The Hybrid Pareto Distribution, Implied Risk-Neutral Density and Option Pricing

Mr. Purin Luanloy

An Independent Study Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Financial Engineering
Department of Banking and Finance
FACULTY OF COMMERCE AND ACCOUNTANCY
Chulalongkorn University
Academic Year 2020

การแจกแจงพาเรโตแบบผสม ความหนาแน่นทางความน่าจะเป็นโดยนัยที่เป็นกลางต่อความเสี่ยง และการคำนวนราคาตราสารสิทธิ

นายภูรินท์ เลื่อนลอย

สารนิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมการเงิน ภาควิชาการธนาคารและการเงิน
คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2563

Independent Study Title	The Hybrid Pareto Distribution, Implied Risk-Neutral Density and Option Pricing
By	Mr. Purin Luanloy
Field of Study	Financial Engineering
Thesis Advisor	Associate Professor THAISIRI WATEWAI, Ph.D.

Accepted by the FACULTY OF COMMERCE AND ACCOUNTANCY, Chulalongkorn University in Partial Fulfillment of the Requirement for the Master of Science

INDEPENDENT STUDY COMMITTEE

.... Chairman
(Associate Professor SIRA SUCHINTABANDID, Ph.D.)

.... Advisor
(Associate Professor THAISIRI WATEWAI, Ph.D.)

.... Examiner
(TANAWIT SAE SUE, Ph.D.)

ภูรินท์ เลื่อนลอย : การแจกแจงพาเรโตแบบผสม ความหนาแน่นทางความน่าจะเป็นโดยนัยที่เป็นกลางต่อความเสี่ยง และการคำนวนราคาตราสารสิทธิ. (The Hybrid Pareto Distribution, Implied Risk-Neutral Density and Option Pricing) อ.ที่ปรึกษาหลัก : รศ. ดร.ไทยศิริ เวทไว

สารนิพนธ์ชิ้นนี้มีจุดประสงค์ที่จะพัฒนาตัวแบบจำลองสำหรับการคำนวณราคาตราสารสิทธิแบบยุโรปโดยมีทฤษฎีค่าสุดขีดเป็นพื้นฐาน วิธีการศึกษาเริ่มจากการตั้งสมมติฐานให้ค่าลบของผลตอบแทนของดัชนี Standard and Poor's 500 มีการแจกแจงแบบพาเรโตแบบผสมในโลกความน่าจะเป็นที่เป็นกลางต่อความเสี่ยง (Risk-neutral probability) รวมถึงมีสมมติฐานให้การแจกแจงดังกล่าวเป็นประเภท fat-tailed จากนั้นจึงสร้างสมการสำหรับคำนวณราคาตราสารสิทธิทั้งประเภท call และ put โดยใช้วิธีการคำนวณราคาในโลกความน่าจะเป็นที่เป็นกลางต่อความเสี่ยง (Risk-neutral pricing) เช่นเดียวกัน การศึกษานี้ใช้ตัวแบบจำลองจากทฤษฎีค่าสุดขีดวางนัยทั่วไปเสนอโดย Markose and Alenton (2011) เป็นตัวแบบวิเคราะห์เชิงเปรียบเทียบ ค่าพารามิเตอร์สำหรับแต่ละตัวแบบจำลองประมาณได้จากการหาค่าเหมาะสมต่ำสุดของรากที่สองของค่าเฉลี่ยความผิดพลาดกำลังสอง ผลการศึกษาแสดงให้เห็นว่าตัวแบบจำลองแบบพาเรโตแบบผสมให้ค่ารากที่สองของค่าเฉลี่ยความผิดพลาดกำลังสองที่ต่ำกว่าตัวแบบวิเคราะห์เชิงเปรียบเทียบในตราสารสิทธิชนิดที่เหลือเวลา 30 วันก่อนหมดอายุ นอกจากนี้ตัวแบบดังกล่าวยังให้ผลลัพธ์ที่ดีกว่าตัวแบบวิเคราะห์เชิงเปรียบเทียบสำหรับตราสารสิทธิประเภท at-the-money แต่สำหรับตราสารสิทธิประเภท out-of-the-money นั้นจะให้ผลลัพธ์ที่ด้อยลงไป

สาขาวิชา	วิศวกรรมการเงิน	ลายมือชื่อนิสิต ....
ปีการศึกษา	2563	ลายมือชื่อ อ.ที่ปรึกษาหลัก ....

6182945326 : MAJOR FINANCIAL ENGINEERING
KEYWORD: option pricing, risk-neutral, Generalized Extreme Value distribution, Hybrid Pareto distribution, Extreme Value theory, risk-neutral pricing

Purin Luanloy : The Hybrid Pareto Distribution, Implied Risk-Neutral Density and Option Pricing. Advisor: Assoc. Prof. THAISIRI WATEWAI, Ph.D.

This paper aims to develop a new European option pricing model based on the Extreme Value Theory (EVT). We assume that, in the risk-neutral probability measure, simple negative returns of the S&P500 index follow the Hybrid Pareto (HP) distribution. Then, we derive closed-form pricing formulas for call and put options according to the risk-neutral pricing method. Additionally, we assume that the distribution has a fat tail. Our study's benchmark model is the Generalized Extreme Value (GEV) model proposed by Markose and Alenton (2011). We estimate model parameters by minimizing the root-mean-square error. The results show that the HP model provides less root-mean-square errors than the GEV model in a 30-day to expiration options. Moreover, it outperforms the benchmark for the at-the-money options, but has a poorer fit for the out-of-the-money options.

Field of Study:	Financial Engineering	Student's Signature ....
Academic Year:	2020	Advisor's Signature ....

ACKNOWLEDGEMENTS

This research paper would not have been possible without the exceptional support of every professor from the Department of Financial Engineering, especially my advisor, Associate Professor Thaisiri Watewai. I would like to express my deep appreciation to him for his diligence, professionalism, and constructive critiques during the past year, which inspire me professionally and personally. Also, I wish to extend my thanks to the MFE #7 class. The assistance of each one has turned the hope of continuing the studies. Although the courses require a great deal of effort and practice, the fact that we leave no one behind empowers me, and makes me believe that we will accomplish it in the most effective ways. Finally, it is a true pleasure to have my family throughout this journey. Even in the most difficult times, they are always there offering me guidance and warm encouragement. To my sister, I am thankful for her contribution in the proofreading stage. I wish I had finished the studies earlier so that we would have graduated at the same time. However, I ensure that I will not keep my supporters waiting any longer for my next commencement.

Purin Luanloy

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

1. INTRODUCTION

In finance, the risk-neutral density (RND) is a probability measure widely used in the pricing framework. It is derived from the first fundamental theorem of asset pricing (FTAP I), stating that this measure must exist for a market to be arbitrage-free. Under the risk-neutral probability measure, any discounted asset prices are martingale. We can use this fact to price any financial products, for example, an option on a single stock. That is, the price of derivative today is equal to the expected value of its payoff function at maturity discounted by the risk-free rate. One can compute this expectation by integrating the payoff function with respect to its density function under the risk-neutral measure. This method is called risk-neutral pricing. Because the risk-neutral itself cannot be observed directly, the limitation here is that we do not know what distribution the risk-neutral density is. One solution is to estimate the implied risk-neutral density from observed prices given by the market data. To do so, we need an assumption of the density function and the model in order to extract its parameters. The more efficient the model is, the more accurate the estimated prices are.

The Black-Scholes (1973) is a well-known model for European option pricing derived from the risk-neutral pricing method. According to the research, it is assumed that the dynamic of stock prices follows Geometric Brownian Motion (GBM) with constant mean and constant volatility across its maturity. It implies that stock prices are distributed as log-normal and the returns of stocks are distributed as the normal distribution, a symmetric distribution. Although this approach is quite interesting, it fails to capture skewness and excess kurtosis of asset returns. Several papers provide empirical results showing that the distribution of asset returns is not symmetric. It has negative skewness and fat tail properties. The drawback of those problems is that its implied volatilities are not consistent across strike prices for options with the same time to maturity. This effect is known as the volatility smile. Consequently, the Black-Scholes model will overprice out-of-the-money call options and underprice in-the-money call options. This result shows how highly sensitive options prices are to extreme events in the tail of its density. In response, later academic researchers have proceeded to overcome mentioned drawbacks by proposing alternative distributions for an asset return.

Based on a review of literature on this topic by Jackwerth (1999), it suggested that numerous methods can be categorized into two main categories, namely, parametric and non-parametric. For parametric methods, they can be divided into three sub-categories. The first method is the expansion method. In this method, we typically start with a general distribution, mostly normal or log-normal. Then we add some conditions to coefficient terms to provide more flexibility to the models. A serious weakness of this sub-category is that the implied risk-neutral density can take on negative value and violate the positivity constraint of density function. The second method is the generalized distribution method. We include more parameters further from two parameters in normal or log-normal distributions. The third method is the mixture method. In this method, we provide greater flexibility by combining a couple of simple distributions using the weighted sum approach. The main disadvantage of mixture models is that the number of parameters is large. To illustrate, the mixture of three log-normal distributions needs us to estimate eight parameters. Non-parametric methods are also divided it into three sub-categories. The first method is the kernel method. We use a regression technique to fit given data without considering a parametric form of the function. The second method is the maximum-entropy method. We try to fit observed data with non-parametric distribution while consider certain constraints. The third method is the curve fitting method. In this method, we basically try to fit the implied volatilities or the risk-neutral density with some flexible function. By using non-parametric methods, because there are not many extraordinarily high or low strike prices traded in the market, most of them fail to capture tail behavior for the density functions. Therefore, parametric methods are more appropriate in this perspective.

According to Markose and Alentorn (2011), they introduced a new generalized distribution in the parametric method, that is, the Generalized Extreme Value (GEV) distribution. It is developed within the Extreme Value Theory (EVT) framework, which is useful in estimating the distribution of extreme events of financial markets. Moreover, it brings more flexibility to capture the negative skewness and excess kurtosis of asset returns. Their article discussed the benefits of using the GEV approach in option pricing as follows: (i) It can provide a closed-form solution for both call and put options. (ii) There are only two parameters in this model to be estimated, the tail shape parameter (ξ) and the scale parameter (σ). This is equal to the number of parameters in log-normal and normal distribution models. (iii) It can be categorized into three special cases depending on the value of the tail shape parameter, namely, the Weibull thin-tailed type, the Gumbel short-tailed type, and the Fréchet fat-tailed type. (iv) For the Fréchet type, it exhibits a fat tail on the right side of the density function. We can adopt Fréchet distribution to model negative returns. (v) The GEV model provides a market-implied tail index (ξ^{-1}), and this index can be used to interpret market conditions in each period. (vi) It removes pricing biases associated with the Black-Scholes model.

There is also another approach within the Extreme Value Theory framework. Instead of considering a whole distribution of sample maxima which results in GEV distribution, we may use a method called Peak over Threshold (Davison and Smith (1990)) to examine the behavior of the exceedances of a random variable over a high threshold value. That is, we narrow our focuses only on the tail of the distribution. Pickands III (1975) states that, for most of any random variables, given a high threshold value, the distribution of their exceedances over that threshold converges to the Generalized Pareto (GP) distribution. Similar to the GEV distribution, the GP distribution provides the tail index (ξ) to determine the tail heaviness of the distribution. However, the main limitation of GP distribution is that its domain does not support the whole real axis. In order to apply GP distribution into the option pricing context, we need an extension method to deal with this problem.

In this paper, we adopt the method called the Hybrid Pareto model by Carreau and Bengio (2009) to price European-style options. It is a combination of the Normal distribution and the Generalized Pareto distribution. Instead of using the basic weighted sum method from previous research, they developed this model by stitching a GP tail to a Normal distribution at some proper threshold. We call the GP tail component above the threshold as a tail model since it contains all extreme values. The Normal distribution component below the threshold is called bulk model describing all the rest non-extreme values. The advantages of using the Hybrid Pareto model in option pricing can be discussed as follows: (i) It provides a closed-form solution. (ii) Initially, this model has five parameters to estimate: the threshold value (α), the mean of Normal component (η), the variance of Normal component (β), the tail index in GP component (ξ) and the scale parameter (σ). However, after adding two constraints in order to make a whole density smooth and continuous at the threshold, the set of parameters reduce to three (η, σ and ξ). (iii) Consequently, it provides more flexibility to model the tail behavior compare to the GEV model.

