var[X_t] = \gamma^2 t + \beta^2 \frac{2}{\lambda \Gamma \left(\frac{1-\alpha}{2}\right)} \Gamma \left(\frac{2-\alpha}{2}\right) t.$$

**2.6 Non-linear dependency**

Dependency between the returns is an important factor in the financial time series analysis. Based on the dependency between returns, which can be linear or nonlinear, one can decide what kind of model can be used for modelling. If the dependency is linear, the correlation coefficient describes precisely the movements of return series, however, it is not always a good measure of dependency. Modelling dependency among returns is more important during a financial crisis, because of the excess correlation that is observed during those periods.

Based on the linear dependency property, a multivariate normal distribution is a good candidate for modeling multivariate financial time series. To do so, the major assump-

\(^{20}\)Detailed information about $G$ can be found in the Online Appendix of Fallahgoul et al. (2018c).
tion is that the dependency is linear and can be explained by variance-covariance matrix. But, a large set of literature have rejected this claim, see McNeil et al. (2005) and Rachev et al. (2011) among others. In other words, let denote the returns of two assets by $X$ and $Y$, which there is dependency between them. If the dependency is linear, then $E[X|Y]$ can be expressed as a linear function of $Y$. But, if the conditional expectation is not a linear function, a linear dependency measure such as Pearson’s correlation coefficient $\rho$ is not an adequate tool for capturing the dependency structure between these two random variables. To overcome this problem, one may use the copula technique.\textsuperscript{21} It should be noted that one can show the existence of the nonlinear dependency by using exceedance correlations methods, too.\textsuperscript{22}

Tail dependency coefficient, as a measure of non-linear dependency, is designed to capture the dependency in the tails. Tempered stable random variables have been used as a flexible class of models for calculating this coefficient.\textsuperscript{23}

3 Evaluating pdf and cdf Based on a Characteristic Function

3.1 Theoretical Setting

Let $X$ be a random variable, then its characteristic function, i.e., $E[e^{iuX}]$, is the Fourier transform of its pdf. Consequently, by obtaining the inverse of the related Fourier transform one can recover the pdf. This approach was introduced by DuMouchel (1975) for calculating the pdf of the stable distribution. Later this approach was applied to some classes of the tempered stable distribution by Kim et al. (2009), Fallahgoul et al. (2016), and Fallahgoul et al. (2018c).

\textsuperscript{21}See McNeil et al. (2005) and Fallahgoul et al. (2016).

\textsuperscript{22}See Longin and Solnik (2001).

\textsuperscript{23}See, McNeil et al. (2005), Fallahgoul et al. (2016), and references therein.
3.1.1 PDF

Let $X$ be a random variable for a class of the tempered stable distribution, then its characteristic function is given by

$$
\phi_X(u) = E[e^{iuX}] = \int_{-\infty}^{\infty} e^{iuX} f_X(X) dX
$$

where $f_X(X)$ is the pdf for the random variable $X$. Consequently, the pdf for the random variable $X$ is given by

$$
f_X(X) = \int_{-\infty}^{\infty} e^{-iuX} \phi_X(u) dX \tag{5}
$$

where $\phi_X(u)$ is the characteristic function. Therefore, by calculating equation (5) one can obtain the pdf for the random variable $X$. The characteristic function of the tempered stable distributions are complicated and an analytic formula for equation (5) does not exist. Based on numerical integration and fast Fourier transform (FFT), numerical approaches are suggested.\(^{24}\)

The first step in numerical approximation for equation (5) is the discrete Fourier transform (DFT), which involves transforming a vector $Y = (y_1, y_2, \cdots, y_n)$ to a vector $X = (x_1, x_2, \cdots, x_n)$ through the DFT

$$
x_j = \sum_{k=1}^{n} y_k e^{\frac{2\pi j (k-1)}{n}}, \quad j = 1, 2, \cdots, n. \tag{6}
$$

The computational cost of equation (6) is high, and the FFT provides a popular way of computing this more efficiently.

Let $a \in \mathbb{R}^+, q \in \mathbb{N}^+$ and for $j, k \in \{1, 2, \cdots, n = 2^q\}$ define

$$
u_k = -a + \frac{2a}{n} (k-1), \quad \nu_k^* = \frac{u_{k+1} + u_k}{2}, \quad x_j = -\frac{n\pi}{2a} + \frac{\pi}{a} (j-1), \quad C_j = \frac{a}{n\pi} (-1)^{j-1} e^{\frac{\pi (j-1)}{n}}.
$$

An approximation for equation (5) at $x_j$ is given by

$$
f_X(x_j) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iuXj} \phi_X(u) du
$$

