ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics
Bachelor Thesis Econometrics and Operational Research

Tail risk and the prediction of stock returns

Ingeborg de Pater, 411989
Supervisor: Francine Gresnigt
Second assessor: X. Xiao

July 2, 2017

Abstract

This paper compares the ability of three tail risk variables to forecast stock returns. One tail risk variable is calculated with the Hill estimator, one with the threshold exceeding method and one is calculated with the price of out-of-the-money European put options. The variables are calculated for two data samples, one with small companies and one with large companies. The forecasts are evaluated using in-sample predictive regressions and out-of-sample regressions with an expanding window. For both the large and the small companies, the variable constructed with the price of out-of-the-money European put options is the best in predicting future stock returns of the three tail risk variables. Some of the benchmark variables, however, perform equally well or even better.

Contents

1 Introduction

This article analyses the predictive power of tail risk measurements to forecast stock returns. Tail risk is the risk of an extreme low return on the stock market. The tail of the return distribution are all returns that fall below some negative threshold u. Kelly and Jiang (2014), found that their tail risk measurement was able to predict future stock returns. Therefore, we suspect that tail risk variables might have predictive power in forecasting stock returns. This paper will look at three different methods to measure tail risk and compare the forecast performance of the variables found.

In the literature, multiple tail risk measurements have been developed and their influence on the stock returns has been investigated. The variable and the methods we will focus on in this paper are from Kelly and Jiang (2014). The authors found a new tail risk variable that estimates common fluctuations across the tail of stock returns based on historical observations of large, negative returns. The main assumption underlying this variable is that the the tail of the return distribution follows a power law. The shape parameter of this power law, determining the shape of the tail, will be the first tail risk variable in this paper. The second tail risk variable is a similar one, used by, among others, Gilli and Këllizi (2006). They assumed that the tail of the return distribution follows a general Pareto distribution, and calculated the shape parameter of this general Pareto distribution with maximum likelihood. This shape parameter is the second tail risk variable. The last tail risk variable is developed by Bollerslev and Todorov (2011). Instead of the realized tail risk, it measures investors' expectations about the tail using option prices. This variable is constructed by comparing the prices of out-of-the-money put options with the future price of an asset. The goal of this paper is to see which of the three tail risk variables is the best predictor of the returns on the stock market. The research question is therefore:

Which tail risk variable is the best predictor of returns on the stock market?

I will evaluate the ability of the variables to predict future stock returns with two different methods. The first method are in-sample, predictive regressions, where the future stock return is regressed on a constant and the tail risk variable. Secondly, the predictive power is analysed with out-of-sample regressions with an expanding window, again with the future stock return on a constant and the tail risk variable. At last, this paper evaluates the predictive power to forecast stock returns of multiple benchmark variables, not related to tail risk, with the same methods and compares the results to the three tail risk variables.

The ability of a variable to predict stock returns depends on, among other factors, the size of the companies in the data sample. Therefore, there are two data samples used in this paper. The first data sample contains the assets that are part of the S&P-500 index and the options on the S&P-500 index. The S&P-500 index contains only large companies with respect to their market capitalization. The second data sample contains the assets that are part of, and the options on the Russell-2000 index. This index contains only small companies.

Kelly and Jiang (2014) argued that a higher tail risk will cause on average higher returns; investors usually do not like risk, and they want to be compensated for high risk by a higher expected return. Thus, if the tail risk in the market is high, the average return should be higher as well. The authors found that this is indeed the case; an increase in their tail risk variable forecasts a higher stock return. The hypothesis is therefore that this paper will find the same effect; an increase in the tail risk should have a positive effect on future stock returns.

Returns on the stock market are very important nowadays, in particular for investors and banks who want to maximize their profit on the stock market while minimizing their risk. But stock prices also have a great influence on the whole society, for example on the interest of a savings account, on retirement funds and during a crisis. Some literature focused on variables that might predict the stock returns. For example, Welch and Goyal (2007) examined the forecast power of a large number of variables suspected to predict stock returns, and showed that finding variables that are consistent in predicting stock returns is still a difficult task. Therefore, it is relevant to look for variables that might be better in predicting stock returns.

The conclusion of this paper is that for our data samples, the tail risk variable of Bollerslev and Todorov (2011), calculated with the prices of out-of-the-money put options, is the best in predicting future stock returns compared to the other two tail risk variables. This is the case for both the large companies in the S&P-500 index and the small companies in the Russell-2000 index. Some of the benchmark variables, not related to tail risk, are equally well or even better in predicting the stock returns than our three tail risk variables, but this depends on the data sample used and the forecast horizon.

In the rest of this paper, I will first discuss previous literature about tail risk on stock markets in section 2. Then, the methods used to calculate the variables and to compare the ability to forecast stock returns are discussed in section 3. In section 4, I will discuss the used data, and continue with the results in section 5. Finally, I will perform a robustness check in section 6 and the final conclusion is in section 7.

2 Literature review

Previous literature has investigated the tail risk on the stock market and the influence of this tail risk on the stock prices. Since observations in the tail are, per definition, rare, special methods are developed to estimate tail risk. There are various strategies to estimate tail risk. One strategy tries to estimate the tail risk with downside shocks in macro-economic variables. Rietz (1988) was one of the first authors using this method, with his hypothesis that the relatively high returns on assets compared to the returns on government bonds can be explained by the risk of market crashes. His model was extended by Barro (2006), who calculated the probability of economic disaster from historical observations. Multiple authors, including Wachter (2013) and Gabaix (2012), in turn extended the model of Barro (2006) to a dynamic model where the tail risk changes over time. The tail risk variables in this paper are different from the models in those papers, since I will not use macro-economic variables.

Other articles try to estimate the tail risk from option prices. The idea is to link option prices with (future) stock prices through calculating the risk premia for tail risk (i.e. the compensation required by investors to hold assets with a certain tail risk). For example, Pan (2002) calculated the tail risk premium demanded by investors, and found that this premium is important in explaining both asset prices and option prices. Eraker (2004) developed a model for option prices including possible jumps, and found that his model was reasonably good in forecasting option prices. Another author working on this topic is Bates, who included tail risk in models to estimate American option prices (Bates, 1996) and future prices (Bates, 2000).

This paper contains one tail risk variable from this strategy, developed by Bollerslev and Todorov (2011). This estimator is calculated with the prices of out-of-the-money European put options that are close to maturity. European put options give you the right to sell an asset for a certain strike price at maturity time. If a put option is out-of-the-money, it means that the current price of the asset is higher than the strike price of the asset. If at the maturity time, the price of the asset is still higher than the strike price, the put option is worthless. In other words, only if the price of an asset falls, the out-of-the-money European put option has some value at maturity. The intuition behind this variable is therefore that when a put option is out-of-the-money and close to maturity, the put option will only be worth anything at maturity if the stock price suddenly falls, i.e. "jumps" downwards. Consequently, the price of the put option might give you some information about the tail risk, the risk of an extreme low return, expected by investors.

Bollerslev and Todorov have developed and analysed many variables related to variance and tail risk in stock returns using, among other data, option prices, and often found that those variables have a high predictive power to forecast stock returns. In Bollerslev and Todorov (2011), the article the tail risk variable in this paper comes from, the authors developed the "Investors fear index", an index for the compensation demanded by investors for their fears of market disasters. To develop this index, they used both intra day high-frequency data to measure expected tail events (an extreme negative return), and option prices to calculate the compensation required by investors for tail events. In another article from Bollerslev, Todorov, and Xu (2015), they found a new procedure to estimate the compensation for tail risk required by investors. Including this new tail risk measure in regressions to predict the future stock returns significantly improves the explanatory power. Another method to estimate a tail risk measurement from option prices is developed by Andersen, Fusari, and Todorov (2015). They conclude that their new tail risk variable captures information about the price of risk that was not captured in other known variables related to option prices before, and is therefore an important variable to use in forecasting stock prices.

All the literature above relied partly on models and assumptions about asset prices and option prices. However, there is also literature about tail risk in stock markets that does not use any models for option prices or asset prices. Articles following this strategy often assume that the tail of the return distribution follows a certain distribution, and then estimate the parameter(s) of this distribution. This is for example done by Gilli and Këllizi (2006), Këllezi and Gilli (2003), Brodin and Klüppelberg (2008) and Longin (2000). One method from this approach is the "threshold exceeding method", focusing on returns that exceed/fall below a certain threshold. This paper will use this method to estimate one tail risk variable. Other literature does not estimate the parameters of the tail distribution, but uses Hill's estimator (Hill, 1975), see, among others, McNeil and

Frey (2000), Huisman, Koedijk, Kool, and Palm (2001) and Kelly and Jiang (2014). This paper estimates one tail risk variable using Hill's estimator as well.