In conclusion, this paper proposes a new alternative model for option pricing. Under the Extreme Value Theory framework, stock losses (or negative return) under the risk-neutral measure are assumed to follow the Hybrid Pareto distribution. Using the risk-neutral pricing method, one can derive a closed-form pricing formula. Given observed historical data on option prices in the market, the implied-risk neutral density can be obtained using the optimization approach. We use the GEV model by Markose and Alentorn (2010) as a benchmark. By using the Hybrid Pareto model, we aim to archive better performances in terms of in-sample pricing evaluated root-mean-square error (RMSE).

1.1 Research Objective

Using the GEV model as a benchmark, this paper proposes a new alternative model with a hybrid method and aims to improve model performance by reducing root-mean-square error and pricing bias.

Table 1: Detail of models used in this paper.

Model Type	Model Name	Distribution		Parameters			
		Bulk Model	Tail Model	η	β	ξ	σ
1. Benchmark Model	GEV	Generalized Extreme Value		-	-	✓	✓
2. Proposed Model	Hybrid Pareto	Normal	Generalized Pareto	✓	✓	✓	-

1.2 Research Hypothesis

Due to an effort we provided for the tail component in the Hybrid Pareto model, we expect that our proposed model would outperform the benchmark model in two aspects.

1.2.1 It provides better performances in terms of average RMSE for deep out-of-the-money options.

1.2.2 It provides better performances in terms of average RMSE for long expiration time (90 days).

The rest of this paper is organized as follows: Section 2 provides the theory of no-arbitrage pricing under risk-neutral and previous related models. Besides, it exhibits some applications of Extreme Value Theory in financial areas. Section 3 describes the data used in our study. Section 4 shows how to derive the closed-form option pricing formula from our proposed model. Then provides a method to extract model parameters and its implied risk-neutral density using an optimization technique. Section 5 discusses about the results we got from mentioned methodology. Section 6 is the conclusion for our study.

2. LITERATURE REVIEWS

2.1 The First Fundamental Theorem of Asset Pricing (FTAP I)

Let S_T denote a price of an underlying asset at time T, $C_t(T,K)$ and $P_t(T,K)$ denote European call and put option price observed at time t with maturity date T and strike price K. In this paper, we assume a constant interest rate r. According to Harrison and Pliska (1981), for a no-arbitrage market, there must exist the probability measure $\widetilde{\mathbb{P}}$ corresponding to the risk-neutral density function $f(\cdot)$ such that the option prices are given by the expected discounted payoff:

$$\begin{aligned} C_t(T,K) &= e^{-r(T-t)}\widetilde{\mathbb{E}}^{\mathbb{P}}\left[(S_T-K)^+ \,\middle|\, S_t\right] \\ &= e^{-r(T-t)}\int_K^{\infty}(S_T-K)\,f(S_T)dS_T \end{aligned} \tag{2.1}$$

and

$$\begin{aligned} P_t(T,K) &= e^{-r(T-t)}\widetilde{\mathbb{E}}^{\mathbb{P}}\left[(K-S_T)^+ \,\middle|\, S_t\right] \\ &= e^{-r(T-t)}\int_{-\infty}^{K}(K-S_T)\,f(S_T)dS_T. \end{aligned} \tag{2.2}$$

2.2 Implied Risk-Neutral Density from Option Price

According to the first fundamental theorem of asset pricing (FTAP I), the existence of risk-neutral density (RND) implies no-arbitrage condition on the market. Following this theory, all the discounted asset prices are martingale under the risk-neutral probability measure $\widetilde{\mathbb{P}}$. In the real world, we cannot observe the RND directly. It needs market data to imply its structure. In this section, we review the previous methods for recovering the RND from the given options price.

Jackwerth (1999) suggested that the methods can be categorized into two approaches: parametric and non-parametric methods.

2.2.1 Parametric Approach

The parametric approach can be divided into three groups.

The first method is the expansion method. This method generally starts with some basic distribution (e.g., log-normal, normal) and then adds correlation terms in order to make the model more flexible and well fit to the observed data. A drawback of the expansion method is that as some conditions are added, the RND function can take on a negative value. Abadir and Rockinger (1997) assumed that RND follows the density functional based on the confluent hypergeometric function (DFCH). Also, Bu and Hadri (2007) used the DFCH method to compare the estimation performance with the smoothed implied volatility smile method (SML) and found that the DFCH method outperforms the SML method by extracting RND function from option price.

The second method is the generalized distribution method. It provides more flexibility to fit market data. In spite of using only two basic parameters, which are mean and volatility, the 3rd moment (skewness) and the 4th moment (kurtosis) are included in this method. Aparicio and Hodges (1998) used a generalized beta distribution. Bookstaber and McDonald (1987) used the log-normal, gamma, and exponential distribution along with several burr type distributions. Markose and Alentorn (2010) introduced the Generalized Extreme Value (GEV) distribution for estimating RND function. They assumed that a simple negative return follows GEV distribution. Then, they derived a closed-form pricing formula from it. The advantage of using a GEV distribution is that we can obtain tail index ξ, which can be used to describe market conditions in each period.

The third method is the mixture method. It is a combination of two or more distributions to create a new distribution. Typically, the weighted average is widely used. Melick and Thomas (1997) used mixtures of three log-normal distributions. Ritchey (1990) applied mixtures of normal distributions for log return. Even though this method is more flexible in terms of the shape of a density function, it suffers from a large number of parameters. Thai is, the mixtures of two log-normal distributions method need five parameters, but, considering one step further, the mixtures of three log-normal distribution method needs eight parameters. Santos and Guerra (2015) compared the performance of the mixture of log-normal distributions (MLN), the smoothed implied volatility smile (SML), the density functional based on the confluent hypergeometric function (DFCH) and the Edgeworth expansion (EE). They found that DFCH and MLN have outperformed the other two.

2.2.2 Non-Parametric Approach

Without any assumption on the RND function, the non-parametric method aims to fit the model with less restriction. It allows a more general form of option pricing formula. This approach can be divided into three groups.

The first method is the Kernel method. It uses a regression technique to fit a function to observed data. The farther away the observed data are from, the less likely a proper function is. Aït-Sahalia and Lo (1998) used five dimensions in the regression process: stock price, strike price, time to maturity, interest rate, and dividend yield. Rookley (1997) used a bivariate kernel estimator across expiration time and moneyness.

The second method is the maximum-entropy method. This method is similar to the Bayesian approach. It requires the prior distribution of the density function. Furthermore, the posterior RND is derived from maximizing a cross-entropy subject to three constraints: positivity of density function, integrating the density to one, and the correction in terms of the option price. Buchen and Kelly (1996) used the uniform and the log-normal distribution as prior distributions. Stutzer (1996) used the historical distribution of the asset price as prior. Only a price constraint was used in his study.

The third method is the curve fitting method. This method usually uses some general polynomial functions to fit the observed data. Shimko (1993) employed the quadratic polynomial function to fit the implied volatility smile and then used it to compute option prices. The risk-neutral density can be derived from taking the second derivative of the option prices with respect to strike prices. Campa, Chang, and Reider (1998) used cubic splines to fit the volatility smile. Rubinstein (1996) tried to estimate the RND directly by minimizing the error of discretized probabilities to the log-normal prior distribution subject to the option prices and underlying asset constraints.

2.3 Extreme Value Framework

2.3.1 Generalized Extreme Value (GEV) Distribution

In this topic, we discuss how generalized extreme value distribution was derived based on the book named Financial Risk Forecasting: The Theory and Practice of Forecasting Market Risk, with implication on R and MATLAB by Danielsson (2011).

Consider IID T random variables $X_1, X_2, \dots, X_T$. Let M_T denote maxima in T samples:

$$M_T = \max(X_1, X_2, \dots, X_T)\,. \tag{2.3}$$

From the Fisher and Tippet (1928) and Gnedenko (1943) theorem, it states that if the sample size is large ($T \to \infty$). Then, standardized maxima asymptotically distributed as generalized extreme value distribution $F(\cdot)$:

$$F(x) = \lim_{T\to\infty} \mathbb{P}\left(\frac{M_T - a_T}{b_T} \le x\right), \tag{2.4}$$

where the constants $a_T > 0$ and $b_T > 0$ exists and are defined as $a_T = T\mathbb{E}(X_1)$ and $b_T = \sqrt{\mathrm{Var}(X_1)}$.

Let ξ denote a shape parameter. The limiting cumulative distribution $F(\cdot)$ becomes the generalized extreme value (GEV) distribution:

$$F(x \mid \xi) = \begin{cases} \exp\left\{-(1+\xi x)^{-\frac{1}{\xi}}\right\}, & \xi \neq 0 \\ \exp(-\exp(-x)), & \xi = 0 \end{cases}. \tag{2.5}$$

According to Mises (1936), given the location parameter $\mu \in \mathbb{R}$, scale parameter $\sigma > 0$ and tail shape parameter $\xi \in \mathbb{R}$, the standardized GEV distribution is given by:

$$G(z \mid \xi, \mu, \sigma) = \begin{cases} \exp\left\{-\left(1+\frac{\xi(z-\mu)}{\sigma}\right)^{-\frac{1}{\xi}}\right\}, & \xi \neq 0 \\ \exp\left\{-e^{\left(-\frac{z-\mu}{\sigma}\right)}\right\}, & \xi = 0 \end{cases}, \tag{2.6}$$

defined on $\{z: 1 + \xi(z-u)/\sigma > 0\}$.

By taking the first derivative of the cumulative density function $G(\cdot)$, we get the following probability density function $g(\cdot)$:

$$g(z \mid \xi, \mu, \sigma) = \begin{cases} \frac{1}{\sigma}\left(1+\frac{\xi(z-\mu)}{\sigma}\right)^{-\frac{1}{\xi}-1} \exp\left\{-\left(1+\frac{\xi(z-\mu)}{\sigma}\right)^{-\frac{1}{\xi}}\right\}, & \xi \neq 0 \\ \frac{1}{\sigma} e^{\left(-\frac{z-\mu}{\sigma}\right)} \exp\left\{-e^{\left(-\frac{z-\mu}{\sigma}\right)}\right\}, & \xi = 0 \end{cases}. \tag{2.7}$$

2.3.2 Extreme Value Theory (EVT) in Financial Market

Rocco (2014) surveyed the uses of extreme value theory for testing the distributional assumption, value-at-risk and expected shortfall calculation, asset allocation under safety-first constraints, and the study of contagion and dependence. In our study, we focus on an application to test the distributional assumption.

Koedijk, Schafgans, and De Vries (1990) employed the extreme value theory approach to examine the amount of tail-fatness. Using bilateral foreign exchange rates of the European monetary system from 1971 to 1987, they found that the tail index parameter is most likely around two and possibly lower than two. This result implies that we should reject the assumption of normal and of the log-normal distribution.

Vilasuso and Katz (2000) studied stock market index prices in 17 countries from 1980 to 1997 and tested the hypothesis of whether their returns favor the heavy-tailed stable distribution or not. They concluded that the results better fit with Student t and ARCH processes, which are in the fat-tailed family.

LeBaron and Samanta (2005) also adopted the EVT to estimate the level of "fatness" in the tails in the equity market. They suggested that, as returns are known to be fat-tailed, we can limit our study only on the Frechét type, where the shape parameter ξ is greater than zero.