\(^{24}\)See, for example, Bailey and Swarztrauber (1994), among others.
\begin{align*}
\sim \frac{1}{2\pi} \int_{-a}^{a} e^{-iux} \phi_X(u) du \\
\sim C_j \sum_{i=1}^{n} (-1)^{k-1} \phi_X(u_k^*) e^{-i 2\pi (j-1)(k-1)/n},
\end{align*}

where parameters \( a \) and \( q \) are the limit of integration for the Fourier transform.\(^{25}\)

### 3.1.2 CDF

Evaluating the cdf of the tempered stable distributions follows the same line as calculating the pdf, but with some differences. The differences between computing the cdf and pdf are related to the form of characteristic function. In this paper, we follow the approach of Kim et al. (2010a) where they show that the cdf of a tempered stable distribution is given by

\[
F_X(x) = \frac{e^{x\rho}}{\pi} \Re \left( \int_{0}^{\infty} e^{-iux} \phi_X(u + i\rho) / c(\rho - iu) du \right), \quad x \in \mathbb{R},
\]

where \( \Re \) and \( \Im \) are the real and imaginary part of a complex number, and for all \( Z \), \( |\phi_X(Z)| < \infty \) and \( \Im(Z) = \rho > 0.\(^{26}\) Let \( g_X(u) = \phi_X(u + i\rho) / \rho - iu \). In order to apply the DFT and FFT, let \( a \in \mathbb{R}^+, q \in \mathbb{N}^+ \) and for \( j, k \in \{1, 2, \cdots, n = 2^q\} \) define

\[
\begin{align*}
  u_k &= \frac{2a}{n} (k - 1), \\
  u_k^* &= \frac{u_k + u_{k+1}}{2} = \frac{a}{n} (2k - 1), \quad \text{and} \quad x_j = -\frac{n\pi}{2a} + \frac{\pi}{a} (j - 1),
\end{align*}
\]

and then,

\[
\int_{0}^{\infty} e^{-ix_j u} g_X(u) du \sim \int_{0}^{a} e^{-ix_j u} g_X(u) du, \\
\sim \sum_{k=1}^{n} e^{-ix_j u_k^*} g_X(u_k^*) \frac{2a}{n}.
\]

By simple computation, it can be shown that for each \( x_j \),

\[
F_X(x_j) \sim \frac{e^{x_j \rho}}{\pi} \Re \left\{ \frac{2a}{n} e^{-i\pi (j-1)/n} \sum_{k=1}^{n} e^{i\pi (2k-1)/n} g_X(u_k^*) e^{-i 2\pi (k-1)(j-1)/n} \right\}.
\]

\(^{25}\)More details about the accuracy of this approach for the stable distribution can be found in Menn and Rachev (2006).

\(^{26}\)For detailed information about deriving (7) see Preposition 1 of Kim et al. (2010a).
It should be noted that the accuracy of this approximation depends on the numerical integration method used.\(^{27}\)

### 3.1.3 Generating a Tempered Stable Random Variable

After calculating the cdf of a random variable, generating a random variable is straightforward. Let \(X\) be a random variable and \(F\) its cdf, let \(U\) be a uniformly distributed random variable on \([0, 1]\), then \(F^{-1}(U)\) returns a random variable with the same law as \(X\). Consequently, in order to generate a tempered stable random variable, by using the inversion of the cdf, one needs to generate a vector of uniform independent random variables and to calculate the inverse of the cdf at these values. The computation of the inverse of the cdf can be done by applying any interpolation method. Therefore, generating a tempered stable random variable can be done through the following steps:

- Given a set of \(y_i\) and \(x_i\), compute the cdf at points \(x_i\) using the characteristic function at points \(y_i\);
- Generate a vector of uniform distributions;
- Compute the inverse of the cdf by interpolation at these points.

### 3.2 Numerical Simulations

Here we generate pdf’s and cdf’s of the tempered stable variables mentioned previously, based on their characteristic function, as explained in the previous section, and examine the effect of some of the parameters. A graphical method to compare two random variables is the so-called qq-plot.\(^{28}\)

#### 3.2.1 Comparing the Classical and Exponentially Tilting (ET) subordinator

Figure 1 shows the qq-plot for the classical and ET stable subordinator for a fixed value of \(\alpha\) and different values of \(\lambda\). They seem to have almost the same quantiles except on the

\(^{27}\)One can increase the accuracy and speed of this approximation by using different numerical integration methods. See Menn and Rachev (2006), Kim et al. (2010a) and Fallahgoul et al. (2018c).

\(^{28}\)Each plot has two coordinates that are given by a given quantile of both distributions, see, McNeil et al. (2005).
tails. By increasing the value of $\lambda$ the classical tempered stable subordinator has larger quantiles on the tails compared to the ET stable subordinator. This property is expected, since the tilting function of the ET stable subordinator puts more weights on the tails in comparison to the classical tempered stable subordinator.