The main difference between the method with the Hill estimator and the threshold exceeding method is that the threshold exceeding method is parametric and the method with the Hill estimator is non-parametric, as explained by Rocco (2014). For the threshold exceeding method, the main assumption is that the tail of the return distribution of all assets in the data sample follows a general Pareto distribution. The idea of the method is to estimate the parameters that fit this general Pareto distribution the best; hence, this method is parametric. For the method with the Hill estimator, the underlying assumption is about the tail of individual assets; the tail of the return distribution of each individual asset follows a power law. However, I do not try to find the parameter that fits this power law the best. Since I do not fit a certain model to the tail, this method is non-parametric.

3 Methods

This section first describes the three methods used to calculate the tail risk variables. Then, it explains how the ability of the three tail risk variables to forecast future stock returns is compared.

3.1 Tail risk variables

3.1.1 Model 1: Hill's estimator (λ)

The first method to estimate a tail risk variable comes from Kelly and Jiang (2014). The rest of the methods in subsection 3.2 come from this article as well. Kelly and Jiang (2014) defined the tail of the return distribution as all the returns that fall under a certain (negative) threshold u_t. The authors made a few key assumptions about the tail of the return distribution. First of all, they assumed that every asset has a different tail, and that this tail changes over time. However, they also assumed that the tails of all assets change over time in the same way, driven by the same process.

Another important assumption the authors made was about the shape of the tail. They assumed that the tail of the return distribution of each individual asset follows a power law, in which the shape parameter is depending on both the time and the characteristics of the individual asset, see equation 1. It is important to notice that this assumption is only made for the *tail* of the return distribution of an individual asset, and that no assumptions about the whole return distribution are made. Many previous researchers found that the tail of the return distribution behaves according to (a distribution following) the power law, among others Jondeau and Rockinger (2003), Kearns and Pagan (1997), and Gopikrishnan, Plerou, Amaral, Meyer, and Stanley (1999).

$$P(R_{i,t+1} < r | R_{i,t+1} < u_t \quad and \quad F_t) = (\frac{r}{u_t})^{\frac{-\alpha_i}{\lambda_t}} \quad with \quad r < u_t < 0 \tag{1}$$

Here, $R_{i,t+1}$ denotes the return of asset i on time $t+1$ and F_t denotes all available information about the stock prices and returns up to time t. u_t is the threshold of the tail at time t, and only returns lower than this (negative) threshold are counted as a return in the tail. Equation 1 therefore denotes the distribution of the probability that the return is lower than a certain return r, given that it is in the tail (i.e. that it is already lower than u_t). The shape parameter of this power law is $\frac{\alpha_i}{\lambda_t}$ and consists of two parts. The first part is the constant α_i and is different for each asset to ensure that each asset has a different tail. The second part is the parameter λ_t, and is the same for all assets. This parameter is the common process that changes the shape of the tail of each asset over time in the same way. If λ_t is high, the tails are more fat and the probability of very negative returns in the tail is high. λ_t will be used as the tail risk variable in this paper, since it is the same for all assets.

To estimate λ_t, I still follow Kelly and Jiang (2014) and use Hill's power law estimator (Hill, 1975). Every month, I re-estimate λ_t, using the daily return data of multiple individual assets of that month. Also, I have a different threshold u_t every month, defined as the fifth percentile of the daily returns of all assets in month t. The idea is that returns under this threshold are in the tail. Hill's power law estimator is then given by:

$$\lambda_t = \frac{1}{K_t} \sum_{k=1}^{K_t} ln(\frac{R_{k,t}}{u_t}) \tag{2}$$

K_t is the total number of daily returns that fall below the threshold u_t (and is the same number each month if the number of assets in the sample is constant) and $R_{k,t}$ is the kth daily return that falls below u_t in month t.

3.1.2 Model 2: Maximum likelihood (ξ)

The second model contains an estimator of tail risk calculated with the threshold exceeding method. Gilli and Këllizi (2006) used this method to estimate the tails of a stock return distribution and this paper will follow their method. A general overview of the use of this method and similar methods in finance is given by Rocco (2014). The threshold exceeding method is similar to the method with the Hill estimator in the first model. As explained in the literature review, the main difference is that the method with the Hill estimator is non-parametric, while the threshold exceeding method is parametric. Also, the assumptions underlying both methods are different. The key-assumption of the threshold exceeding model is that the tail of the return distribution of *all* assets follows a generalized Pareto distribution, a distribution that satisfies the power law. The generalized Pareto distribution is given by the following equation:

$$P(R_{i,t} - u_t < r' | R_{i,t} < u_t) = \begin{cases} 1 - (1 + \frac{\xi_t * r'}{\sigma_t})^{-\frac{1}{\xi_t}} & if \quad \xi_t \neq 0 \\ 1 - e^{-\frac{r'}{\sigma_t}} & if \quad \xi_t = 0 \end{cases} \tag{3}$$

With $r' < u_t$. ξ_t is the shape parameter, also called the tail index, and is the parameter representing tail risk this paper will use as the second tail risk variable. σ is the scale parameter, $R_{i,t}$ is the return on asset i at time t and u_t is the threshold value. r' is the excess return under the threshold. Therefore, equation 3 gives the distribution of the probability that a return is at least r' lower than the threshold u_t, given that it is already lower than u_t. Since a strictly negative shape parameter ξ_t implies an upper bound to the returns, while in finance all returns are possible, we only consider the case where $\xi_t \geqslant 0$.

The idea of this model is that all returns lower than a threshold u are an extreme return in the tail of the return distribution. According to the extreme value theory, if this threshold goes to infinity (or in this case, minus infinity), the distribution of the returns that exceed of u (or in this case, fall below u) follows a general Pareto distribution (Rocco, 2014). The shape parameter of the general Pareto distribution therefore represents the shape of the tail of the return distribution.

The parameters of the distribution are estimated with maximum likelihood. The log-likelihood function corresponding to the generalized Pareto distribution is:

$$L(\xi_t, \sigma_t | r) = \begin{cases} -nln(\sigma) - (\frac{1}{\xi_t} + 1) \sum_{k=1}^{K_t} (1 + \frac{\xi_t}{\sigma_t} R'_{k,t}) & if \quad \xi_t \neq 0 \\ -nln(\sigma_t) - \frac{1}{\sigma_t} \sum_{k=1}^{K_t} R'_{k,t} & if \quad \xi_t = 0 \end{cases} \tag{4}$$

The variables are as defined in equation 3, except $R'_{k,t}$. $R'_{k,t}$ is the kth daily excess return under the threshold, $R'_{k,t} = R_{k,t} - u_t$, given that the return falls below the threshold. The maximum likelihood function is re-estimated every month and the shape parameter ξ_t is taken as the second tail risk parameter. For the threshold u_t we again take the fifth percentile of the daily returns of all assets in month t.

3.1.3 Fama-French regression

In the first model with the Hill estimator λ_t, and the second model with the maximum likelihood variable ξ_t, there might be a bias in the results due to the common fluctuations in stock prices. To prevent this bias, I will follow Kelly and Jiang (2014) and not use the stock returns in the first and second model, but the residuals of a Fama-French regression (Fama and French, 1993) instead. For each asset, I use daily data to perform an OLS-regression of the daily return on a constant and the three Fama-French factors: Small-Minus-Big (SMB), High-Minus-Low (HML) and the excess market return r* (the market return minus the risk-free interest rate), see equation 5. The factors are the same for all assets, but the coefficients are different for each asset. The residuals $e_{k,t}$ of asset k at time t replace the return $R_{k,t}$ in equation 2 and the threshold u_t is replaced by the fifth quantile of the residuals $e_{k,t}$ in each month t. The excess return $R'_{k,t}$ in the equation 4 is replaced by the excess residual $e'_{k,t} = e_{k,t} - u_t$

$$R_{k,t} = \beta_{0,k} + \beta_{1,k} SMB_t + \beta_{2,k} HML_t + \beta_{3,k} r^*_t \tag{5}$$

3.1.4 Model 3: Option prices ($LT(k)$ and $LT'(k)$)

The third tail risk variable comes from Bollerslev and Todorov (2011). The models above measure the realized tail risk with a variable that measures the realized shape of the tail of the return distribution. In contrast, the variable of the third model measures the shape of the tail expected by investors (ex-ante), and not the realized shape of the tail. Furthermore, this method estimates a tail risk variable for only one asset, which means I have to use an index to get a variable representing the tail risk in the whole stock market. Also, this method assumes a risk-neutral world. In the risk-neutral world, investors do not need to be compensated for risk.