2.3.3 Implied RND Function from Option Price Using GEV Distribution

Markose and Alentorn (2010) employed the GEV distribution approach to obtain an option pricing formula. First, they assumed that negative simple returns L_T follow the GEV distribution in (2.7) when it is a Frechét type ($\xi > 0$):

$$f\left(L_T \mid \xi, \mu, \sigma\right) = \frac{1}{\sigma}\left(1 + \frac{\xi(L_T - \mu)}{\sigma}\right)^{-\frac{1}{\xi}-1} \exp\left\{-\left(1 + \frac{\xi(L_T - \mu)}{\sigma}\right)^{-\frac{1}{\xi}}\right\}, \tag{2.8}$$

where we define L_T:

$$L_T = -\frac{S_T - S_t}{S_t} = 1 - \frac{S_T}{S_t}. \tag{2.9}$$

We can transform the density function of L_T in (2.8) to the density function of an underlying price at time T by:

$$g(S_T) = f(L_T)\left|\frac{\partial L_T}{\partial S_T}\right| = \frac{1}{S_t} f(L_T). \tag{2.10}$$

In order to obtain the RND function of S_T, we substitute (2.8) into (2.10). Then $g(S_T)$ becomes:

$$g(S_T) = \frac{1}{S_t \sigma}\left(1 + \frac{\xi(L_T - \mu)}{\sigma}\right)^{-\frac{1}{\xi}-1} \exp\left\{-\left(1 + \frac{\xi(L_T - \mu)}{\sigma}\right)^{-\frac{1}{\xi}}\right\}, \tag{2.11}$$

with the condition:

$$1 + \frac{\xi(L_T - \mu)}{\sigma} = 1 + \frac{\xi}{\sigma}\left(1 - \frac{S_T}{S_t} - \mu\right) > 0. \tag{2.12}$$

From (2.12), the upper bound of S_T becomes:

$$S_T < S_t\left(1-\mu+\frac{\sigma}{\xi}\right). \tag{2.13}$$

We can compute a European call option price by using the risk-neutral pricing equation in (2.1) along with the RND function of underlying asset price given by (2.11) and the upper limit of S_T in (2.13):

$$C_t(T,K) = e^{-r(T-t)}\int_K^{S_t\left(1-\mu+\frac{\sigma}{\xi}\right)}(S_T-K)\frac{1}{S_t\sigma}\left(1+\frac{\xi(L_T-\mu)}{\sigma}\right)^{-\frac{1}{\xi}-1}$$
$$\times \exp\left\{-\left(1+\frac{\xi(L_T-\mu)}{\sigma}\right)^{-\frac{1}{\xi}}\right\}dS_T \tag{2.14}$$

For simplicity, we do the change of variable given by:

$$y = 1+\frac{\xi(L_T-\mu)}{\sigma} = 1+\frac{\xi}{\sigma}\left(1-\frac{S_T}{S_t}-\mu\right). \tag{2.15}$$

Then we have:

$$C_t(T,K) = e^{-r(T-t)}\int_H^0\left(S_t\left(1-\mu-\frac{\sigma}{\xi}(y-1)\right)-K\right)$$
$$\times\frac{1}{S_t\sigma}\left(y^{-1-\frac{1}{\xi}}\right)\exp\left(-y^{-\frac{1}{\xi}}\right)\left(-\frac{S_t\sigma}{\xi}\right)dy, \tag{2.16}$$

where:

$$H = 1+\frac{\xi}{\sigma}\left(1-\frac{K}{S_t}-\mu\right). \tag{2.17}$$

(2.16) can be simplified to:

$$C_t(T,K) = e^{-r(T-t)}\left\{S_t\left(\left(1-\mu+\frac{\sigma}{\xi}\right)e^{-H^{-1/\xi}}-\frac{\sigma}{\xi}\Gamma\left(1-\xi,H^{-1/\xi}\right)\right)-Ke^{-H^{-1/\xi}}\right\}, \tag{2.18}$$

where $\Gamma(a,b)$ is an incomplete Gamma function given by:

$$\Gamma(a,b) = \int_b^\infty x^{a-1}e^{-x}dx\,. \tag{2.19}$$

Following by the same method above, the formula for the European put option yields:

$$P_t(T,K) = e^{-r(T-t)}$$
$$\times\left(K\left(e^{-h^{-\frac{1}{\xi}}}-e^{-H^{-\frac{1}{\xi}}}\right)-S_t\left(\left(1-\mu+\frac{\sigma}{\xi}\right)\left(e^{-H^{-\frac{1}{\xi}}}-e^{-h^{-\frac{1}{\xi}}}\right)-\frac{\sigma}{\xi}\Gamma\left(1-\xi,h^{-\frac{1}{\xi}},H^{-\frac{1}{\xi}}\right)\right)\right) \tag{2.20}$$

where $\Gamma(a,b,c)$ is the generalized Gamma function given by:

$$\Gamma(a,b,c) = \int_b^c x^{a-1}e^{-x}dx\,, \tag{2.21}$$

and

$$h = 1 + \frac{\xi}{\sigma}(1-\mu) > 0. \tag{2.22}$$

Here, we can use the fact that the futures price of an underlying at time t (F_t) is equal to the expectation of underlying prices at maturity date T under the risk-neutral measure:

$$\begin{aligned} F_t &= \widetilde{\mathbb{E}}^{\mathbb{P}}(S_T \mid S_t) \\ &= 1 - S_t \widetilde{\mathbb{E}}^{\mathbb{P}}(L_T) \\ &= 1 - S_t\left(\mu + \left(\frac{\Gamma(1-\xi)-1}{\xi}\right)\sigma\right). \end{aligned} \tag{2.23}$$

From (2.23), we can get rid of the location parameter μ from the model by using the observed futures prices as follow:

$$u = 1 - \frac{F_{t,T}}{S_t} - \left(\frac{\Gamma(1-\xi)-1}{\xi}\right)\sigma. \tag{2.24}$$

In the end, one can estimate the model parameters by minimizing the sum squared error (SSE) between the option prices given by market data and the prices estimated by the GEV model. Denote $\hat{C}_t(T,K)$, and $\hat{P}_t(T,K)$ as the estimated call and put option prices from the GEV model. The SSE can be computed by:

$$SSE_t = \min_{\xi,\sigma}\left\{\left(\sum_{i=1}^{N}\left(\hat{C}_{i,t}(T,K \mid \xi,\sigma) - C_{i,t}(T,K)\right)^2\right) + \left(\sum_{j=1}^{M}\left(\hat{P}_{j,t}(T,K \mid \xi,\sigma) - P_{j,t}(T,K)\right)^2\right)\right\}, \tag{2.25}$$

where N and M are the number of observed call and put option prices at time t.

2.3.4 Peak Over Threshold (PoT) Method and Generalized Pareto Distribution

This method is first developed by Todorovic and Zelenhasic (1970). Instead of using the maxima, it considers the exceedance over a given high threshold (α) to model extreme events. Given any random variable x, we consider a probability of x exceeding α conditional on x being greater than α:

$$P(X-\alpha | X > \alpha). \tag{2.26}$$

Davision and Smith (1990) suggested that we can model those excesses by using Generalized Pareto (GP) distribution, which is also a limiting distribution as with the GEV distribution case. The GP density function above α is given by Carreau and Bengio, 2009:

$$g(x-\alpha \mid \xi,\sigma) = \begin{cases} \frac{1}{\sigma}\left(1+\frac{\xi}{\sigma}(x-\alpha)\right)^{-1/\xi-1}, & \xi \neq 0 \\ \frac{1}{\sigma}\exp\left(-\frac{x-\alpha}{\sigma}\right), & \xi = 0 \end{cases}, \tag{2.27}$$

define on $\{x : x \geq \alpha\}$ when $\xi \geq 0$ and $\{x : \alpha \leq x \leq \alpha - \beta/\xi\}$ when $\xi < 0$.

2.4 Extremal Mixture Model

In statistics, extreme events occur with a small probability. In this section, we review papers that studied the extremal mixture models. In order to describe the tail event more efficiently, we split the density function into two components. First, the bulk model contains all non-extreme events. Second, the tail model contains all the excess events from the bulk model at a given threshold α.

2.4.1 Behrens et al. (2004)

They proposed a two-component model by using the Bayesian approach. Under a threshold α, the cumulative density function is given by a particular distribution $H(\cdot)$ with a set of parameters Θ. For the data above a given threshold, they were assumed to follow generalized Pareto distribution $G(\cdot)$ with parameters ξ, σ, and μ. Therefore, the whole cumulative density function $F(\cdot)$ can be written as:

$$F(x \mid \Theta, \xi, \sigma, \mu) = \begin{cases} H(x|\Theta), & x \leq \alpha \\ H(\alpha|\Theta) + [1 - H(\alpha|\Theta)]G(x \mid \xi, \sigma, \mu), & x > \alpha \end{cases}. \tag{2.28}$$

This model is simple and straightforward. However, it fails to satisfy the continuity condition and suffers from obtaining an explicit form of density function $f(\cdot \mid \Theta, \xi, \sigma, \mu)$.

2.4.2 Carreay and Bengio (2009)

Their study employed the Peak over Threshold (PoT) method to model events that excess from a threshold α (tail model). The probability of a random variable X excessing α given $X > \alpha$ is modeled by generalized Pareto distribution $g(\cdot)$ with two parameters ξ and σ:

$$\mathbb{P}(x - \alpha \mid x > \alpha) \sim g(x - \alpha \mid \xi, \sigma), \tag{2.29}$$

where the generalized Pareto distribution $g(\cdot)$ is given by (2.27).

For the bulk model, the author used the normal distribution $h(\cdot \mid \eta, \beta)$ with two parameters, which are mean (η) and variance (β). In order to make a model smooth and continuous at $x = \alpha$, the two conditions must hold. The first condition is:

$$h(\alpha|\eta, \beta) = g(0 \mid \xi, \sigma)$$

$$\frac{1}{\sqrt{2\pi}\beta} \exp\left(-\frac{(\alpha - \eta)^2}{2\beta^2}\right) = \frac{1}{\sigma}$$

$$\exp\left(-\frac{(\alpha - \eta)^2}{2\beta^2}\right) = \frac{\sqrt{2\pi}\beta}{\sigma}, \tag{2.30}$$

and the second one is:

$$h'(\alpha|\eta, \beta) = g'(0 \mid \xi, \sigma)$$

$$-\frac{(\alpha - \eta)}{\sqrt{2\pi}\beta^3} \exp\left(-\frac{(\alpha - \eta)^2}{2\beta^2}\right) = -\frac{(1 + \xi)}{\sigma^2}. \tag{2.31}$$

As a result of the two constraints above, we can obtain the values of σ and α in terms of other parameters by combining (2.30) and (2.31):

$$\sigma(\xi,\beta)=\frac{\beta(1+\xi)}{\sqrt{W(z)}}, \text{ and} \tag{2.32}$$

$$\alpha(\xi,\eta,\beta)=\eta+\beta\sqrt{W(z)}, \tag{2.33}$$

where $z=(1+\xi)^2/2\pi$ and $W(z)$ is the Lambert W function.

We leave three parameters ξ,η, and β to be model parameters. Therefore, the whole hybrid Pareto distribution $f(\cdot)$ is given by:

$$f\left(x\,\middle|\,\eta,\beta,\xi\right)=\begin{cases}\dfrac{1}{\gamma}h(x|\eta,\beta), & x\le\alpha\\ \dfrac{1}{\gamma}g\left(x-\alpha\,\middle|\,\xi\right), & x>\alpha\end{cases}, \tag{2.34}$$

where γ is added so that the density integrates to one and is given by:

$$\gamma(\xi)=1+\frac{1}{2}\left(1+Erf\left(\sqrt{\frac{W(z)}{2}}\right)\right), \tag{2.35}$$

where $Erf(\cdot)$ is the error function defined by:

$$Erf(a)=\frac{2}{\sqrt{\pi}}\int_o^a e^{-t^2}dt. \tag{2.36}$$

3. DATA

In this study, all the financial products used are priced in USD. In addition, quarterly -period data are selected for calibrating the model parameters. There are four categories.

3.1 Underlying Asset

S&P 500 index closing prices from December 2001 to December 2015 are used. During this period, they are two significant crises: the 9/11 attack in 2001 and the subprime crisis in 2008. These events can be considered as extreme events.