Figure 5 shows the pdf function for the ET and classical tempered stable subordinator. A visualization for the pdf function reveals that the pdf function for the ET distribution decays faster than the classical tempered stable subordinator. In other words, the tail for the classical tempered stable is heavier than the tail for the ET subordinator, which is the consist with our expectation.

3.2.2 Behaviour of the ET Subordinator with respect to parameters

In order to illustrate the theoretical results and understand the role of each parameter for the ET stable subordinator, we present some plots of its pdf and cdf. Figures 2-4 show the pdf and cdf plots for the ET stable subordinator. The effect of parameter $\alpha$ for fixed values of $\lambda$ and $t$ is presented in Figure 2a, which confirms that by increasing its value from 1.1 to 1.9, a heavier tail is obtained. The parameter $\alpha$ controls the heavy-tail property.

Figure 2b shows the effect of tempering parameter, i.e., $\lambda$. Increasing its value from 0.1 to 1 creates heavier tails and it skews the pdf to the right. Consequently, the parameter $\lambda$ has effect on both the decay of the tail and the skewness.

Figure 3 shows the effect of time parameter, i.e., $t$, on the pdf and cdf functions. As expected, by increasing the value of parameter $t$ the mass of the pdf function is transferred to the positive side and flattens the pdf. Finally, Figure 4 shows the effect of $\alpha$ and $\lambda$ on the cdf.

3.2.3 Study of the Univariate ET stable distribution

In order to visualize the role of parameters on the pdf and cdf function for the ET stable distribution, we provide some plots by using the FFT algorithm.

Figures 6-8 show the effect of each parameter on the pdf function. Figure 6a presents the effect of tail parameter, i.e., $\alpha$, on the pdf function, which indicates that the tails decay slower by increasing the value of $\alpha$. Figure 6b shows the effect of $\gamma$ on the pdf function,
which controls the dispersion/ kurtosis. By increasing the values of $\gamma$, the pdf has a smaller mode and heavier tails. Figure 7a presents the effect of the tempering parameter, i.e., $\lambda$, which has the same role as in the ET subordinator: it controls both the tail decay and the skewness of the distribution.

The effect of parameters $\beta$ and $\mu$ are presented in Figures 7b and 8a. It reveals that the $\beta$ parameter affects the skewness while the $\mu$ parameter changes the location of the distribution. Finally, the role of time parameter, i.e., $t$, is presented in Figure 8b. It reveals that by increasing the values of $t$, the pdf has a smaller mode and heavier tails.

4 Empirical Analysis

In this section, we study the performance of different types of distributions to capture the empirical regularities of the equity returns. For our purpose, we will use 18 years of daily prices for two major equity market indexes: (1) S&P500, and; (2) DAX. Figure 9 shows the daily price processes of S&P500 and DAX indexes.

4.1 GJR-GARCH

4.1.1 Conditional vs. Unconditional Data

It has been shown that financial asset returns share common features such as heavy tail, stochastic volatility, nonlinear dependency, leverage effect, and skewness. The leverage effect is the correlation between returns and change of volatilities, which is likely to be negative or close to zero for equities. However, stochastic volatility can also be referred to as volatility clustering: the standard deviation of the returns (or any other measure of their dispersion) follows a time evolving process, which needs to be filtered out to study the unconditional data.\footnote{By unconditional we mean all observations are independent and identical distributed. Recently, Fallahgoul et al. (2018a) have introduced a novel approach for filtering out the latent variable from a proxy of it.}

There are several ways to show that a historical time series of returns has a volatility clustering property: (i) visual inspection on autocorrelation function plot, and; (ii) statistical tests such as Ljung-Box and Engle Ljung-Box.\footnote{See McNeil et al. (2005) and Rachev et al. (2011).} Stochastic volatility and leverage effect induce a non-stationarity of the return distribution, i.e., at a given time the return dis-
tribution is conditional to the past events. To capture all of these "empirical regularities" one has to fit a suitable transitional density. There is a long literature about this topic from the discrete- and continuous-time point of views.\footnote{Detailed information about these works can be find in Bates (2012), Fallahgoul et al. (2016), Fallahgoul et al. (2017), Fallahgoul et al. (2018a), and reference therein.} To do so, we first filter the data with a discrete-time model such as the GARCH process, (see, Bollerslev (1986) and Glosten et al. (1993)), to remove as much as possible the path-dependent effects. We use here the Glosten-Jagannathan-Runkle GARCH (GJR-GARCH) which was introduced by Glosten et al. (1993).