To measure the tail risk, Bollerslev and Todorov (2011) used the slightly adjusted version of the standard Black-Scholes model by developed Carr and Wu (2003) in equation 6. In equation 6, S_t is the spot price of an asset at time t, and S_{t-} is the spot price of an asset at time t just before a jump occurs. Therefore, $\frac{dS_t}{S_{t-}}$ is the (relative) price change of an asset just after a jump. This price change is driven by three processes. The first term, $(r-q)dt$ is the expected increase of the stock-price during time dt in the risk-neutral world. r is the risk-free interest rate, and q is the dividend yield. Since investors do not need to be compensated for risk in the risk-neutral world, the expected growth of the asset price equals $(r-q)$. However, there is also a continuous random element involved, $\sigma_t dW_t$. W_t is a standard Brownian motion with mean zero and variance t. This random variable is multiplied by σ_t, the volatility of the asset price at time t, to get some random variation in the asset price change. Those first two terms form a standard Brownian motion with drift, and are the basis of the "normal" Black-Scholes model. However, the random variable W_t usually does not make large, sudden jumps and is therefore not suited to represent jumps in asset prices. Jumps however, are often used to model extremely large negative returns in asset prices. If an asset prices suddenly falls, it "jumps" downwards. Tail risk is therefore sometimes called "jump risk" instead. To include jumps, this standard model is extended with the term $\int_{\mathbb{R}^0}(e^x-1)(\mu(dx,dt)-\upsilon_t(x)dxdt)$. In this jump part, $\mu(dx,dt)$ "counts" all the jumps of size x that occur at time t. For every jump of size x, the stock price increases with (e^x-1). The jumps can have every size except 0, hence $\mathbb{R}^0$ (all the real numbers except zero). The term $\upsilon_t(x)dxdt$ is called the "local density" of the jumps, and is a discontinuous process with the density (i.e. probability) of the jumps, determining the form of the tails. It's discontinuous, since jumps do not occur that often, while the random variation in the asset price change from the term W_t is continuous.

$$\frac{dS_t}{S_{t-}} = (r-q)dt + \sigma_t dW_t + \int_{\mathbb{R}^0}(e^x-1)(\mu(dx,dt)-\upsilon_t(x)dxdt) \tag{6}$$

To estimate the tail risk expected by investors, Bollerslev and Todorov (2011) used the prices of deep-out-of-the-money European put options that are close to maturity. The authors argued that, if an option is sufficient out-of-money and the time to maturity goes to zero in the limit ($T \downarrow t$) small changes in the stock price due to the variation coming from the standard Brownian motion will not change the value of the option. Those put options will only be worth anything at maturity if the stock price suddenly makes a large fall, i.e. "jumps" downwards. If there is no large negative jump in the stock price, the options will be worthless at maturity. Therefore, the authors assumed that to determine the value of a close-to-maturity, deep-out-of-money put option only the expected downward jumps in the stock price are important.

Carr and Wu (2003) divided the risk-neutral valuation of an European option into two parts; one part related to the expected variation from the Brownian motion and one part related to the expected jumps. Since Bollerslev and Todorov (2011) assumed that the (small) changes in the stock price due to the Brownian motion will not influence the value of the close-to-maturity, deep-out-of-money European put options, they only used the part related to the expected jumps in the valuation of options. Under this assumption, the price of a close-to-maturity, out-of-money put option with strike price K at time t ($P_t(K)$) in the risk neutral world can be calculated with the following formula:

$$P_t(K) \approx e^{-r_{(t,T]}} \int_t^T E_t^{\mathbb{Q}} \int_{\mathbb{R}} ((K-F_{s-}e^x)^+(\upsilon_s^{\mathbb{Q}}(x)dx))ds \quad if \quad K < F_{s-} \quad and \quad T \downarrow t \tag{7}$$

Here, K is the strike price, T is the time at maturity , t is the current time, F_{s-} is the forward price just before the jump, $E_t^{\mathbb{Q}}$ is the risk-neutral expectation and $r_{(t,T]}$ denotes the risk-free interest rate at the current time t towards the maturity date T. The right hand side of equation 7 can be seen as the discounted expectation of the pay-off of the option, conditional on the value of K and on $T \downarrow t$. Intuitively, $\upsilon_s^{\mathbb{Q}}(dx)$ is the probability of a jump of size x from the density function. For every jump of size x ($\int_{\mathbb{R}}$), one calculates the pay-off of the option if this jump occurs ($(K-F_{s-}e^x)^+$) and multiplies this with the probability that the jump occurs to get the expected value of the pay-off of the option at time t. Then, this is integrated over the time to maturity $T-t$ ($\int_t^T$), since the jump may occur at every time, and discounted ($e^{-r_{(t,T]}}$) to the present. The term $\mu(dx,dt)$ from equation 6 does not occur in equation 7, since Bollerslev and Todorov assumed that only one jump can occur if $T \downarrow t$.

From equation 7, Bollerslev and Todorov (2011) derived a variable related to the density of the (left) tail of the return distribution, $LT_t(k)$:

$$LT_t(k) = \frac{1}{T-t}\int_t^T \int_{\mathbb{R}} (e^x - \frac{K}{F_{t-}})^+ (E_t^{\mathbb{Q}}(v_s^{\mathbb{Q}}(x)dx)ds \approx \frac{e^{r(T-t)} * P_t(k)}{(T-t)F_{t-}} \tag{8}$$

given that $T \downarrow t$. In those equations, k is the moneyness of the option, ($k = \frac{K}{F_t}$, the strike price divided by the future price) and $LT_t(k)$ is the variable related to the density of the tail of the return distribution as expected by the investors. Intuitively, $LT_t(k)$ is the expected value of the pay-off of the put option, conditional on the value of k and on $T \downarrow t$, per day (since it is multiplied by $\frac{1}{T-t}$) and scaled by the current future price F_{t-} to get a variable measuring the tail independent of the current future price.

Time to maturity and moneyness Bollerslev and Todorov (2011) fixed the moneyness k on 0.9 to use the variable $LT(k)$ in the rest of their article. To test which level of moneyness delivers the best results with respect to the forecasting of stock returns, I will fix the moneyness at different levels k, namely at k=0,6, 0.65, 0.7, 0.75, 0.8,0.85, 0.9 and 0.95. Furthermore, I will combine multiple estimations of $LT(k)$ by taking the mean, $LT(\bar{k})$.

Unfortunately, on most days there is no option with the exact moneyness k. To overcome this problem, Bollerslev and Todorov (2011) calculate the option price themselves with the Black-Scholes model. In the data there is for each option toe corresponding implied volatility, calculated with the standard Black-Scholes model. If there is no option with exactly moneyness k, I take the implied volatility σ_1 of the option with a moneyness k_1 lower than, but closest to k, and the implied volatility σ_2 of the option with a moneyness k_2 higher than, but closest to k. If, for example, k equals 0.75, and the levels of moneyness of the available options are 0.62, 0.73, 0.78 and 0.84, I take the implied volatility belonging to the option with moneyness 0.73 and the implied volatility belonging to the option with moneyness 0.78. Then, one can estimate the implied volatility of an option with moneyness k, σ, by linear interpolation:

$$\sigma = \sigma_1 + \frac{\sigma_2 - \sigma_1}{k_2 - k_1} * k \tag{9}$$

If there is no option with a lower/higher moneyness than k, I take the implied volatility of the option with the lowest/highest moneyness available. With the estimated implied volatility, the option price is estimated using the Black-Scholes formula:

$$P_t(k) = N(-d_2)Ke^{-r(T-t)} - N(-d_1)S_t \tag{10}$$

with

$$d_1 = \frac{1}{\sigma\sqrt{T-t}}(ln(\frac{S_t}{K}) + (r + \frac{\sigma^2}{2})(T-t)) \tag{11}$$

and

$$d_2 = d_1 - \sigma\sqrt{T-t} \tag{12}$$

N() is the CDF of the standard normal distribution, and σ the implied volatility of the returns of the asset, as estimated in equation 9. The other variables are as defined above.

Furthermore, this paper looks at two different kind of maturities. For the first kind of maturity, I follow Bollerslev and Todorov (2011) and take the shortest time-to-maturity option available to estimate $LT_t(k)$, with the restriction that the time-to-maturity is at least 8 days. I will also use only options with exactly 30 days to maturity, since I am looking for a monthly estimator, $LT_t'(k)$.

3.2 Predictive power in forecasting stock returns

3.2.1 Predictive regressions

To compare the forecast performance of λ_t, $LT_t(k)$ and ξ_t, I use the methods from Kelly and Jiang (2014). They calculated the predictive power in forecasting stock returns of an independent variable with a linear regression of the return in the next month on a constant and the independent variable in month t. They did the same for three other forecast horizons; the return in the next year, three years and five years. This paper will do the same for all variables.