3.2 Futures Contract

The futures contracts can be used in the GEV model (benchmark) to reduce the number of parameters, as suggested in (2.23). Let $F_{t,T}$ be the S&P 500 futures prices observed at time t with T years to maturity. Their expiration date is on the 3rd Friday on a 3-month basis (March, June, September, and December). The data are provided by Bloomberg.

3.3 European option on S&P 500

Along with the futures contracts, the options also expire on the same schedule. We apply four conditions suggested by Markose and Alentorn (2010) to filter the data. Options with the following properties are eliminated. **First**: Options with zero volume traded on any given days. **Second**: Options quoted at zero prices. **Third:** Option prices that violate the monotonicity condition. Here, two cross-section analyses are made to the options data.

3.3.1 Time-to-Maturity

We divide options into three types:

(i) Short time to maturity: options with 30 days left until the expiration date.

(ii) Medium time to maturity: options with 60 days left until the expiration date.

(iii) Long time to maturity: options with 90 days left until the expiration date.

3.3.2 Moneyness

According to Figlewski (2002), the moneyness of an option can be computed by:

$$\text{Moneyness} = \frac{1}{\sigma_{BS}\sqrt{T}} \ln\left(\frac{S_t}{Ke^{-rT}}\right), \tag{3.1}$$

where σ_{BS} is an implied volatility obtained by the Black-Scholes model.

The options can be divided into five moneyness categories based on the current underlying asset price S_t:

(i) Deep out-of-the-money (Deep OTM): be more than 1.5σ out of the money.

(ii) Out-of-the-money (OTM): be within $0.5\sigma - 1.5\sigma$ out of the money.

(iii) At-the-money (ATM): be within 0.5σ for both directions.

(iv) In-the-money (ITM): be within $0.5\sigma - 1.5\sigma$ in the money.

(v) Deep in-the-money (Deep ITM): be more than 1.5σ in the money.

3.4 Risk-Free rate

The London InterBank Offered Rate (LIBOR) in USD provided by global-rates.com is used. At each quarterly time t, let $r_{t,30}, r_{t,60}$ and $r_{t,90}$ denote the monthly-average spot risk-free rates with 1 month, 2 months, and 3 months maturities respectively.

4. METHODOLOGY

4.1 Option Pricing with GEV Distribution (Benchmark model)

We use the method from Section **2.3.3**. The model parameters can be obtained by solving the optimization in (2.25), while the estimated call and put prices are given by (2.18) and (2.20).

4.2 Option Pricing with Hybrid Pareto Distribution (Proposed Model)

Similar to the GEV model, negative returns are assumed to follow the hybrid Pareto distribution in (2.34). Hence, the density function of L_T becomes:

$$f(L_T \mid \xi, \eta, \beta) = \begin{cases} \dfrac{1}{\gamma} h(L_T|\eta, \beta), & L_T \leq \alpha \\ \dfrac{1}{\gamma} g(L_T - \alpha \mid \xi, \sigma), & L_T > \alpha \end{cases}, \tag{4.1}$$

where $h(\cdot \,|\eta, \beta)$ is the density function of normal distribution and $g(\cdot \mid \xi, \sigma)$ is the density function of generalized Pareto distribution.

Using the same technique in (2.10), the density function of an underlying asset $j(\cdot)$ is given by:

$$j(S_T) = \frac{1}{S_t} f(L_T \mid \xi, \eta, \beta) = \begin{cases} \dfrac{1}{\gamma S_t} h(L_T \mid \eta, \beta), & S_T \geq S_t(1-\alpha) \\ \dfrac{1}{\gamma S_t} g(L_T - \alpha | \xi, \sigma), & S_T < S_t(1-\alpha) \end{cases}. \tag{4.2}$$

Applying the no-arbitrage condition for a European call option from (2.1), the pricing formula becomes:

$$C_t(T, K) = e^{-r(T-t)} \int_K^\infty (S_T - K)\, j(S_T) dS_T. \tag{4.3}$$

Because the integral depends on whether K is below or above a threshold $S_t(1-\alpha)$, then the limits of the integral in equation (4.3) can be written with min , max functions as follows:

$$C_t(T, K) = \frac{1}{\gamma} e^{-r(T-t)}$$
$$\times \left(\left(\int_{\min(S_t(1-\alpha),K)}^{S_t(1-\alpha)} (S_T - K) \frac{1}{S_t} g(L_T - \alpha) dS_T \right) + \left(\int_{\max(S_t(1-\alpha),K)}^{\infty} (S_T - K) \frac{1}{S_t} h(L_T) dS_T \right) \right). \tag{4.4}$$

Let A denote the first integral. Again, we consider the case where $\xi > 0$:

$$A = \int_{\min(S_t(1-\alpha),K)}^{S_t(1-\alpha)} (S_T - K) \frac{1}{S_t} g(L_T - \alpha) dS_T$$
$$= \int_{\min(S_t(1-\alpha),K)}^{S_t(1-\alpha)} (S_T - K) \frac{1}{S_t \sigma} \left(1 + \frac{\xi}{\sigma}(L_T - \alpha) \right)^{-\frac{1}{\xi}-1} dS_T. \tag{4.5}$$

Consider the change of variable:

$$y = 1 + \frac{\xi}{\sigma}(L_T - \alpha) = 1 + \frac{\xi}{\sigma}\left(1 - \frac{S_T}{S_t} - \alpha\right), \tag{4.6}$$

(4.5) becomes:

$$A=-\int_{\max\left(1,1+\frac{\xi}{\sigma}\left(1-\frac{K}{S_t}-\alpha\right)\right)}^{1}\left(S_t\left(1-\alpha-\frac{\sigma}{\xi}(y-1)\right)-K\right)\frac{1}{S_t\sigma}(y^{-\frac{1}{\xi}-1})\frac{S_t\sigma}{\xi}dy. \qquad (4.7)$$

After grouping the terms, (4.7) becomes:

$$A=\frac{1}{\xi}\left(\frac{S_t\sigma}{\xi}\int_{\max\left(1,1+\frac{\xi}{\sigma}\left(1-\frac{K}{S_t}-\alpha\right)\right)}^{1} y\left(y^{-\frac{1}{\xi}-1}\right)dy-\left(S_t\left(1-\alpha+\frac{\sigma}{\xi}\right)-K\right)\int_{\max\left(1,1+\frac{\xi}{\sigma}\left(1-\frac{K}{S_t}-\alpha\right)\right)}^{1} y^{-\frac{1}{\xi}-1}dy\right). \qquad (4.8)$$

Let $d_1=\max\left(1,1+\frac{\xi}{\sigma}\left(1-\frac{K}{S_t}-\alpha\right)\right)$. The solution of integral in (4.8) becomes:

$$\begin{aligned} A&=\frac{1}{\xi}\left[\frac{S_t\sigma}{(\xi-1)}\left(1-d_1^{-\frac{1}{\xi}+1}\right)-\left(S_t\left(1-\alpha+\frac{\sigma}{\xi}\right)-K\right)(-\xi)\left(1-d_1^{-\frac{1}{\xi}}\right)\right] \\ &=\frac{S_t\sigma}{\xi(\xi-1)}\left(1-d_1^{-\frac{1}{\xi}+1}\right)+\left(S_t\left(1-\alpha+\frac{\sigma}{\xi}\right)-K\right)\left(1-d_1^{-\frac{1}{\xi}}\right). \qquad (4.9)\end{aligned}$$

Let B denote the second integral with the normal distribution:

$$\begin{aligned} B&=\int_{\max(S_t(1-\alpha),K)}^{\infty}(S_T-K)\frac{1}{S_t}f(L_T)dS_T \\ &=\int_{\max(S_t(1-\alpha),K)}^{\infty}(S_T-K)\frac{1}{S_t}\frac{1}{\sqrt{2\pi}\beta}\exp\left(-\frac{1}{2}\left(\frac{L_T-\eta}{\beta}\right)^2\right)dS_T \\ &=\int_{\max(S_t(1-\alpha),K)}^{\infty}\frac{S_T}{S_t\sqrt{2\pi}\beta}\exp\left(-\frac{1}{2}\left(\frac{L_T-\eta}{\beta}\right)^2\right)dS_T \\ &\quad -K\int_{\max(S_t(1-\alpha),K)}^{\infty}\frac{1}{S_t\sqrt{2\pi}\beta}\exp\left(-\frac{1}{2}\left(\frac{L_T-\eta}{\beta}\right)^2\right)dS_T. \qquad (4.10)\end{aligned}$$

Consider the change of variable:

$$y=\frac{(L_T-\eta)}{\beta}=\frac{1-\frac{S_T}{S_t}-\eta}{\beta}. \qquad (4.11)$$

The equation in (4.11) becomes:

$$B=-\int_{\min\left(\frac{\alpha-\eta}{\beta},\frac{1-\frac{K}{S_t}-\eta}{\beta}\right)}^{-\infty}S_T\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{y^2}{2}\right)dy+K\int_{\min\left(\frac{\alpha-\eta}{\beta},\frac{1-\frac{K}{S_t}-\eta}{\beta}\right)}^{-\infty}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{y^2}{2}\right)dy$$

$$=-S_t\int_{\min\left(\frac{\alpha-\eta}{\beta},\frac{1-\frac{K}{S_t}-\eta}{\beta}\right)}^{-\infty}(1-\eta-y\beta)\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{y^2}{2}\right)dy-K\mathcal{N}\left(\min\left(\frac{\alpha-\eta}{\beta},\frac{1-\frac{K}{S_t}-\eta}{\beta}\right)\right). \qquad (4.12)$$

Let $d_2=\min\left(\frac{\alpha-\eta}{\beta},\frac{1-\frac{K}{S_t}-\eta}{\beta}\right)$. Then (4.12) becomes:

$$\begin{aligned} B&=S_t(1-\eta)\mathcal{N}(d_2)-\frac{\beta S_t}{\sqrt{2\pi}}\int_{-\infty}^{d_2}y\exp\left(-\frac{y^2}{2}\right)dy-K\mathcal{N}(d_2) \\ &=S_t(1-\eta)\mathcal{N}(d_2)+\frac{\beta S_t}{\sqrt{2\pi}}\exp\left(-\frac{d_2^2}{2}\right)-K\mathcal{N}(d_2)\end{aligned}$$

$$= (S_t(1-\eta) - K)\mathcal{N}(d_2) + \frac{\beta S_t}{\sqrt{2\pi}} \exp\left(-\frac{d_2^2}{2}\right). \tag{4.13}$$

Where $\mathcal{N}(\cdot)$ is a cumulative density function (CDF) of a standard normal distribution. Therefore, the call option pricing formula under hybrid Pareto distribution becomes:

$$C_t\big(T, K \,\big|\, \xi, \eta, \beta\big) = \frac{1}{\gamma} e^{-r(T-t)}[A + B], \tag{4.14}$$

where:

$$A = \frac{S_t \sigma}{\xi(\xi - 1)}\left(1 - d_1^{-\frac{1}{\xi}+1}\right) + \left(S_t\left(1 - \alpha + \frac{\sigma}{\xi}\right) - K\right)\left(1 - d_1^{-\frac{1}{\xi}}\right),$$

$$B = (S_t(1-\eta) - K)\mathcal{N}(d_2) + \frac{\beta S_t}{\sqrt{2\pi}} \exp\left(-\frac{d_2^2}{2}\right),$$

$$d_1 = \max\left(1, 1 + \frac{\xi}{\sigma}\left(1 - \frac{K}{S_t} - \alpha\right)\right),$$

$$d_2 = \min\left(\frac{\alpha - \eta}{\beta}, \frac{1 - \frac{K}{S_t} - \eta}{\beta}\right),$$

for the continuous condition (2.32), (2.33) and (2.35) to hold. Note that this model needs only three parameters (ξ, η, β) to be estimated.