Let $r_t$ be a time series of an asset returns. Then $r_t = \mu + \epsilon_t$, where $\mu \in \mathbb{R}$ is the mean of returns and $\epsilon_t$ is a zero-mean white noise. Empirically the value of $\mu$ for low/high frequency data, i.e., such as daily, is very small (almost zero). We say that $r_t - \mu$ is a GJR-GARCH(1,1) process if $\epsilon_t = r_t - \mu = \sigma_t z_t$, where $z_t$ is a standard Gaussian distribution and $\sigma_t$ follows

$$\sigma_t^2 = \omega + (\alpha + \gamma I_{t-1})\epsilon_{t-1} + \beta \sigma_{t-1}^2$$

where

$$I_{t-1} = \begin{cases} 0, & r_{t-1} \geq \mu \\ 1, & r_{t-1} < \mu. \end{cases}$$

To ensure that $\sigma_t$ remains positive, we restrict to non-negative values of $\omega, \alpha, \gamma, \beta$. Note that the usual GARCH process is a special case of GJR-GARCH when $\gamma = 0$, (see, Bollerslev (1986)).

### 4.2 Filtering the Data

Figure 10 shows the daily log-return process of the S&P500 and DAX indexes. A quick visual inspection show that stochastic volatility and volatility clustering are very likely to happen. It is known that (unconditional) heavy tails generated by (conditional) volatility clustering can be mistakenly interpreted as evidence in favour of fat tailed distributions. To safeguard against conditional volatility (and possible mean reversion) we filter each
return series with a GJR-GARCH model such that the remaining tail dependencies are not conditional. To eliminate the volatility clustering feature, one can fit a GJR-GARCH model. Since the residuals of the GJR-GARCH process are independent identical distributed, we can use the tempered stable distributions as an unconditional distribution of these residuals.

Figure 11 shows the filtered daily log-returns for S&P500 and DAX indexes. A visual comparison between Figure 10 and Figure 11 reveals that the filtered data does not have volatility clustering. They can be seen as independent identically distributed random variables.

4.3 Empirical Evidence

To estimate the parameters of the stable and tempered stable distributions, we use the MLE method. To do so, we first recover the related pdfs by using the fast Fourier transform (FFT), as we explained in Section 3. Second, the log-likelihood function is constructed via the recovered pdfs from their characteristic functions. Finally, the parameters of the model of interest are obtained by maximizing the log-likelihood function over unknown parameters.

We first compute the skewness and kurtosis of both the raw and filtered returns. If a dataset follows the normal distribution, due to the symmetric property of the normal distribution, it should have zero skewness parameter. The value of skewness for raw (filtered) daily returns of S&P500 and DAX are $-0.2082 (-0.4847)$ and $-0.0531 (-0.3107)$, respectively, which means negative skewness, a strong evidence against normality. On the other hand, high kurtosis for return series usually means that it has heavy tail property. And also, excess kurtosis is defined as kurtosis above 3. A kurtosis significantly higher 3 is a strong evidence against normality. The value of kurtosis for raw (filtered) daily returns of S&P500 and DAX are 11.6076 (4.7335) and 7.3591 (3.9268), respectively, which shows the kurtosis parameter for both indexes of raw and filtered data, is higher than 3 (normality).

There are also statistical tests based on the skewness and kurtosis for testing the normality assumption. Two of these tests are the Jarque-Bera (JB) and Kolmogorov-Smirnov
(KS) tests, the later being based on a minimum distance estimation comparing a sample with a given probability distribution. The KS test makes no assumption about the data distribution, it is more flexible and its implementation is easy.

The KS statistics for raw (filtered) daily returns of S&P500 and DAX are 0.4611 (0.6880) and 0.4900 (0.6718), respectively, which implies that all of them reject the null hypothesis at level 5%. For implementing the KS test given normality assumption, we estimate the parameters $\mu$ and $\sigma$ from the observations, which leads to some inaccuracy. There is a modified version of the KS test, named Lilliefors (hereafter Lill) test, that follow the same procedure as KS test but with a correction to give a more accurate approximation of the distribution of the test statistics. The results of the Lill test are in the same line as the KS test.

We close this section via some graphical tests. There are a number of graphical methods that can be used to detect the heavy tail property from data. Graphical methods do not provide a precise statistical inference, but they can show the nature of deviation (if any) of data and they can be used as a prototype.

Quantile-Quantile plot (hereafter qq-plot) is a commonly graphical tool for identifying the heavy tail property. Via a qq-plot, one can see that a sample data is similar to a specific distribution or whether two datasets have similar distributions. The idea in identifying the heavy tail property via qq-plot is to compare the sample data with some known heavy tail distribution such as Student’s $t$, Stable, tempered stable, and etc.