Since I calculate the independent variable at a monthly frequency, there are overlapping observations for the forecast horizon of one year and longer. Therefore, I will follow Kelly and Jiang (2014) and use Hodrick's standard error correction for overlapping data (Hodrick, 1992) to calculate the variance and thereby the t-statistic of the coefficients.

There is one last complication in comparing the predictive power in forecasting stock returns of the three models; while λ_t and ξ_t change every month, $LT_t(k)$ is re-estimated every day. For a fair comparison, I will calculate the monthly average of $LT_t(k)$ and $LT'_t(k)$ and use this average in the regressions. Something similar is done by Gao, Gao, and Song (2016). They directly estimated the compensation for tail risk demanded by investors based on daily option prices, and then take the monthly average.

3.2.2 Out-of-sample regressions

Another way to compare the ability to forecast stock returns of the variables is by looking at their out-of-sample performance. Following Kelly and Jiang (2014), the out-of-sample performance of the variables is evaluated with an expanding window estimation. First, I perform an OLS-regression of the return on a constant and an independent variable, with an initial period of 120 months. With the coefficients from this regression, the expected return during the next forecast horizon is estimated. Afterwards, the sample period is extended with one extra month, and I re-estimate the coefficients and estimate the expected return during the next forecast horizon. This continues until the end of the sample period. The R^2 is calculated by comparing the errors of the forecasts from the regressions to the errors of a forecast that equals the mean return:

$$R^2 = 1 - \frac{\sum_t (\hat{R}_{m,t+1|t} - R_{m,t+1})^2}{\sum_t (\bar{R}_{m,t} - R_{m,t+1})^2} \tag{13}$$

In this equation, $\hat{R}_{m,t+1|t}$ is the estimation of the return on time $t+1$ (during one month, one year, three years or five years), made on time t, $R_{m,t+1}$ is the actual return of the forecast horizon on time $t+1$, and $\bar{R}_{m,t}$ is the average return on time t. If this R^2 is negative, the mean return is better in forecasting future stock returns than the estimations from the OLS-regression. If the R^2 is positive, the OLS-regression gives a better prediction for future stock returns than the average return. I calculate the significance of the R^2 (i.e. if R^2 is significantly different from zero) with the ENC-New test (Clark and McCracken, 2001).

For this test, I make two regressions. First, I regress the future return $R_{m,t+1}$ at time $t+1$ on only a constant, which gives the mean return $\bar{R}_{m,t}$, and calculate the error term $\hat{u}^1_{t+1} = R_{m,t+1} - \bar{R}_{m,t}$. Secondly, I regress the future return on a constant and the independent variable. This gives the error term $\hat{u}^2_{t+1} = R_{m,t+1} - \hat{R}_{m,t+1|t}$. The ENC-New test statistic is then calculated by

$$ENC - New = P \frac{\sum_t (\hat{u}^1_{t+1} - \hat{u}^1_{t+1}\hat{u}^2_{t+1})}{\sum_t (\hat{u}^2_{t+1})^2} \tag{14}$$

In this equation, P is the total number of out-of-sample observations. The critical values depend on the fraction $\pi = \frac{P}{R}$, where R equals the number of observations in the initialization period. I will use the critical values that were estimated in a simulation by Clark and McCracken (2001).

3.3 Benchmark parameters

I want to know if the tail risk variables are better or worse in predicting the future stock returns than other variables often used in the literature to predict stock returns. Therefore, the results of the three tail risk variables are compared to nearly all of the benchmark parameters used by Kelly and Jiang (2014). They used the variables from the article of Welch and Goyal (2007), who compared the ability of many standard variables to forecast stock returns, and the variance-risk-premium from the article of Bollerslev, Tauchen, and Zhou (2009). The ability of those benchmark variables to forecast stock returns is calculated with the same two methods as above; predictive regressions and out-of-sample regressions.

4 Data

For comparison reasons, I would like to use the same data in all three models. However, the kind of data needed for each model is different. The first model with the Hill estimator and the second model with the maximum likelihood estimator use data of individual assets on a daily basis. However, for the third model with the option prices, I cannot use the data of individual assets. Equation 8 only gives a variable for the tail of one option. To obtain a tail risk variable representing the whole market from this model, I need to use an option on an index representing the whole market.

Hence, I will use data of the option on an index for the third model, and I will use data of the individual assets that are part of this index for the first and second model. I will use two indices; the first index is the S&P-500 index. This index contains the 500 largest companies of the United States with respect to their market capitalization. However, smaller firms might have different tail risk than large firms. Therefore, I will also use the Russell-2000 index. This index contains around 2000 assets of companies with a relatively small market capitalization. It is made by taking the 2000 smallest companies from the Russell-3000 index.

For the first and second model, I need the data of individual assets. For the S&P-500 index, the list with the assets in the S&P-500 index and the period they belonged to this index comes from "Compustat- Capital IQ from Standard & Poor". As Kelly and Jiang (2014), I use this list to get the data of the individual assets in the S&P-500 index from "The Centre for Research in Security Prices" (CRSP). Also the return of the S&P-500 index comes from CRSP. For the Russell-2000 index, the individual assets that are part of this index and their return data come from "Bloomberg Finance". Finally, the daily Fama-French factors are downloaded from "Fama-French Portfolios and Factors".

For the third model with the option prices, the data about the option price, the dividend yield and the current price of the S&P-500 index and the Russell-2000 index are from OptionMetrics. Also the return on the Russell-2000 index and the risk-free interest rate come from OptionMetrics. When there is no risk-free interest rate with exactly the same time-to-maturity as the option, I use the closest risk-free interest rate available (and interpolate this interest rate for the time-to-maturity of our own option). Since OptionMetrics only has option data from January 1996 until April 2016, our sample period is from 1996 to 04-2016.

The benchmark variables from Welch and Goyal (2007) are downloaded from the website of Amit Goyal, and the variance risk premium of Bollerslev et al. (2009) is downloaded from the website of Hao Zhou.

5 Results

5.1 Model 1: Hill's estimator (λ)

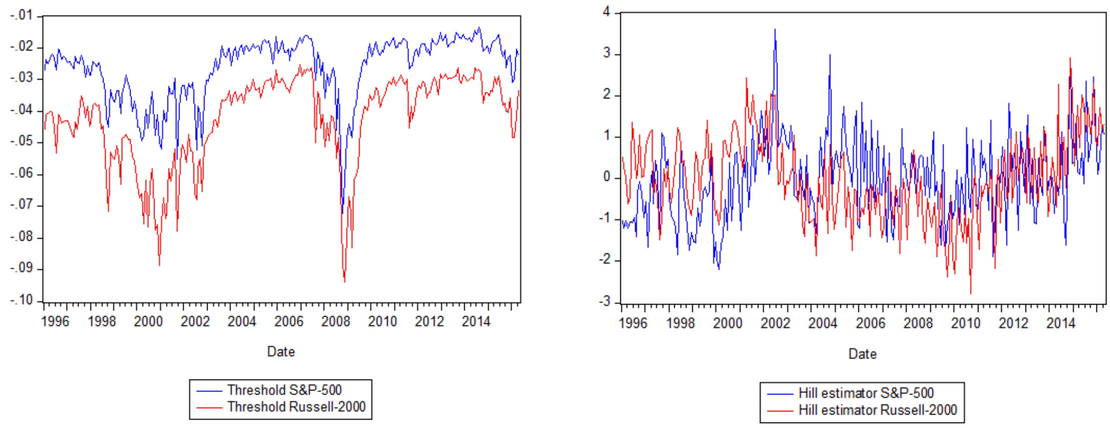
The first model estimated the Hill estimator λ representing the shape of the tail with all returns below the threshold value u_t. The Hill estimators for the large companies in the S&P-500 index, λ_{SP}, and for the small companies in the Russell-2000 index, λ_{Rus} and their threshold values u_t are plotted in figure 1b and 1a respectively. From figure 1a it appears that the threshold values of the Russell-2000 index and the S&P-500 index follow the same pattern. This is confirmed by the correlation of 96% between those two threshold values. The threshold of the S&P-500 index, however, is nearly always higher than the threshold of the Russell-2000 index. Figure 1b plots λ_{SP} and λ_{Rus}, both scaled to have a mean of zero and a standard deviation of 1. Both estimators look quite "messy", with many ups and downs . The general pattern of λ_{SP} and λ_{Rus} seems to be the same, though it is less clear than with the thresholds. The correlation between λ_{SP} and λ_{Rus} is indeed only 35%.

In figure 2a is the scaled three-year return (recalculated every month) of the S&P-500 index plotted together with the scaled λ_{SP}. The returns are less volatile and "spiky" than λ_{SP}. For λ_{Rus} and the return on the Russell-2000 index, the graph is quite similar. In figure 2b is a plot of the scaled absolute value of u_{Rus} and the scaled, monthly realized volatility of the returns on the Russell-2000 index. u_{Rus} and the volatility appear to follow the same pattern. If the volatility peaks, the absolute threshold value peaks as well.