4.3 Model Estimation

Quarterly period data of option prices from December 2001 to December 2015 gives us 60 data points. Let $t \in \{1,2, \dots ,60\}$ be the time-period. Denote $\Theta_{1,t.30} = \{\xi_{1,t,30}, \sigma_{t,30}\}$ as the parameter set of the GEV model estimated by 30 days to expiration options at time t and $\Theta_{2,t,30} = \{\xi_{2,t,30}, \eta_{t,30}, \beta_{t,30}\}$ as that of the hybrid Pareto model. At each time t, we calibrate the model parameters by the observed prices with 30, 60, and 90 days to expiration respectively. By minimizing the sum square errors (SSE_t), $\Theta_{i,t}$ can be obtained by this optimization:

$$SSE_t = \min_{\Theta_t}\left\{\left(\sum_{i=1}^{N}\left(\hat{C}_{i,t}\big(T, K \,\big|\, \Theta_t\big) - C_{i,t}(T, K)\right)^2\right) + \left(\sum_{j=1}^{M}\left(\hat{P}_{j,t}\big(T, K \,\big|\, \Theta_t\big) - P_{j,t}(T, K)\right)^2\right)\right\}. \tag{4.15}$$

This optimization is archived by the non-linear least square method from the MATLAB optimization toolbox. If there are no options matched in that time horizons, the nearest date is used instead. Note that, T is a time-to-maturity in years.

4.4 In-Sample Pricing

For the in-sample pricing, we use all the data provided at period t to estimate $\Theta_{1,t}$, and $\Theta_{2,t}$. Then use them to compute the estimated call and put option prices.

4.5 Model Performance

The model performance can be examined in three values:

4.5.1 The Root Mean Square Error (RMSE)

$$RMSE_t = \sqrt{\frac{SSE_t}{\tilde{N}}}, \tag{4.16}$$

where SSE_t is the sum square errors computed at time t and is given by (4.15).

This value measures how far estimated prices deviate from the market value. The less this number is, the better the model fits to the data. $\tilde{N}$ here is the number of observed prices at period t.

4.5.2 Pricing Bias

$$\text{Price Bias} = \text{Market Price} - \text{Estimated Price} \tag{4.17}$$

Pricing biases report the error in a unit of USD. If it turns out to be a negative value, then it implies that our model overprices the option and vice versa.

4.5.3 Percentage of Pricing Error

$$\text{Percentage of Pricing Error} = \frac{|\text{Market Price} - \text{Estimated Price}|}{\text{Market Price}} \tag{4.18}$$

This value reports how much the pricing error is comparing to its market price.

5. RESULTS

5.1 Observed Data

Table 2: Observed Data

Categories	Number of observations		
	calls	**puts**	**total**
Time to Maturity			
30 days	2,673	4,151	**6,824**
60 days	1,684	2,663	**4,347**
90 days	1,641	2,279	**3,920**
total	**5,998**	**9,093**	**15,091**
Moneyness			
Deep ITM	1,726	1,020	**2,746**
ITM	436	214	**650**
ATM	397	374	**771**
OTM	307	539	**846**
Deep OTM	3,132	6,946	**10,078**
total	**5,998**	**9,093**	**15,091**

Table 2 above summarizes the total numbers of options used in this study, which are selected by the criteria described in Section 3. In addition, we use mid-prices as representatives of option prices. Overall, 60.25% of data are puts, and the rest 39.75% are calls. We can see that the number of data points for the put options is about 1.5 times more than that for the call options. One possible reason here is that people use the puts as protection for their investment. The 30-day options account for 45.21% of the entire data set. In terms of the moneyness, 66.78% of observed data are in the deep OTM category. In contrast, ITM options have the least numbers of observation, which is 4.30%. The average number of options in each period is 251. In comparison, the maximum and minimum numbers are 546 and 93, respectively.

5.2 In-sample Pricing Performance

The in-sample model performances can be analyzed in four aspects:

5.2.1 Time -to-maturity

Table 3: In-sample pricing performance (time-to-maturity)

Time-to-expiration		30 days		60 days		90 days		All	
Model		GEV	HP	GEV	HP	GEV	HP	GEV	HP
Avg. RMSE	Calls	2.07	1.04	1.58	1.81	0.88	1.75	1.51	1.53
	Puts	1.77	0.86	0.99	1.17	0.81	1.46	1.19	1.16
	All	**1.90**	**0.94**	**1.32**	**1.53**	**0.85**	**1.60**	**1.36**	**1.36**
Avg. Pricing Error		37.05%	31.74%	16.49%	21.97%	16.46%	23.51%	25.06%	26.10%
Avg. Absolute Pricing Biases		1.59	0.71	0.92	0.95	0.86	1.13	1.16	0.90

Table 3 displays two model performances of each time horizon. The performances evaluated by RMSE of the two models is about the same, but the overall average pricing bias from the HP model is 22.41% lower compared to the GEV model. In addition, the Hybrid Pareto pricing model (HP) outperforms the Generalized Extreme Value pricing model (GEV) by 50.52% in a 30-day to expiration. While the HP model has an average RMSE of 0.94 USD, GEV model's RMSE is approximately two times bigger, at 1.90 USD. However, the GEV model performance improves when the time to maturity increases. The HP model performs poorly with a longer time to maturity. At a 90-day horizon, the GEV model produces the lowest RMSE and pricing bias at 0.85 and 0.86 USD. The HP model's fitting performance decreases by about 70.21% to 1.60 USD. Because the short time-to-maturity option contains less time value, then its density clusters around the bulk part of its distribution. Therefore, the HP model with a taller body has an advantage in this time horizon. Another observation is that both models produce smaller error in put options than in the calls. This might be caused by the huge differences between calls and puts in the observed data. The percentage errors in 30-day to expiration is high compared the others.

5.2.2 Moneyness

Table 4: In-sample pricing performance (moneyness)

Moneyness		Deep ITM		ITM		ATM		OTM		Deep OTM	
Model		GEV	HP	GEV	HP	GEV	HP	GEV	HP	GEV	HP
Avg. RMSE	Call	2.78	1.30	1.59	0.74	1.34	0.80	1.35	1.25	1.10	1.78
	Put	2.87	1.46	1.85	0.98	1.44	0.59	1.46	0.85	0.97	1.23
	All	2.83	1.38	1.72	0.86	1.39	0.70	1.40	1.05	1.04	1.50
Avg. Pricing Error		3.56%	1.24%	4.11%	1.69%	3.45%	1.75%	39.52%	45.88%	35.77%	38.27%
Avg. Absolute Pricing Biases		2.54	0.90	2.36	1.08	1.18	0.65	0.95	1.26	0.76	0.90

Table 4 reports the performance of two models for five moneyness categories. The HP model outperforms the GEV model in all moneyness except the deep OTM options. Besides, it also produces lower pricing biases by 64.57%, 54.23% and 44.92% in the Deep ITM, ITM and ATM options. While the GEV model shows less pricing biases by 24.60% and 15.55% in the OTM and deep OTM options compared to the HP model. The average RMSE of the HP model decreases when options are at-the-money. On the other hand, the GEV model performs well when the options move out of the money. It shows 30.67% less RMSE compared to the HP model. This implies that the GEV model performs better with extreme samples than a normal one. However, we can see the high percentage of pricing errors in the OTM and deep OTM options. The explanation is that options in this type typically trade at low prices. When we compute the error compared to their market prices, it turns out to be huge numbers. While in three other cases, all of them have less than five percent pricing error. (see more details in Section 5.2.4)

5.2.3 Pricing Biases

(i) Average calls price biases in terms of moneyness

Figure 1: Pricing biases of 30 days to expiration in-sample call options

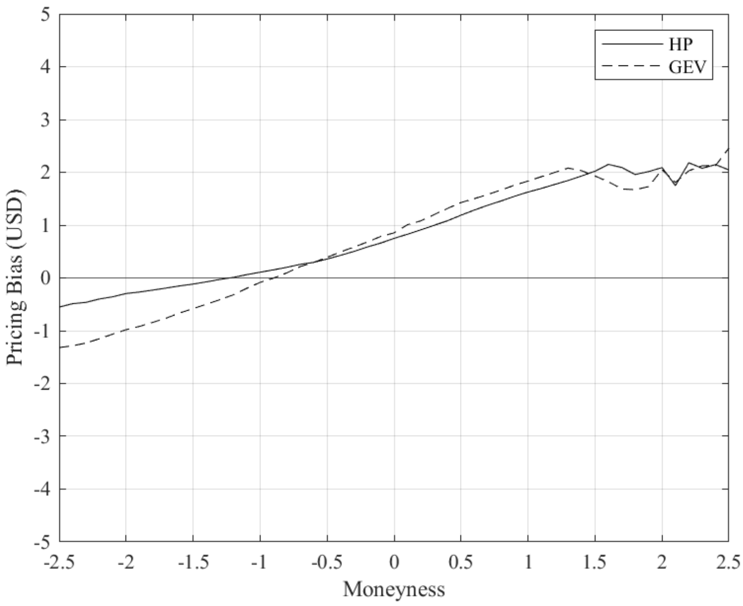


Figure 2: Pricing biases of 60 days to expiration in-sample call options

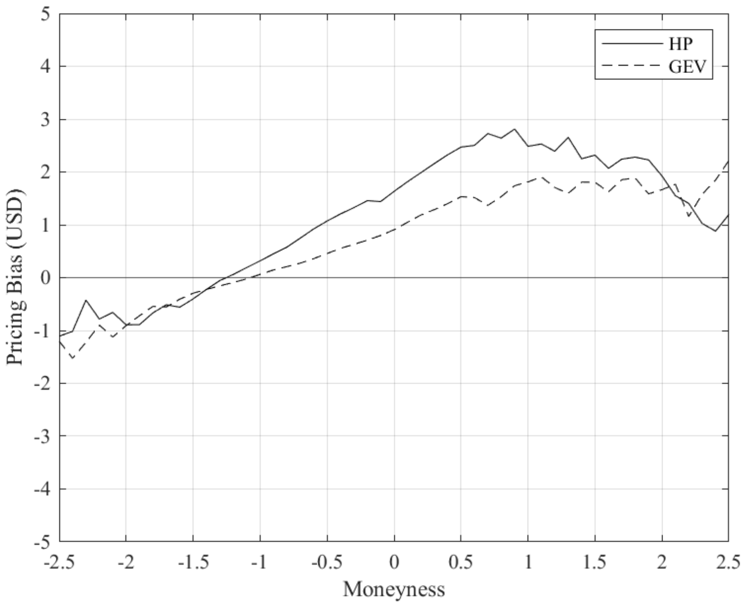
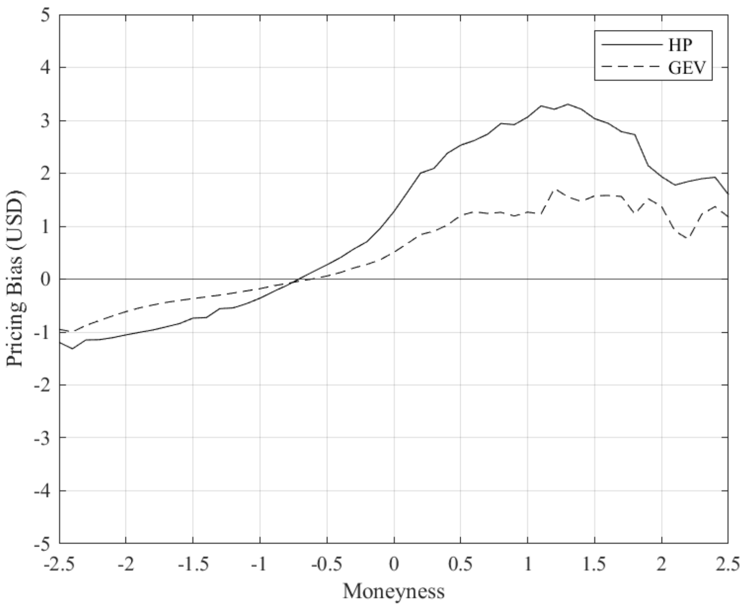


Figure 3: Pricing biases of 90 days to expiration in-sample call options



Figures 1-3 show pricing errors in terms of pricing biases defined in (4.17). The graphs above are derived by the same technique suggested by Markose and Alenton (2011). That is, for each observations date, we use a Curve Fitting Toolbox form MATLAB in order to fit a spline to pricing biases as a function of moneyness on that date. In the end, we take an average value of these 60 splines across moneyness between -2.5 to $+2.5$. The results show that both models overprice ITM call options and underprice OTM call options. Besides, the magnitude of biases in the deep OTM side is relatively small compared to the deep ITM. This evidence supports the idea that the fat-tailed model favors extreme outcomes.