Figure 12 shows the qq-plot of S&P500 and DAX returns against the fitted normal distribution. The x-axis shows the quantiles of fitted normal distribution while the y-axis represents the empirical quantiles of the sample. If the quantiles of sample data and fitted distribution follow the same probability law, the blue points should map on the straight line. It is clear that on the downside and upside the observation is deviate from normality. At least a visualization reveals that the behavior of tails of both samples are totally different. By using a heavy tail model such as Student’s $t$ distribution instead of normal distribution, we reach the better results, which its qq-plot reveal of a better

\[ N(\mu, \sigma^2) \]

We obtained the same inference via the JB test.

\[ N(\mu, \sigma^2) \]
consistency of the Student’s $t$ distribution with the returns. The bottom panel of Figure 12 demonstrates this fact.\textsuperscript{34}

Figures 13 and 15 show the pdf and log-pdf of the different models to the filtered log-returns. It appears that the kurtosis of the sample data is much larger than the normal distribution. This can be seen by looking at the peak (or mood) of distribution. Next, the decay of the tails of the normal distribution is much faster than the stable and tempered stable distribution.

Both, stable and tempered stable, have high kurtosis and heavy tails. An closer visual inspection of the tails confirms the theoretical fact that tempered stable distributions have a thinner tail than stable distributions, and heavier tails than a normal distribution. Overall, a comparison among all distributions and empirical distribution shows that the tempered stable distribution has a better performance in fitting the dynamic of the filtered sample returns.

4.4 Which Tempered Stable Distribution should you choose?

As we discussed in Section 4.3, there are statistical tests such as KS and Lill test for comparing the fitting performance.\textsuperscript{35} All statistical tests are in favor of tempered stable distributions. There are strong evidences, such as large test statistics and low p-values (almost zero), against normality assumption.

The problem of selecting one tempered stable model over the rest of others can be seen as a special case of a general problem. In fact, when there are several models/approaches for modeling a sample data, the main question is which model is better, in other words, statistically speaking is there a significant difference among them? To address this question in econometrics, several tests have been introduced.

Figure 16 shows detailed procedure to finding the best model/approach in fitting to the sample data when there are several candidates. The first step is to use the GoF tests such as KS, AD, Lill, and $\chi^2$ tests. However, one may not be able to reject any of competing

\textsuperscript{34}The same results have been obtained for filtered data, results are available upon request. It should be noted that there are other graphical tests such as sequential moments, box plot, and histogram that show the same results. The former is based on the extreme value theory and can be used as a suitable alternative to qq-plot, see McNeil et al. (2005).

\textsuperscript{35}Detailed information about these tests as well as the performance of stable and tempered stable distribution can be found in Fallahgoul et al. (2016), Fallahgoul et al. (2017) and Fallahgoul et al. (2018c).
models with certainty. More precisely, the difference of p-values may not be significant, indeed, it is hard to select one model over one another.

After being failed to select the best model from first step, one can check the models/approaches are nested (N) or non-nested (NN). If they are nested, the likelihood ratio test (LRT) can be applied and a model/approach with the highest likelihood ratio will be selected. In this case, the best one can be either a complex or simple model/approach. However, for the non-nested case, one can use the Akaike information criterion (AIC) or Bayesian information criterion (BIC), which a model/approach with the smallest AIC or BIC will be selected.

5 Conclusion

Based on the statistical facts of the Section 4.3, heavy tail models are crucial for capturing the empirical regularities of the return distribution, in particular for a good assessment of the tail risk. Stable distribution plays an important role in modeling tail risk as a heavy tail distribution, but some critical problems remain on the ground. Firstly, since a closed-form formula of its pdf and cdf do not exist, the implementation of a Maximum Likelihood Estimation is difficult. Secondly, a stable distribution has infinite $q$−th moment for all $q \geq \alpha$, where $\alpha < 2$ controls the level of decay of tails, which is why tempered stable distribution have been introduced. Detailed information about different classes of the tempered stable distribution is provided in Section 2. A tempered stable random variable can be constructed either by multiplying the Lévy measure of a symmetric stable random variable by a tempering function or by a subordinator. Tempered stable distributions have heavy tail property and their tails are thinner than the stable and heavier than the normal. Still, tempered stable distributions do not have a closed-form formula for their pdf and cdf. However, their characteristic functions are available in a closed-form. By using the methodology discussed in Section 3, one can calculate estimate the parameters of interest. The results of an empirical study show the superiority of the tempered stable distributions.
A Lévy Measure and Characteristic Function

In this section, we describe the Lévy measure and characteristic function for the stable and different classes of the tempered stable processes.

- The Lévy measure and characteristic function of a stable random variable $X_S$ are given respectively by

$$
\nu(dx) = \left( \frac{C_+}{x^{1+\alpha}} 1_{x>0} + \frac{C_-}{x^{1+\alpha}} 1_{x<0} \right) dx.
$$

$$
\Phi(u; S) = \mathbb{E}[e^{iuX_S}]
= \begin{cases} 
\exp \left( i\mu u - |\sigma u|^\alpha \left( 1 - i\beta \text{sign}(u) \tan \left( \frac{\alpha \pi}{2} \right) \right) \right), & \alpha \neq 1 \\
\exp \left( i\mu u - \sigma |u| \left( 1 + i\beta \frac{2}{\pi} \text{sign}(u) \ln |u| \right) \right), & \alpha = 1
\end{cases}
$$

where

$$
\text{sign } u = \begin{cases} 
1, & u > 0 \\
0, & u = 0 \\
-1, & u < 0,
\end{cases}
$$

$0 < \alpha \leq 2, \sigma \geq 0, -1 \geq \beta \geq 1,$ and $\mu \in \mathbb{R}$.