The results of the in-sample predictive regression and the out-of-sample regressions are in table 1 and table 2 respectively. Both λ_{SP} and λ_{Rus} appear to be not that good in predicting the future stock returns. For the out-of-sample regressions the R^2 is usually negative. This means the the average return is better in predicting future stock returns than λ_{SP} and λ_{Rus}. In the article of Kelly and Jiang (2014), the Hill estimator λ was much better in predicting the stock returns. There are three possible reasons that our λ's perform worse in predicting the stock returns; it might be due to the data period, to the size of the data sample or to the company characteristics in the data sample.

In section 6, I evaluate whether the data period and the size of the data sample might explain the difference between the results in this paper and the results found by Kelly and Jiang (2014). The difference in the results might also be due to the different kind of assets in the data samples. The two data samples of this paper contain

Figure 1: The scaled Hill estimator λ and the corresponding threshold u_t for the S&P-500 index and the Russell-2000 index



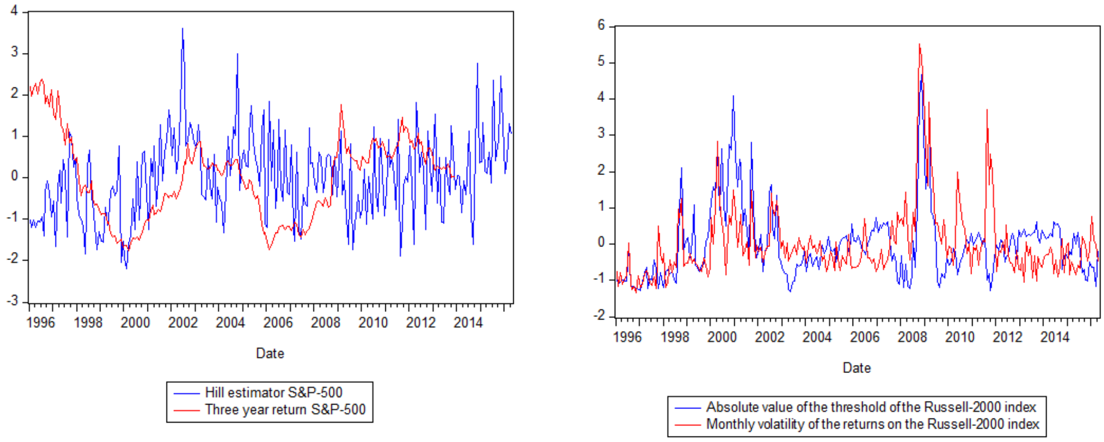
(a) u_{SP} and u_{Rus}

(b) λ_{SP} and λ_{Rus}

Figure 2

(a) λ_{SP} and the three year return on the S&P-500 index, both scaled

(b) Absolute value of u_{Rus} and the monthly realized volatility of the returns on the Russell-2000 index, both scaled



only large companies (S&P-500) or small companies (Russell-2000), while the original data sample used by Kelly and Jiang (2014) (with all the assets from NYSE,AMEX or NASDAQ with share code 10 or 11) contains both large, small and middle-sized companies.

Table 1: **Predictive regression results for λ**

Variable	One month			One year			Three years			Five years		
	β	T-stat.	R^2	β	T-stat.	R^2	β	T-stat.	R^2	β	T-stat.	R^2
λ^1_{SP}	-0.079	-2.35	2.24	-0.015	-0.68	0.70	0.002	0.16	0.03	0.003	0.23	0.08
λ^2_{Rus}	-0.020	-0.45	0.08	-0.021	-0.57	1.02	0.00	-0.08	0.06	0.00	-0.029	0.00

1. This Hill estimator is made with the data from the S&P-500 index and the dependent variable is the return on the S&P-500 index
2. This Hill estimator is made with the data from the Russell-2000 index and the dependent variable is the return on the Russell-2000 index

All the coefficients β in this table are scaled so that they represent the expected increase (in numeric values) in the annualized future stock returns when the independent variable increases with one standard deviation.

Table 2: **Out-of-sample results for λ**

Variable	One month		One year		Three years		Five years	
π	1.03		0.93		0.73		0.53	
Crit. value[1]	1.584		1.584		1.584		1.079	
	R^2	ENC-New	R^2	ENC-New	R^2	ENC-New	R^2	ENC-New
λ^2_{SP}	0.14	0.17	-1.29	-0.62	-0.63	-0.26	-2.89	-0.76
λ^3_{Rus}	-1.91	-0.97	0.39	0.68	-1.30	0.30	-17.19	-3.11

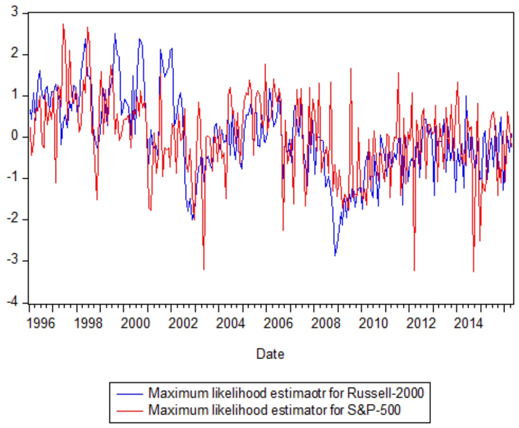
1. The approximate critical value comes from Clark and McCracken (2001) at a 5% significance level
2. This Hill estimator is made with the data from the S&P-500 index and the dependent variable is the return on the S&P-500 index
3. This Hill estimator is made with the data from the Russell-2000 index and the dependent variable is the return on the Russell-2000 index

5.2 Model 2: Maximum likelihood (ξ)

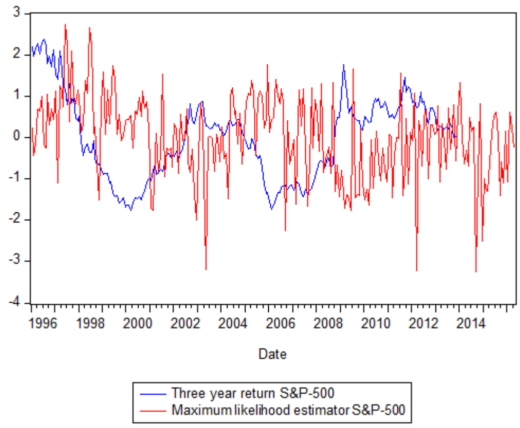
For the second model, I estimated the tail risk parameter ξ with maximum likelihood and the threshold exceeding method. Figure 3a plots ξ_{Rus}, estimated with the assets of the Russell-2000 index, and ξ_{SP}, estimated with the assets of the S&P-500 index, both scaled to have a mean of zero and a standard deviation of 1. ξ_{Rus} and ξ_{SP} seem to roughly follow the same pattern, which is confirmed by the correlation of 47%, though ξ_{SP} is much more "spiky" than ξ_{Rus}. Figure 3b plots ξ_{SP} and the scaled three year return on the S&P-500 index. The maximum likelihood estimator ξ_{SP} is much more volatile than the returns, and there is no clear relation between the returns and ξ_{SP} from the plot.

Figure 3

(a) ξ_{SP} and ξ_{Rus}, both scaled



(b) ξ_{SP} and the three year return on the S&P-500 index, both scaled



In table 3 and 4 are the results for the predictive regression and the out-of-sample regression of ξ respectively. The maximum likelihood estimator appears to be quite good in forecasting the returns in the next three and especially five years. The R^2 in the predictive regression is relatively high for this horizon (compared to the other time horizons), and moreover, the out-of-sample results indicate that ξ is significantly better at predicting the stock returns three or five years forward than the mean return. ξ_{Rus} seems to be a bit better in predicting the returns on the Russell-2000 index than ξ_{SP} in predicting the returns of the S&P-500 index.

Table 3: **Predictive regression results for ξ**

Variable	One month			One year			Three years			Five years		
	β	T-stat.	R^2	β	T-stat.	R^2	β	T-stat.	R^2	β	T-stat.	R^2
ξ^1_{SP}	-0.035	-1.02	0.43	0.00	0.19	0.04	-0.017	-1.09	1.83	-0.035	2.83	17.95
ξ^2_{Rus}	-0.045	-1.01	0.42	-0.055	-1.34	8.15	-0.046	-1.43	21.82	-0.040	-1.61	26.46

1. This maximum likelihood estimator is made with the data from the S&P-500 index and the dependent variable is the return on the S&P-500 index
2. This maximum likelihood estimator is made with the data from the Russell-2000 index and the dependent variable is the return on the Russell-2000 index

All the coefficients β in this table are scaled so that they represent the expected increase (in numeric values) in the annualized future stock returns when the independent variable increases with one standard deviation.