Another observation is that the pricing biases are larger when moneyness moves away from around -1 or when the time horizon increases. Nevertheless the GEV model appears to be less dependent on the time horizon. The pricing biases of the HP model are between $+3$ USD and -1 USD. For the GEV model, they are between $+2$ USD and -1 USD.

(ii) Average puts price biases in terms of moneyness

Figure 4: Pricing biases of 30 days to expiration in-sample put options

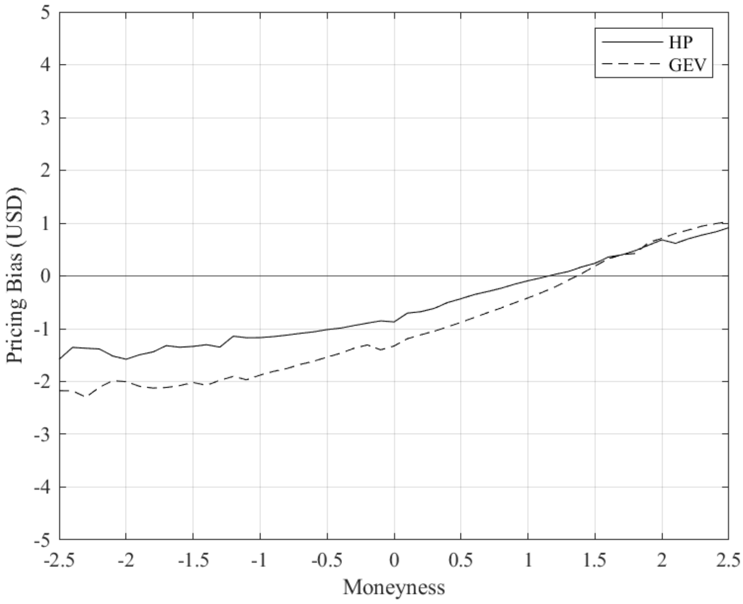


Figure 5: Pricing biases of 60 days to expiration in-sample put options

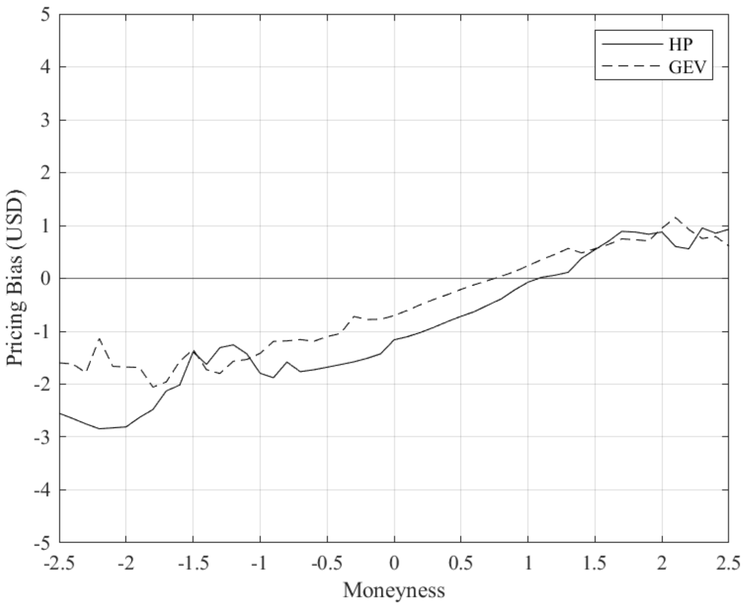
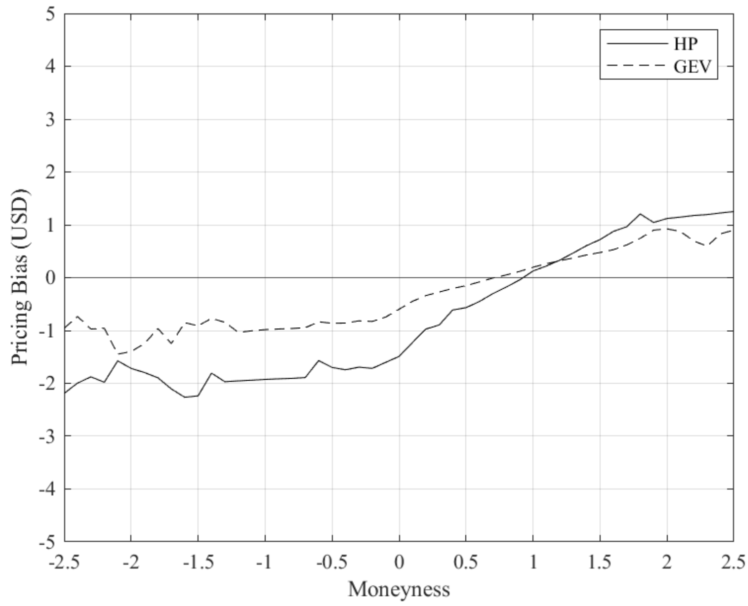


Figure 6: Pricing biases of 90 days to expiration in-sample put options



Figures 4-6 show pricing errors in terms of pricing bias of put options. As we can see, both models overprice OTM put options and underprice ITM put options. Again, the pricing biases are larger when the moneyness moves away from around 1 to 1.5. The pricing biases of the HP model are between around $+1$ USD and -3 USD. For the GEV model, they are between $+1$ USD and -2 USD. The magnitude of biases in the deep OTM side are smaller than the deep ITM. This confirms previous finding in the call option.

5.2.4 Pricing Error and Option Prices

Figure 7: Log of estimated price and market price from the GEV model

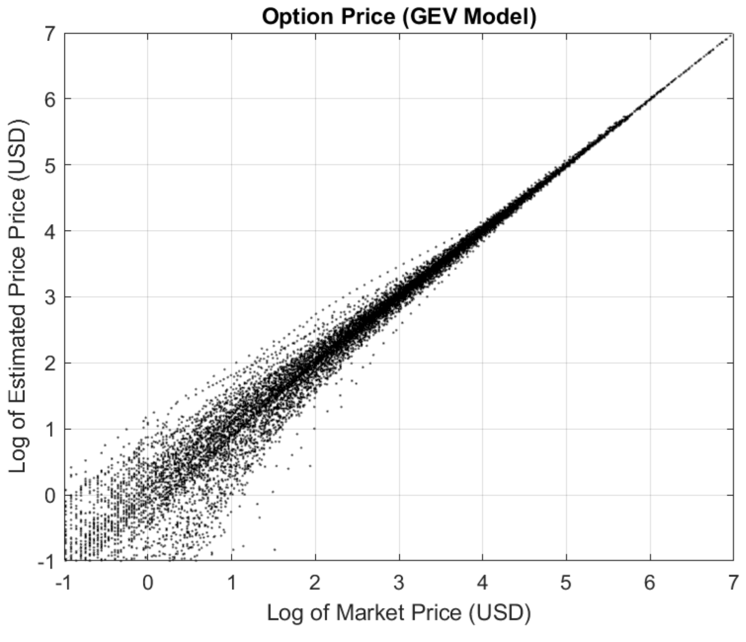


Figure 8: Log of estimated price and market price from the HP model

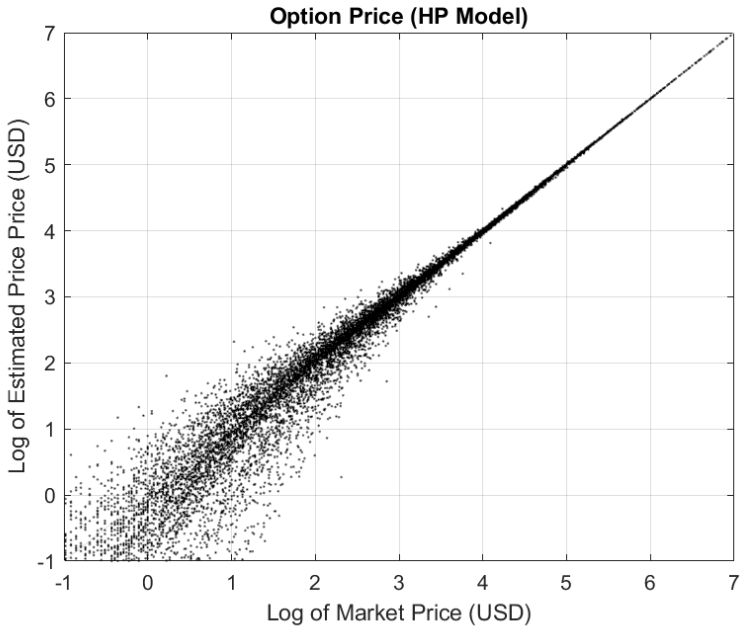


Table 5: Pricing errors and option prices

	Low		**Mid**		**Large**		**Very Large**	
Price Range (USD)	0.025-1		1-10		10-50		> 50	
Number of data	2718		4676		4811		2761	
Model	**GEV**	**HP**	**GEV**	**HP**	**GEV**	**HP**	**GEV**	**HP**
Avg. Pricing Error (%)	74.51%	75.12%	27.92%	32.12%	6.22%	5.75%	2.37%	1.07%
Avg. Absolute Pricing Biases	0.26	0.27	0.89	1.05	1.34	1.08	2.22	0.94

Figure 7-8 display scatter plots of the natural log of estimated prices and market prices. We can see that the low-priced options plots are more dispersed compare to the high-priced. Also, Table 5 reports the average percentage of pricing error defined in (4.18). We divide our 60-period observed data by their price into four groups and compare them across all time-to-maturity and moneyness. Both models show that the bigger price the options are, the lower the pricing error is. When the models try to fit the low-priced option, most of the time, they come up with zero USD instead. For example, 557 estimated prices (20.49%) from the low-priced have less than 0.001 USD. Consequently, this makes the percentage of pricing error go higher. On the other hand the HP provides less pricing biases by 19.40% and 57.66% for the large and very large option prices compared to the GEV model. Because the high-priced options are usually in the in-the-money, this evidence supports results from Table 4 that the HP model prefers ITM options to OTM options.

5.3 Implied Risk-Neutral Density During the Crisis Event

5.3.1 Implied Risk-Neutral Density in 2005 and 2008

Figure 9: 30 days implied risk-neutral density in June 2005 and 2008

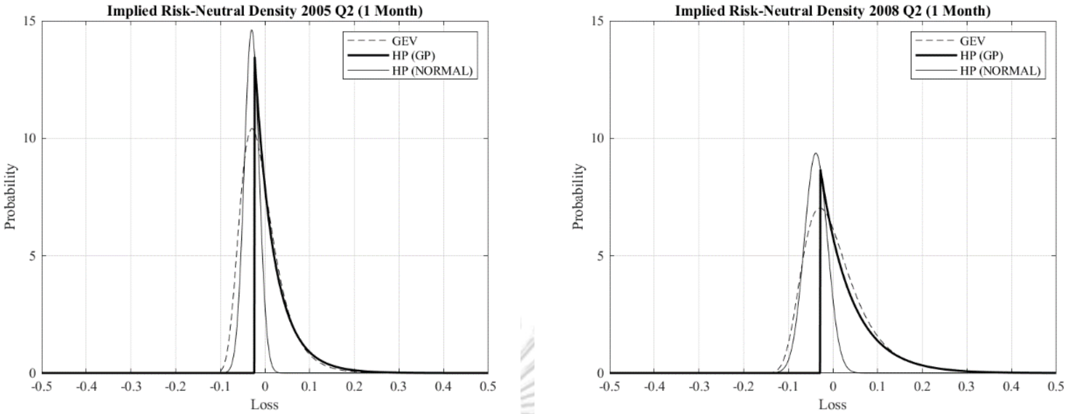


Figure 9 displays one-year implied risk-neutral density of the loss (negative returns) of the S&P 500 index estimated in June 2005 and 2008. In 2005, a stable period, the GEV distribution exhibits a fatter tail on the left-hand side comparing to the HP model. In contrast, the density of the HP model is much taller in the body part. On the right-hand side, the HP model exhibits a fatter tail than that of the GEV model, but the difference between them is relatively small. This is because the tail shape parameters of these two models are similar in terms of the density function.