- The Lévy measure and characteristic function of a Classical Tempered Stable (CTS) random variable $X_{CTS}$ are given respectively by

$$
\nu(dx) = \left( \frac{C_+ e^{-\lambda_+ x}}{x^{1+\alpha}} 1_{x>0} + \frac{C_- e^{-\lambda_- x}}{x^{1+\alpha}} 1_{x<0} \right) dx.
$$

$$
\Phi(u; X_{CTS}) = \mathbb{E}[e^{iuX_{CTS}}]
= \int_{\mathbb{R}} e^{iuX} f_X(x) dx
$$

27
\[
= \exp \left( -iu \Gamma(1 - \alpha)(\lambda_+^{\alpha-1} - \lambda_-^{\alpha-1}) + \Gamma(-\alpha) \left( (\lambda_+ - iu)^\alpha - \lambda_+^{\alpha} + (\lambda_- + iu)^\alpha - \lambda_-^{\alpha} \right) \right).
\]

- The Lévy measure and characteristic function of a Generalised Tempered Stable (GCTS) random variable are given respectively by

\[
\nu(dx) = \left( C_+ \frac{e^{-\lambda_+ x}}{x^{1+\alpha}} 1_{x > 0} + C_- \frac{e^{-\lambda_- x}}{x^{1+\alpha}} 1_{x < 0} \right) dx.
\]

\[
\Phi(u; X_{\text{GCTS}}) = \mathbb{E}[e^{iuX_{\text{GCTS}}}] = \int_{\mathbb{R}} e^{iux} f_X(x) dx = \exp \left( G_R(u; \alpha, C_+, \lambda_+, \lambda_-) + G_I(u; \alpha, C_+, \lambda_+, \lambda_-) \right).
\]

where \( \alpha_+, \alpha_- \in (0, 1) \cup (1, 2), C_+, C_-, \lambda_+, \text{ and } \lambda_- > 0. \)

- The Lévy measure and characteristic function of a Modified Tempered Stable (MTS) random variable are given respectively by

\[
\nu(dx) = C \left( \frac{\lambda^{\alpha+1}_+ K_{\frac{\alpha+1}{2}}(\lambda_+ x)}{x^{\frac{\alpha+1}{2}}} 1_{x > 0} + \frac{\lambda^{\alpha+1}_- K_{\frac{\alpha+1}{2}}(\lambda_- x)}{x^{\frac{\alpha+1}{2}}} 1_{x < 0} \right) dx.
\]

\[
\Phi(u; X_{\text{MTS}}) = \mathbb{E}[e^{iuX_{\text{MTS}}}] = \int_{\mathbb{R}} e^{iux} f_X(x) dx = \exp \left( G_R(u; \alpha, C, \lambda_+, \lambda_-) + G_I(u; \alpha, C, \lambda_+, \lambda_-) \right)
\]
where for $u \in \mathbb{R}$

$$G_R(u; \alpha, C, \lambda_+, \lambda_-) = \sqrt{\pi} 2^{-\frac{\alpha}{2} - \frac{3}{2}} \Gamma(-\frac{\alpha}{2}) \left( (\lambda_+^2 + u^2)^{\frac{\alpha}{2}} - \lambda_+^\alpha + (\lambda_-^2 + u^2)^{\frac{\alpha}{2}} - \lambda_-^\alpha \right)$$

and

$$G_I(u; \alpha, C, \lambda_+, \lambda_-) = \frac{iu \Gamma(\frac{1-\alpha}{2})}{2^{\frac{\alpha+1}{2}}} \left( \lambda_+^{\alpha-1} F \left( 1, \frac{1-\alpha}{2}, \frac{3}{2}, -\frac{u^2}{\lambda_+^2} \right) - \lambda_-^{\alpha-1} F \left( 1, \frac{1-\alpha}{2}, \frac{3}{2}, -\frac{u^2}{\lambda_-^2} \right) \right)$$

where $F$ is the hypergeometric function.$^{36}$

- The Lévy measure and characteristic function of a Rapidly Decreasing Tempered Stable (RDTS random) variable are given respectively by

$$v(dx) = \left( C \frac{e^{-\frac{\lambda_+^2 x^2}{2}}}{x^{\alpha+1} 1_{x>0}} + C \frac{e^{-\frac{\lambda_-^2 x^2}{2}}}{x^{\alpha+1} 1_{|x|<0}} \right) dx.$$  