Table 4: **Out-of-sample results for ξ**

Variable	One month		One year		Three years		Five years	
π	1.03		0.93		0.73		0.53	
Cric. value[1]	1.584		1.584		1.584		1.079	
	R^2	ENC-New	R^2	ENC-New	R^2	ENC-New	R^2	ENC-New
ξ^2_{SP}	0.03	0.20	-2.13	-0.67	5.12	2.91	23.34	12.00
ξ^3_{Rus}	-2.16	-0.97	-2.09	3.63	18.15	16.93	37.77	24.11

1. The approximate critical value comes from Clark and McCracken (2001) at a 5% significance level
2. This maximum likelihood estimator is made with the data from the S&P-500 index and the dependent variable is the return on the S&P-500 index
3. This maximum likelihood estimator is made with the data from the Russell-2000 index and the dependent variable is the return on the Russell-2000 index

5.3 Model 3: Option prices

5.3.1 Minimum time-to-maturity ($LT(k)$)

In the third model, I estimated a tail risk variable from the option prices for different levels of moneyness k and for different maturity times; one with the shortest time-to-maturity available, $LT(k)$, and one with a time-to-maturity of 30 days, $LT'(k)$. In this paragraph, I will focus on the variable $LT(k)$ with the minimum time-to-maturity. Figure 4a is a plot of $LT(k)$ for all levels of moneyness k from the S&P-500 index, all scaled to have a mean of zero and a standard deviation of 1. All variables $LT(k)_{SP}$ appear to follow the same pattern. This is confirmed by the correlation between the variables. The lowest correlation between $LT(0.6)_{SP}$ and $LT(0.95)_{SP}$ is still 72%. For the Russell-2000 index, the variables move together very closely as well; the lowest correlation for the Russell-2000 index between $LT(0.6)_{Rus}$ and $LT(0.95)_{Rus}$ is 72 % as well.

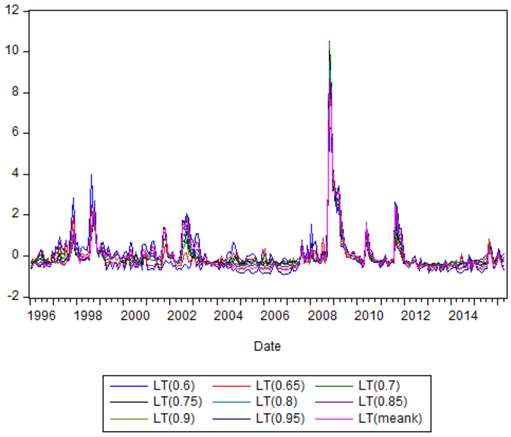
Figure 4b is the plot of $LT(\bar{k})_{Rus}$, $LT(\bar{k})_{SP}$, and the three year returns on the S&P-500 and the Russell-2000 index, all scaled. In this figure the two tail risk variables estimated with the Russell-2000 index and the S&P-500 index appear to be quite close together; the correlation is 94%. Also the three year returns of the S&P-500 index and the Russell-2000 index are quite close together in the graph. Furthermore, there is one large peak for both $LT(\bar{k})_{SP}$ and $LT(\bar{k})_{Rus}$ in October 2008, during the financial crisis. This peak was also visible in figure 4a.

In table 5 and 6 are the results of the in-sample predictive regressions and the out-of-sample regressions respectively. For both the large companies of the S&P-500 index and the small companies of the Russell-2000 index, the highest level of moneyness (k=0.95) is the best in predicting the stock returns in the next five years. Unfortunately, since I need to use out-of-the-money options, we cannot calculate $LT(k)$ for a higher level of moneyness. For the other time horizons the results are less straight forward. The highest level of moneyness still performs quite well, but for the in-sample regressions, the lowest level of moneyness (k=0.6) seems to be better. For the out-of-sample regressions, medium levels of moneyness like k=0.8 are the best. Taking the mean of $LT(k)$ does not improve the results.

$LT(k)$ is not a good variable to estimate the stock returns one month forward; for both the large and the small companies, and for all levels of moneyness, the mean return is better in predicting the stock returns the next month than $LT(k)$. For the other time horizons $LT(k)$ is better in predicting the stock returns than the mean return. Furthermore, for all time horizons $LT(k)$ is better in predicting the stock returns of the small companies than the stock returns of the large companies.

Figure 4

(a) $LT(k)$ (scaled) for the S&P-500 index and all levels of moneyness k



(b) $LT(\bar{k})_{SP}$, $LT(\bar{k})_{Rus}$, the three year return on the S&P-500 and the Russell-2000 index, all scaled

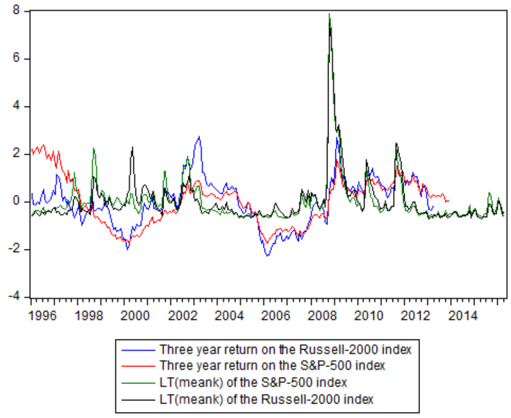


Table 5: **Predictive regression results for $LT(k)$**

Variable	One month			One year			Three years			Five years		
	β	T-stat.	R^2	β	T-stat.	R^2	β	T-stat.	R^2	β	T-stat.	R^2
Return on the S&P-500 index												
$LT(0.6)_{SP}$	-0.23	-0.67	0.19	0.041	1.55	**5.45**	0.019	1.36	**2.64**	0.018	2.40	5.61
$LT(0.65)_{SP}$	-0.025	-0.73	0.22	0.036	1.38	4.13	0.019	1.41	2.62	0.023	3.05	9.60
$LT(0.7)_{SP}$	-0.030	-0.89	0.32	0.031	1.23	3.05	0.016	1.28	1.96	0.026	3.39	11.76
$LT(0.75)_{SP}$	-0.031	-0.90	0.33	0.029	1.16	2.65	0.015	1.17	1.68	0.028	3.52	13.41
$L(0.8)_{SP}$	-0.029	-0.86	0.30	0.028	1.15	2.62	0.015	1.12	1.56	0.029	3.51	14.47
$LT(0.85)_{SP}$	-0.026	-0.77	**0.25**	0.029	1.15	2.68	0.015	1.15	1.73	0.030	3.45	15.51
$LT(0.9)_{SP}$	-0.018	-0.54	0.12	0.028	1.11	2.55	0.016	1.16	1.94	0.032	3.28	17.14
$LT(0.95)_{SP}$	0.00	-0.14	0.00	0.021	0.82	1.48	0.017	1.09	2.13	0.035	2.93	**19.29**
$LT(\bar{k})^1_{SP}$	-0.017	-0.48	0.10	0.026	1.03	2.26	0.017	1.16	2.05	0.032	3.24	17.42
$LT(k)^2_{SP}$	-0.024	-0.72	0.21	0.029	1.17	2.82	0.016	1.19	1.90	0.030	3.43	15.44
Return on the Russell-2000 index												
$LT(0.6)_{Rus}$	0.016	0.36	0.05	0.043	1.26	**5.04**	0.033	1.39	**11.21**	0.039	2.49	26.13
$LT(0.65)_{Rus}$	0.013	0.28	0.03	0.034	1.03	3.16	0.026	1.27	7.44	0.035	2.63	21.20
$LT(0.7)_{Rus}$	0.018	0.40	0.06	0.025	0.75	1.75	0.022	1.11	5.22	0.034	2.75	19.90
$LT(0.75)_{Rus}$	0.011	0.26	0.03	0.026	0.77	1.80	0.022	1.15	5.13	0.035	3.07	21.67
$L(0.8)_{Rus}$	0.010	0.23	0.02	0.029	0.86	2.26	0.024	1.20	5.96	0.039	3.26	25.57
$LT(0.85)_{Rus}$	0.020	0.44	0.08	0.032	0.93	2.74	0.027	1.31	7.71	0.043	3.39	30.81
$LT(0.9)_{Rus}$	0.030	0.68	0.19	0.033	0.95	2.94	0.030	1.37	9.30	0.047	3.42	35.24
$LT(0.95)_{Rus}$	0.040	0.91	**0.34**	0.032	0.90	2.78	0.033	1.42	11.08	0.051	3.39	**38.98**
$LT(\bar{k})^1_{Rus}$	0.030	0.67	0.19	0.033	0.92	2.86	0.030	1.36	9.49	0.047	3.37	35.02

1. $LT(\bar{k})$ is the mean of the LT(k) for all levels of moneyness k.
2. This $LT(k)$ is made with the mean of $LT(k)$ for all levels of moneyness k, except k=0.95, based on the in-sample performance
All the coefficients β in this table are scaled so that they represent the expected increase (in numeric values) in the annualized future stock returns when the independent variable increases with one standard deviation. In each column, the highest R^2 is blue and bold.