In 2008, during the subprime crisis, the body parts of the two models decrease in their sizes. Both tails are also thicker than those in the stable period. This result implies that, when a crisis occurs, it is less likely to observe the usual outcome than in the stable period.

5.3.2 Model Performance in 2005 and 2008

Table 6: Model performances in terms of average RMSE in 2005 and 2008

	2005		2008	
	GEV	HP	GEV	HP
Calls	1.11	1.03	1.50	2.41
Puts	0.90	0.73	1.62	2.63
All	1.00	0.88	1.56	2.52

Table 6 reports the performance comparison between two models in 2005, stable period, and 2008, crisis period. The HP model outperforms the GEV model in the stable period, by 13.63% However, in the subprime crisis period in 2008, the GEV model has less pricing error, which is 38.09% lower than the HP model. Since the two models' tails are about the same, the differences between their performances mainly come from the body part. As shown in Figure 7, the GEV model exhibits small density in the middle compared to the HP model. This implies that the GEV model provides more possibilities for the return to deviate from its mean. Consequently, during the crisis, the GEV model is preferred.

6. CONCLUSIONS

This paper develops a new European option pricing model based on the Extreme Value Theory. We use the Generalized Extreme Value (GEV) model by Markose and Alenton (2011) as a benchmark model. Then, we adopt the Hybrid Pareto (HP) model by Carreau and Bengio (2009) as our proposed model. We derive closed-form pricing formulas for call and put options with an assumption that the tail shape parameter (ξ) is greater than zero (the Frechét type). Using the optimization method, in each period, we estimate the model parameters by minimizing the root-mean-squared error (RMSE). In the end, we compare the results in two aspects, time-to-maturity and moneyness.

According to the results in Section 5, the HP model clearly outperforms the GEV model in a short time horizon, a 30-day to expiration option. This might be because option with a short time to maturity contains less time value and less uncertainty. Hence, if it is a deep OTM option, its price will be a small number. If there is mispricing from a model, an error possibly be s small number too. In the deep ITM case, its price mainly comes from the intrinsic value and will not move much till the expiration date. Therefore, the best model could be decided by which one is good at pricing ATM options. As displayed in Figure 9, our result shows that the HP distribution always exhibits a taller body part comparing with the GEV distribution. This result implies that the HP model provides more possibilities that options will be strict at their current prices. Moreover, our fitting result supports the conclusion above since it shows that the HP model is the best fit for ATM options. Besides, the HP model provides lower absolute pricing biases by 22.41% compared to the GEV model across all time-to maturity and moneyness, although the fitting performances from two models evaluated by RMSE are about the same. However, our proposed model did not perform well with the low-priced options (deep-out-of-the-money options). Having a special case of pricing formula might solve this issue.

In the crisis period, the GEV model still has an advantage. The reason that the HP model is outperformed by the GEV model can be described by the shape of its distribution explained above. When an extreme event occurs, there is less probability that we see the normal outcome.

For further studies, to improve the performance of the Hybrid Pareto model, one may drop out an assumption that the tail shape parameter (ξ) is greater than zero. As a result, it allows the model to be either thin-tailed or normal-tailed. Also, this provides more flexibility to the model to be well-performed in any economic scenario. Consequently, there are two pricing formulas to be derived. The first one is the pricing formula when the tail shape parameter is equal to zero. The second is when it is less than zeros. Furthermore, we may study more details about when we should use a fat, thin, or normal tail.

REFERENCES

Abadir, K. M., & Rockinger, M. (1997). *Density-embedding functions*: University of York, Department of Economics and Related Studies.

Aït-Sahalia, Y., & Lo, A. W. (1998). Nonparametric estimation of state-price densities implicit in financial asset prices. *The journal of finance, 53*(2), 499-547.

Aparicio, S., & Hodges, S. (1998). Implied risk-neutral distribution: A comparison of estimation methods. *FORC Preprint, University of Warwick*.

Bookstaber, R. M., & McDonald, J. B. (1987). A general distribution for describing security price returns. *Journal of business*, 401-424.

Bu, R., & Hadri, K. (2007). Estimating option implied risk-neutral densities using spline and hypergeometric functions. *The Econometrics Journal, 10*(2), 216-244.

Carreau, J., & Bengio, Y. (2009). A hybrid Pareto model for asymmetric fat-tailed data: the univariate case. *Extremes, 12*(1), 53-76.

Davison, A. C., & Smith, R. L. (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society: Series B, 52*(3), 393-425.

Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of mathematics*, 423-453.

Jackwerth, J. C. (1999). Option-implied risk-neutral distributions and implied binomial trees: A literature review. *The Journal of Derivatives, 7*(2), 66-82.

Koedijk, K. G., Schafgans, M. M., & De Vries, C. G. (1990). The tail index of exchange rate returns. *Journal of international economics, 29*(1-2), 93-108.

LeBaron, B., & Samanta, R. (2005). Extreme value theory and fat tails in equity markets. *Available at SSRN 873656*.

Markose, S., & Alentorn, A. (2011). The generalized extreme value distribution, implied tail index, and option pricing. *The Journal of Derivatives, 18*(3), 35-60.

Melick, W. R., & Thomas, C. P. (1997). Recovering an asset's implied PDF from option prices: an application to crude oil during the Gulf crisis. *Journal of Financial Quantitative Analysis, 32*(1), 91-115.

Pickands III, J. (1975). Statistical inference using extreme order statistics. *the Annals of Statistics, 3*(1), 119-131.

Ritchey, R. J. (1990). Call option valuation for discrete normal mixtures. *Journal of Financial Research, 13*(4), 285-296.

Rocco, M. (2014). Extreme value theory in finance: A survey. *Journal of Economic Surveys, 28*(1), 82-108.

Santos, A., & Guerra, J. (2015). Implied risk neutral densities from option prices: Hypergeometric, spline, lognormal, and Edgeworth functions. *Journal of Futures Markets, 35*(7), 655-678.

Todorovic, P., & Zelenhasic, E. (1970). A stochastic model for flood analysis. *Water Resources Research, 6*(6), 1641-1648.

Vilasuso, J., & Katz, D. (2000). Estimates of the likelihood of extreme returns in international stock markets. *Journal of Applied Statistics, 27*(1), 119-130.

APPENDIX

A1. The formula for European put option pricing

Applying the no-arbitrage condition for a European put option, the pricing formula becomes:

$$P_t(T,K) = e^{-r(T-t)} \int_{-\infty}^{K} (K - S_T)\, j(S_T) dS_T. \qquad (A\ 1.1)$$

The integral depends on the value of K and a threshold $S_t(1-\alpha)$, then limits of the integral in equation (4.3) can be written with min, max functions as:

$$P_t(T,K) = \frac{1}{\gamma} e^{-r(T-t)}$$
$$\times \left[\left(\int_{-\infty}^{\min(S_t(1-\alpha),K)} (K - S_T) \frac{1}{S_t} g(L_T - \alpha) dS_T \right) + \left(\int_{\min(S_t(1-\alpha),K)}^{K} (K - S_T) \frac{1}{S_t} h(L_T) dS_T \right) \right] \qquad (A\ 1.2)$$

Let C denote the first integral:

$$C = \int_{-\infty}^{\min(S_t(1-\alpha),K)} (K - S_T) \frac{1}{S_t} g(L_T - \alpha) dS_T$$

$$C = \int_{-\infty}^{\min(S_t(1-\alpha),K)} (K - S_T) \frac{1}{S_t \sigma} \left(1 + \frac{\xi}{\sigma} (L_T - \alpha) \right)^{-\frac{1}{\xi} - 1} dS_T. \qquad (A\ 1.3)$$

Consider the change of variable:

$$y = 1 + \frac{\xi}{\sigma}(L_T - \alpha) = 1 + \frac{\xi}{\sigma}\left(1 - \frac{S_T}{S_t} - \alpha\right), \qquad (A\ 1.4)$$

$(A\ 1.4)$ becomes:

$$C = \int_{\max\left(1,1+\frac{\xi}{\sigma}\left(1-\frac{K}{S_t}-\alpha\right)\right)}^{\infty} \left(K - S_t \left(1 - \alpha - \frac{\sigma}{\xi}(y-1) \right) \right) \frac{1}{S_t \sigma} (y)^{-\frac{1}{\xi}-1} \left(\frac{S_t \sigma}{\xi} \right) dy. \qquad (A\ 1.5)$$

After grouping the terms, $(A\ 1.5)$ becomes:

$$C = \frac{1}{\xi} \left[\frac{S_t \sigma}{\xi} \int_{d_1}^{\infty} y \left(y^{-\frac{1}{\xi}-1} \right) dy + \left(K - S_t \left(1 - \alpha + \frac{\sigma}{\xi} \right) \right) \int_{d_1}^{\infty} y^{-\frac{1}{\xi}-1} dy \right]$$
$$= \frac{1}{\xi} \left[\frac{S_t \sigma}{(\xi - 1)} \left(0 - d_1^{-\frac{1}{\xi}+1} \right) + \left(K - S_t \left(1 - \alpha + \frac{\sigma}{\xi} \right) \right) (-\xi) \left(0 - d_1^{-\frac{1}{\xi}} \right) \right]$$
$$= \left(K - S_t \left(1 - \alpha + \frac{\sigma}{\xi} \right) \right) d_1^{-\frac{1}{\xi}} - \frac{S_t \sigma}{\xi(\xi - 1)} d_1^{-\frac{1}{\xi}+1}. \qquad (A\ 1.6)$$

Let D denote the second integral with the normal distribution:

$$\begin{aligned} D &= \int_{\min(S_t(1-\alpha),K)}^{K} (K - S_T)\frac{1}{S_t} h(L_T) dS_T \\ &= \int_{\min(S_t(1-\alpha),K)}^{K} (K - S_T)\frac{1}{S_t}\frac{1}{\sqrt{2\pi}\beta}\exp\left(-\frac{1}{2}\left(\frac{L_T-\eta}{\beta}\right)^2\right) dS_T \\ &= K\int_{\min(S_t(1-\alpha),K)}^{K} \frac{1}{S_t\sqrt{2\pi}\beta}\exp\left(-\frac{1}{2}\left(\frac{L_T-\eta}{\beta}\right)^2\right) dS_T \\ &- \int_{\min(S_t(1-\alpha),K)}^{K} \frac{S_T}{S_t\sqrt{2\pi}\beta}\exp\left(-\frac{1}{2}\left(\frac{L_T-\eta}{\beta}\right)^2\right) dS_T. \end{aligned} \quad (A\ 1.7)$$

Consider the change of variable:

$$y = \frac{(L_T - \eta)}{\beta} = \frac{1 - \frac{S_T}{S_t} - \eta}{\beta}, \quad (A\ 1.8)$$