$$\Phi(u; X_{RDTS}) = \mathbb{E}[e^{iuX_{RDTS}}]$$

$$= \int_{\mathbb{R}} e^{iu x} f_X(x) dx$$

$$= \exp \left( C \left( G(iu; \alpha, \lambda_+) + G(-iu; \alpha, \lambda_-) \right) \right)$$

where

$$G(x; \alpha, \lambda) = 2^{-\frac{\alpha}{2} - 1} \lambda^\alpha \Gamma \left( -\frac{\alpha}{2} \right) \left( M \left( -\frac{\alpha}{2}, 1/2, \frac{x^2}{2\lambda^2} \right) - 1 \right)$$

$$+ 2^{-\frac{\alpha}{2} - \frac{1}{2}} \lambda^{\alpha-1} \Gamma \left( 1 - \frac{\alpha}{2} \right) \left( M \left( \frac{1-\alpha}{2}, 3/2, \frac{x^2}{2\lambda^2} \right) - 1 \right)$$

and $M$ is the confluent hypergeometric function.

- The Lévy measure and characteristic function of a Normal tempered stable (NTS

$^{36}$Detailed information about this hypergeometric function can be found in Andrews and Andrews (1992).
random) variable are given respectively by

\[
v(dx) = \left( \frac{\sqrt{2}\theta^{1-\frac{\alpha}{2}}(\beta^2 + 2\gamma \lambda)^{\frac{\alpha+1}{2}}}{\sqrt{\pi \gamma \Gamma(1 - \frac{\alpha}{2})}} \exp \left( \frac{x\gamma}{\gamma^2} \right) \frac{K_{\frac{\alpha+1}{2}} \left( \frac{|x|\sqrt{\beta^2 + 2\sigma^2 \lambda}}{\gamma^2} \right)}{|x|^{\frac{\alpha+1}{2}}} \right) dx.
\]

\[
\Phi(u; X_t) = \exp \left( iu(\mu - \beta)t - \frac{2\theta^{1-\frac{\alpha}{2}}}{\alpha} \left( (\theta - i\beta u + \frac{\sigma^2 u^2}{2})^{\frac{\alpha}{2}} - \theta^{\frac{\alpha}{2}} \right) t \right).
\]

- The characteristic function of an exponential tilting stable (ETS random) variable is given by

\[
\Phi(u; X_t) = \exp \left( iu(\mu - \beta)t - \frac{2\theta^{1-\frac{\alpha}{2}}}{\alpha} \left( (\theta - i\beta u + \frac{\sigma^2 u^2}{2})^{\frac{\alpha}{2}} - \theta^{\frac{\alpha}{2}} \right) t \right).
\]

The Lévy measure of the ETS random variable is not available in a closed-form (see, Fallahgoul et al. (2018c)).
B Modified Bessel function of the second kind

The modified Bessel function is the solution of a differential equation called modified Bessel’s equation. It is given by

\[
K_{\nu}(z) = \left( \frac{\pi}{2} \right) \frac{I_{-\nu}(z) - I_{\nu}(z)}{\sin(\nu \pi)}
\]

where \( I_{-\nu}(z) \) and \( I_{\nu}(z) \) are the fundamental solutions of the modified Bessel’s equation. \( I_{\nu}(z) \) is given by

\[
I_{\nu}(z) = \left( \frac{z}{2} \right)^{\nu} \sum_{k=0}^{\infty} \frac{\left( \frac{z^2}{4} \right)^k}{k! \Gamma(\nu + k + 1)}
\]

where \( \Gamma(a) \) is the Gamma function.
C Figures

(a) QQ-plot for the classical and ET subordinator for fixed values of $\alpha = 1.9$ and $\lambda = 0.5$.

(b) QQ-plot for the classical and ET subordinator for fixed values of $\alpha = 1.9$ and $\lambda = 1.5$.

Figure 1: QQ-plot: quantile quantile plot. ET: exponential tilting. Dashed red line (−) represents 45 degree line. Blue plus sign (+) shows the value of quantiles. If two random variable follow the same probability law, their quantiles must be on the dashed red line. Deviating from the 45 degree line means their probability law are different.
(a) pdf of ET subordinator: the effect of $\alpha$ for fixed values of $\lambda = 0.5$ and $t = 1$.

(b) pdf of ET subordinator: the effect of $\lambda$ for fixed values of $\alpha = 1.75$ and $t = 1$.

**Figure 2:** pdf: probability density function. ET: exponential tilting. The effect of $\alpha$ and $\lambda$ on the pdf of ET subordinator: panel (a) shows the effect of $\alpha$ and panel (b) shows the effect of $\lambda$. 
(a) pdf of ET subordinator: the effect of $t$ for fixed values of $\alpha = 1.75$ and $\lambda = 0.5$.

(b) cdf of ET subordinator: the effect of $t$ for fixed values of $\alpha = 1.75$ and $\lambda = 0.5$.