Table 6: **Out-of-sample results for $LT(k)$**

Variable	One month		One year		Three years		Five years	
π	1.03		0.93		0.73		0.53	
Crit. value[1]	1.584		1.584		1.584		1.079	
	R^2	ENC-New	R^2	ENC-New	R^2	ENC-New	R^2	ENC-New
Return on the S&P-500 index								
$LT(0.6)_{SP}$	-11.84	-4.05	-9.29	8.38	2.28	5.59	2.45	1.67
$LT(0.65)_{SP}$	-40.96	7.00	-22.91	9.60	-8.73	7.59	14.39	6.79
$LT(0.7)_{SP}$	-71.39	-8.36	-17.9	10.51	-71.39	-8.36	22.27	13.20
$LT(0.75)_{SP}$	-47.44	-7.39	4.94	8.84	**3.73**	2.43	24.56	16.12
$LT(0.8)_{SP}$	-24.46	-6.03	7.36	7.94	0.28	0.70	26.68	17.31
$LT(0.85)_{SP}$	-14.62	-4.81	**7.93**	7.37	1.16	1.11	28.44	18.50
$LT(0.9)_{SP}$	-9.92	-3.97	6.79	5.73	2.01	1.65	30.46	20.29
$LT(0.95)_{SP}$	**-6.39**	-3.00	2.02	2.16	1.58	1.66	**33.10**	22.37
$LT(\bar{k})^2_{SP}$	-10.67	-4.19	5.61	4.63	2.55	2.00	30.98	21.06
$LT(k)^3_{SP}$	-17.73	-5.37	7.64	7.73	3.31	2.27	28.15	18.59
Return on the Russell-2000 index								
$LT_(0.6)_{Rus}$	-2.73	2.23	5.37	10.71	-8.81	17.58	45.18	33.56
$LT_(0.65)_{Rus}$	-2.03	5.19	7.31	7.37	8.50	12.59	38.81	26.30
$LT_(0.7)_{Rus}$	-3.57	8.28	4.52	4.03	12.87	8.61	38.11	25.49
$LT(0.75)_{Rus}$	-2.04	5.64	5.74	4.56	12.96	8.89	41.47	34.12
$LT(0.8)_{Rus}$	**-1.95**	1.21	5.47	4.66	14.13	11.20	41.69	45.25
$LT(0.85)_{Rus}$	-3.80	-0.74	**7.55**	6.12	17.78	16.31	46.73	57.09
$LT(0.9)_{Rus}$	-4.83	-1.89	7.07	6.05	21.30	18.85	54.46	68.20
$LT(0.95)_{Rus}$	-4.43	-1.99	5.39	5.05	**24.33**	21.92	**59.05**	75.43
$LT(\bar{k})^2_{Rus}$	-4.56	-1.41	6.54	5.81	21.09	19.83	52.25	67.48

1. The approximate critical value comes from Clark and McCracken (2001) at a 5% significance level
2. $LT(\bar{k})$ is the mean of $LT(k)$ for all levels of moneyness k
3. This $LT(k)$ is made with the mean of $LT(k)$ for all levels of moneyness k, except k=0.95, based on the in-sample performance.

In each column, the highest R^2 is blue and bold.

5.3.2 Time-to-maturity of 30 days ($LT'(k)$)

I also estimated the tail risk variable $LT'(k)$ with a time-to-maturity of exactly 30 days (or 29 days when there is no option with a time-to-maturity of 30 days in a month). In figure 5) is a plot of the scaled $LT'(k)_{Rus}$, estimated with the data from the Russell-2000 index. All variables $LT'(k)_{Rus}$ appear to be quite close together, which was the same in figure 4a. The smallest correlation between $LT'(0.6)_{Rus}$ and $LT'(0.95)_{Rus}$ is now 73 %.

From the results in table 7 and 8, it is clear that $LT(k)$ (estimated with the minimum time-to-maturity) is better in forecasting the stock returns than $LT'(k)$ (estimated with 30 days to maturity) for both the large and the small companies. The only exception is the time horizon of one month; then, $LT'(k)$ is much better in forecasting the returns than $LT(k)$. The prices of put options reflect the expectations of investors about the development of the stock prices during the time to maturity. For $LT'(k)$, this period is 30 days, and therefore part of the time-to-maturity falls into the next month. For $LT(k)$, the mean time-to-maturity is only ten trading days. Therefore, many observations of options in the beginning of the month do not have any days during the time-to-maturity in the next month. This might explain why $LT(k)$ is worse in predicting the stock returns in the next month than the mean return, while many $LT'(k)$'s are better.

The reason $LT(k)$ is so much better than $LT'(k)$ in predicting returns more than one month ahead might be due to the number of observation used while estimating $LT(k)$ and $LT'(k)$. In both data samples, there are not that many options with a time-to-maturity of exactly 30 days. In the beginning of the sample, there is often only one day each month with an option with a time to maturity of 30 days. In contrast, for $LT(k)$, we have a data point on (nearly) every trading day of the month. More data points might increase the reliability and the predictive power with respect to forecasting stock returns of $LT(k)$ compared to $LT'(k)$.

Figure 5: $LT'(k)$ (with a time-to-maturity of 30 days) for the Russell-2000 index and all levels of moneyness k, all scaled

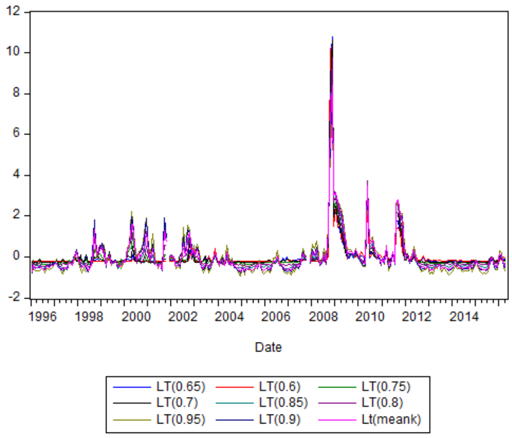


Table 7: **Predictive regression results for $LT'(k)$**

Variable	One month			One year			Three years			Five years		
	β	T-stat.	R^2	β	T-stat.	R^2	β	T-stat.	R^2	β	T-stat.	R^2
Return on the S&P-500 index												
$LT'(0.6)$	-0.055	-1.63	**1.11**	0.030	1.21	**2.85**	0.016	1.24	1.77	0.022	2.97	8.26
$LT'(0.65)$	-0.049	-1.43	0.85	0.028	1.14	2.49	0.017	1.31	**2.02**	0.024	3.16	10.18
$LT'(0.7)$	-0.040	-1.19	0.58	0.025	1.05	2.03	0.015	1.22	1.68	0.025	3.35	11.19
$LT'(0.75)$	-0.033	-0.97	0.39	0.026	1.09	2.18	0.014	1.17	1.47	0.026	3.52	11.95
$LT'(0.8)$	-0.029	-0.87	0.31	0.027	1.12	2.32	0.014	1.18	1.52	0.027	3.54	12.84
$LT'(0.85)$	-0.022	-0.66	0.18	0.026	1.07	2.19	0.014	1.14	1.50	0.028	3.45	13.65
$LT'(0.9)$	-0.018	-0.54	0.12	0.024	0.95	1.81	0.015	1.07	1.52	0.030	3.22	14.70
$LT'(0.95)$	-0.013	-0.38	0.06	0.018	0.69	1.01	0.015	0.98	1.57	0.031	2.81	**15.74**
$LT'(\bar{k})^1$	-0.023	-0.66	0.18	0.024	0.93	1.76	0.015	1.10	1.62	0.030	3.26	14.55
Return on the Russell-2000 index												
$LT'(0.6)$	-0.015	-0.33	0.05	0.026	0.84	1.76	0.020	1.15	4.54	0.028	2.57	14.67
$LT'(0.65)$	0.00	-0.17	0.01	0.025	0.81	1.65	0.020	1.13	4.35	0.029	2.56	14.58
$LT'(0.7)$	0.00	-0.03	0.00	0.026	0.83	1.80	0.021	1.16	4.82	0.031	2.66	16.67
$LT'(0.75)$	0.00	0.05	0.00	0.028	0.89	2.08	0.022	1.23	5.43	0.034	2.88	19.35
$LT'(0.8)$	0.017	0.38	0.06	0.030	0.95	2.40	0.024	1.30	6.12	0.037	3.18	22.70
$LT'(0.85)$	0.025	0.56	0.13	0.030	0.95	**2.48**	0.027	1.40	7.61	0.040	3.38	26.68
$LT'(0.9)$	0.033	0.75	0.23	0.030	0.93	**2.48**	0.030	1.49	9.41	0.044	3.46	30.47
$LT'(0.95)$	0.041	0.93	**0.36**	0.030	0.87	2.37	0.032	1.49	**10.68**	0.046	3.46	**33.07**
$LT'(\bar{k})^1$	0.030	0.67	0.19	0.031	0.93	2.57	0.030	1.45	9.50	0.044	3.31	30.45

1.$LT'(\bar{k})$ is the mean of the $LT'(k)$ for all levels of moneyness k
All the coefficients β in this table are scaled so that they represent the expected increase (in numeric value) in the annualized future stock returns when the independent variable increases with one standard deviation. In each column, the highest R^2 is blue and bold.