Let $d_3 = \max\left(\frac{\alpha-\eta}{\beta}, \frac{1-\frac{K}{S_t}-\eta}{\beta}\right)$. Then the equation in $(A\ 1.7)$ becomes:

$$\begin{aligned} D &= S_t\int_{d_3}^{\frac{1-\frac{K}{S_t}-\eta}{\beta}} (1-\eta-y\beta)\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{y^2}{2}\right)dy - K\int_{d_3}^{\frac{1-\frac{K}{S_t}-\eta}{\beta}}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{y^2}{2}\right)dy \\ &= S_t(1-\eta)\left(\mathcal{N}\left(\frac{1-\frac{K}{S_t}-\eta}{\beta}\right) - \mathcal{N}(d_3)\right) - \frac{\beta S_t}{\sqrt{2\pi}}\int_{d_3}^{\frac{1-\frac{K}{S_t}-\eta}{\beta}} y\exp\left(-\frac{y^2}{2}\right)dy \\ &\quad - K\left(\mathcal{N}\left(\frac{1-\frac{K}{S_t}-\eta}{\beta}\right) - \mathcal{N}(d_3)\right) \\ &= \left(\mathcal{N}\left(\frac{1-\frac{K}{S_t}-\eta}{\beta}\right) - \mathcal{N}(d_3)\right)(S_t(1-\eta)-K) - \frac{\beta S_t}{\sqrt{2\pi}}\left(e^{-\frac{d_3^2}{2}} - e^{-\frac{\left(1-\frac{K}{S_t}-\eta\right)^2}{2\beta^2}}\right). \end{aligned} \quad (A\ 1.9)$$

Therefore, the put option pricing formula under the hybrid Pareto distribution becomes:

$$P_t(T, K \mid \xi, \eta, \beta) = \frac{1}{\gamma}e^{-r(T-t)}[C + D], \quad (A\ 1.10)$$

where:

$$C = \left(K - S_t\left(1 - \alpha + \frac{\sigma}{\xi}\right)\right)d_1^{-\frac{1}{\xi}} - \frac{S_t\sigma}{\xi(\xi-1)}d_1^{-\frac{1}{\xi}+1},$$

$$D = \left(\mathcal{N}\left(\frac{1-\frac{K}{S_t}-\eta}{\beta}\right) - \mathcal{N}(d_3)\right)(S_t(1-\eta)-K) - \frac{\beta S_t}{\sqrt{2\pi}}\left(e^{-\frac{d_3^2}{2}} - e^{-\frac{\left(1-\frac{K}{S_t}-\eta\right)^2}{2\beta^2}}\right).$$

A2. Time Series of estimated parameters from the Generalized Extreme Value model

Figure 10: Estimated tail shape parameter from the GEV model

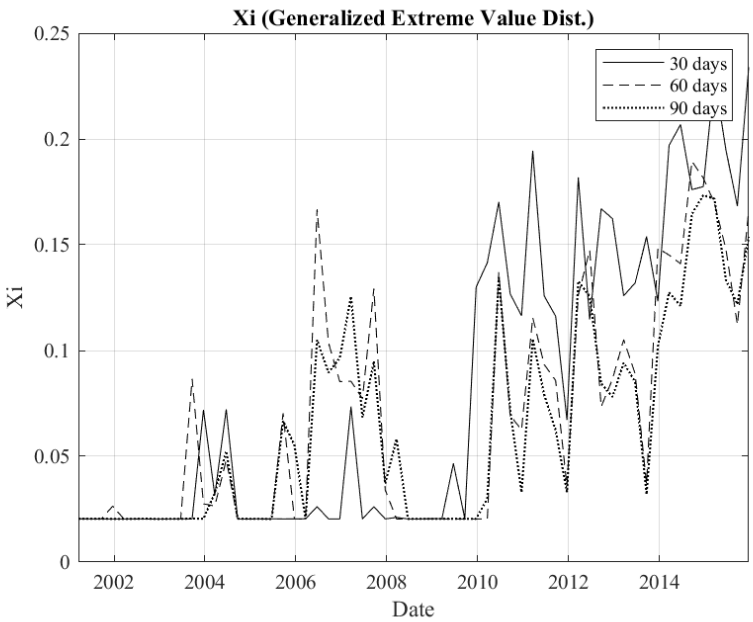
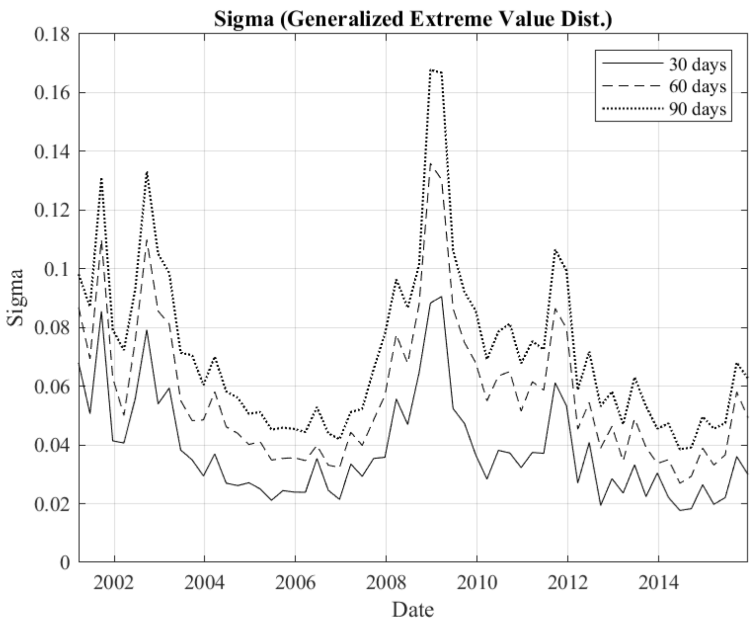


Figure 11: Estimated scale parameter from the GEV model



A3. Time series of estimated parameters from the hybrid Pareto model

Figure 12: Estimated tail shape parameter from the HP model

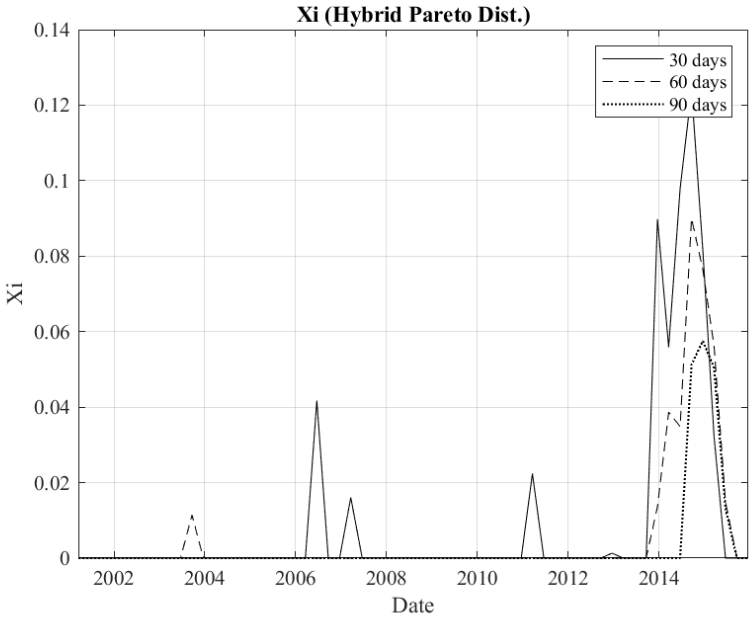


Figure 13: Estimated scale parameter from the HP model

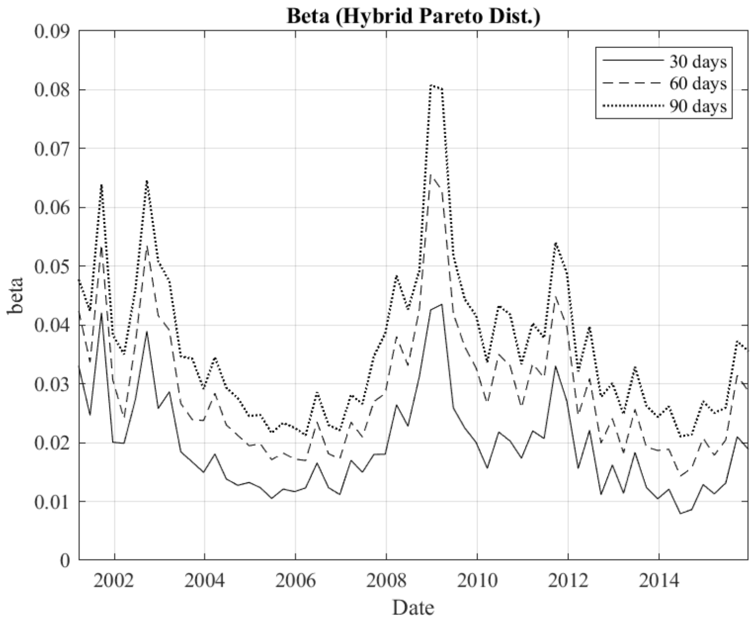
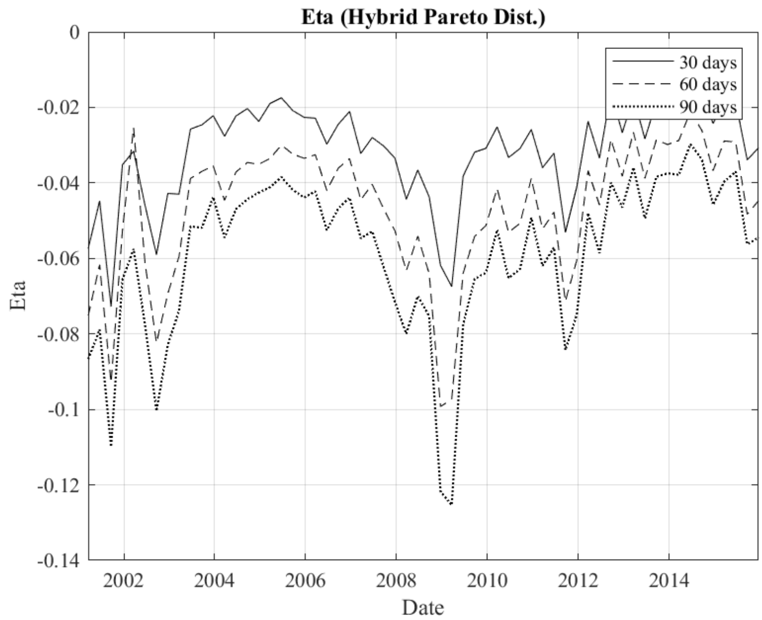
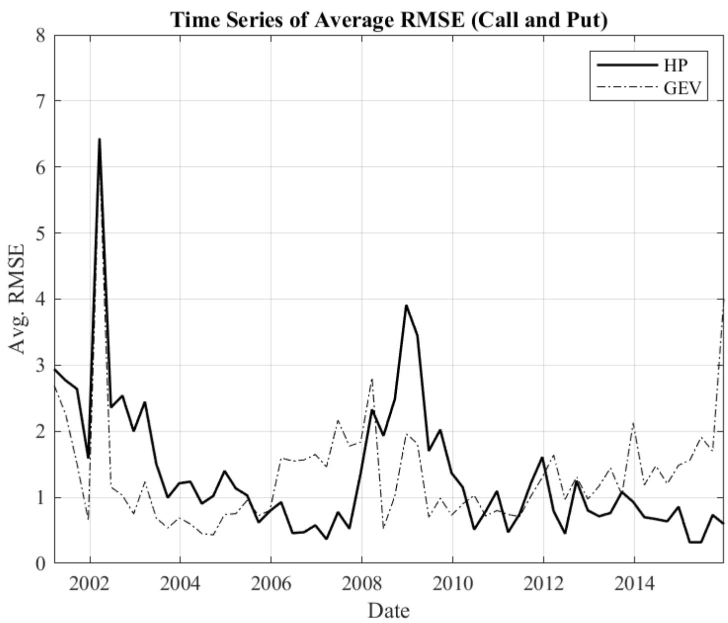


Figure 14: Estimated location parameter from the HP model



A4. Time series of average RMSE

Figure 15: Time series of average RMSE



VITA

NAME	Purin Luanloy
DATE OF BIRTH	25 June 1991
PLACE OF BIRTH	Rayong
HOME ADDRESS	32/45, Ivy Suthisan, Ratchadapisek 20, Huai Khwang, Samsen Nok, Bangkok, 10310