Figure 3: pdf: probability density function. cdf: cumulative distribution function. ET: exponential tilting. The effect of $t$ on the pdf and cdf of ET subordinator: panel (a) shows the effect of $t$ on pdf and panel (b) shows the effect of $t$ on cdf.
(a) cdf of ET subordinator: the effect of $\alpha$ for fixed values of $\lambda = 0.5$ and $t = 1$

(b) cdf of ET subordinator: the effect of $\lambda$ for fixed values of $\alpha = 1.75$ and $t = 1$.

**Figure 4**: cdf: cumulative distribution function. ET: exponential tilting. The effect of $\alpha$ and $\lambda$ on the cdf of ET subordinator: panel (a) shows the effect of $\alpha$ and panel (b) shows the effect of $\lambda$. 
(a) pdf for the ET and classical tempered stable distribution for fixed values of \( \alpha = 1.75, \lambda = 0.5 \) and \( t = 2 \).

(b) Tail behavior for pdf for the ET and classical tempered stable distribution for fixed values of \( \alpha = 1.75, \lambda = 0.5 \) and \( t = 2 \).

**Figure 5:** C: classical. ET: exponential tilting. RD: rapidly decreasing. pdf: probability density function. Comparison of behavior tail of ET and classical tempered stable subordinator.
(a) pdf for the ETS distribution: the effect of $\alpha$ for fixed values of $\gamma = 5$, $\lambda = 0.5$, $\beta = 0$, $\mu = 0$ and $t = 2$.

(b) pdf for the ETS distribution: the effect of $\gamma$ for fixed values of $\alpha = 1.75$, $\lambda = 0.5$, $\beta = 0$, $\mu = 0$ and $t = 2$.

**Figure 6**: pdf: probability density function. ETS: exponential tilting stable. The effect of $\alpha$ and $\gamma$ on the pdf of ETS: panel (a) shows the effect of $\alpha$ and panel (b) shows the effect of $\gamma$. 
(a) pdf for the ETS distribution: the effect of $\lambda$ for fixed values of $\alpha = 1.75$, $\gamma = 5$, $\beta = 0$, $\mu = 0$ and $t = 2$.

(b) pdf for the ETS distribution: the effect of $\beta$ for fixed values of $\alpha = 1.75$, $\gamma = 5$, $\lambda = 0.5$, $\mu = 0$ and $t = 2$.

**Figure 7:** pdf: probability density function. ETS: exponential tilting stable. The effect of $\lambda$ and $\beta$ on the pdf of ETS: panel (a) shows the effect of $\lambda$ and panel (b) shows the effect of $\beta$. 
(a) pdf for the univariate ETS distribution: the effect of $\mu$ for fixed values of $\alpha = 1.75$, $\gamma = 5$, $\lambda = 0.5$, $\beta = 0$ and $t = 2$.

(b) pdf for the univariate ETS distribution: the effect of $t$ for fixed values of $\alpha = 1.75$, $\gamma = 5$, $\lambda = 0.5$, $\beta = 0$ and $\mu = 0$.

**Figure 8:** pdf: probability density function. ETS: exponential tilting stable. The effect of $\mu$ and $t$ on the pdf of ETS: panel (a) shows the effect of $\mu$ and panel (b) shows the effect of $t$. 

39
Figure 9: Daily price process for S&P500 and DAX indexes from January 1, 2000 to December 29, 2017.
Figure 10: Daily log-return process for S&P500 and DAX indexes from January 1, 2000 to December 29, 2017.
Figure 11: Filtered daily log-return process for S&P500 and DAX indexes from January 1, 2000 to December 29, 2017.
**Figure 12:** QQ-plots: quantile quantile plot. QQ-plots for daily returns of four indexes of both indexes: S&P500 and DAX. The x-axis shows the quantiles of fitted Normal and Student ’t. The y-axis shows the sample quantiles.
Figure 13: TS: tempered stable. Fitted probability density function (pdf) for different types of distributions. The range of x-axis is 3 standard deviation, i.e., $3 \times \sigma$, of the sample.
Figure 14: The log-plot of the fitted probability density function (pdf) for different types of distributions.
Figure 15: The log-plot of the fitted probability density function (pdf) for different types of distributions. The range of x-axis is 3 standard deviation, i.e., $3 \times \sigma$, of the sample.
Fitted Models

Decision Based on GoF

Rejected Models from GoF

Models are equivalent

LRT for N Models

Models are N/NN

Complex Model is Preferred

Simple Model

Model with Low BIC/AIC is Preferred

Fails to Reject Models from GoF

AIC/BIC for NN Models

Figure 16: GoF: goodness-of-fit. LRT: Likelihood Ratio Test. AIC: Akaike Information Criteria. BIC: Bayesian Information Criteria. N: nested. NN: not nested. The algorithm for comparison two different fitted models,
References