5.4 Overall comparison of all variables

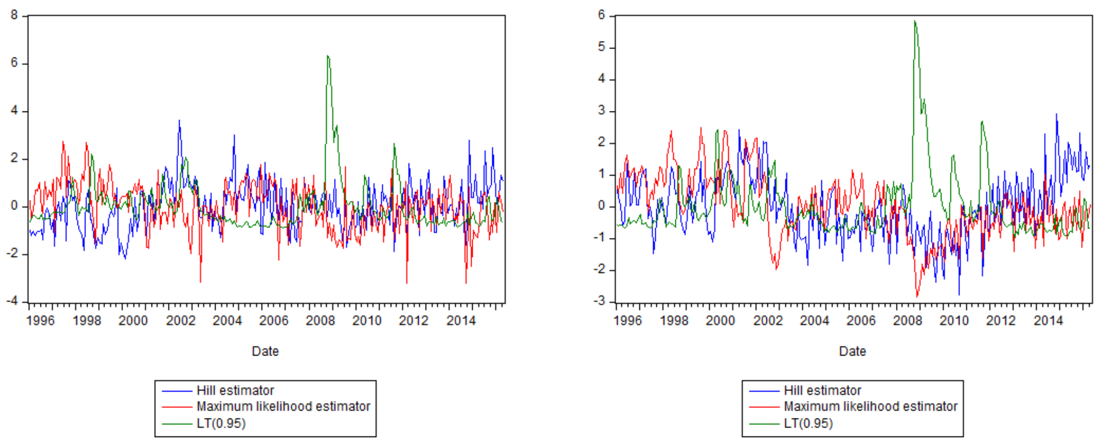
In this section, I will compare all three tail risk variables with each other. In figure 6 is a plot of λ, ξ and $LT(0.95)$ for both the Russell-index and the S&P-500 index, all scaled to have a mean of zero and a standard deviation of 1. $LT(0.95)$ differs the most from the other two variables, mainly due to its peak in 2008 and since it's less "spiky" than the other two tail risk variables. λ and ξ are closer together. The reason might be that λ and ξ have both been constructed with the same, changing threshold u_t, while $LT(0.95)$ has been estimated with a fixed level of moneyness k=0.95. This is in line with the results of Almeida, Ardison, Garcia, and Vicente (2017), who compared their tail risk variable to the variable of Kelly and Jiang (2014) and found that their tail risk variable is quite different from λ as well, especially when there is a shock in the stock market.

Table 8: **Out-of-sample results for $LT'(k)$**

Variable	One month		One year		Three years		Five years	
π	1.03		0.93		0.73		0.53	
Crit. value[1]	1.584		1.584		1.584		1.079	
	R^2	ENC-New	R^2	ENC-New	R^2	ENC-New	R^2	ENC-New
Return on the S&P-500 index								
$LT'(0.6)_{SP}$	-7.16	3.08	-109.02	-3.16	-33.87	-0.27	0.98	1.22
$LT'(0.65)_{SP}$	-14.67	3.14	-76.19	-3.71	-37.32	-0.10	8.73	4.07
$LT'(0.7)_{SP}$	-7.79	5.18	-12.24	-1.80	-4.00	-0.43	6.37	5.83
$LT('(0.75)_{SP}$	8.79	11.83	-2.68	-0.54	**-2.33**	-0.81	8.95	6.41
$LT'(0.8)_{SP}$	12.89	14.60	-1.59	-0.25	-3.37	-1.06	10.67	7.20
$LT'(0.85)_{SP}$	14.23	15.45	**-0.26**	0.07	-3.94	-1.29	13.07	7.75
$LT'(0.9)_{SP}$	15.67	17.13	-0.28	0.23	-3.97	-1.38	15.38	8.84
$LT'(0.95)_{SP}$	**16.54**	18.46	-1.13	0.39	-4.05	-1.44	**18.55**	10.26
$LT'(\bar{k})^2_{SP}$	16.52	18.27	-0.39	0.25	-3.65	-1.29	14.74	8.81
Return on the Russell-2000 index								
$LT'(0.6)_{Rus}$	-284.00	-5.21	-64.87	-3.87	-199.67	-2.45	-2.31	5.71
$LT'(0.65)_{Rus}$	-115.05	-4.26	-2.28	0.90	-13.76	-1.66	3.43	5.50
$LT'(0.7)_{Rus}$	-35.01	-0.59	-15.71	3.74	0.55	0.68	-12.19	6.95
$LT('(0.75)_{Rus}$	0.78	5.83	-2.71	1.85	-0.00	0.50	-12.70	9.04
$LT'(0.8)_{Rus}$	9.04	10.24	-0.25	0.47	0.71	1.28	7.02	13.09
$LT'(0.85)_{Rus}$	10.45	12.29	0.18	0.59	3.08	3.39	17.53	17.57
$LT'(0.9)_{Rus}$	10.84	13.59	**0.58**	0.78	6.82	6.07	26.89	22.58
$LT'(0.95)_{Rus}$	**11.05**	14.78	**0.58**	0.83	**9.87**	8.11	**33.46**	26.76
$LT'(\bar{k})^2_{Rus}$	10.85	13.69	0.34	0.91	4.57	6.15	22.15	21.34

1. The approximate critical value comes from Clark and McCracken (2001) at a 5% significance level
2. $LT'(\bar{k})$ is the mean of $LT'(k)$ for all levels of moneyness k
In each column, the highest R^2 is blue and bold.

Figure 6: λ, ξ and $LT(0.95)$ for the S&P-500 index and the Russell-2000 index, all scaled



(a) S&P-500 index

(b) Russell-2000 index

For the predictive regressions, there is often a negative coefficient when I regress the return in the next month on a tail risk variable. This is only the case for the time-horizon of one month, and not for (most) $LT(k)_{Rus}$ and $LT'(k)_{Rus}$. Furthermore, only one coefficient is significantly different from zero and negative, namely the coefficient of λ_{SP}. All the other negative coefficients are not significantly different from zero. Still, those results are quite strange; I expected positive coefficients. As explained in the introduction, if the tail risk is high, investors want to be compensated for the risk with a higher expected return. Kelly and Jiang (2014) indeed found only positive coefficients when they used their tail risk variable λ in the predictive regressions. However, similar results with a negative, and only sometimes significant coefficient when regressing the return in the next month on a tail risk variable were found by Bali, Cakici, and Whitelaw (2014). In appendix A, I looked whether the coefficient of λ_{SP} is still significantly negative if I estimate λ_{SP} for the period 1963-2010, and I found that the coefficient is still negative, but not significantly different from zero any more. I also re-estimated the original results of Kelly and Jiang (2014) in this appendix, and found indeed only positive coefficients. In appendix B, I analysed what would happen if I used the value-weighted return of CRSP (as done by Kelly and Jiang, 2014). All coefficients that were negative before, are still negative in those regressions.

When comparing the ability of the three tail risk variables to predict future stock returns, $LT(k)$ (estimated with the minimum time-to-maturity), $LT(0.95)$, $LT(0.85)$ and $LT(0.75)$ in particular, has the best results in both data samples. The R^2 in both the predictive and the out-of-sample regressions is usually the highest for this variable compared to the other variables, for both the large companies in the S&P-500 index and the small companies in the Russell-2000 index. Furthermore, the results for predicting the stock returns of the Russell-2000 index are better than the results for predicting the stock returns of the S&P-500 index when using $LT(k)$. This is probably due to the differences in the size of the firms in both indices. One exception is the prediction of the stock returns in the next month. $LT'(k)$ (estimated with a time-to-maturity of 30 days), $LT'(0.95)$ in particular, is the best in forecasting the stock returns in the next month. Moreover, $LT'(k)$ is better in forecasting the returns of the S&P-500 index in the next month than the returns of the Russell-2000 index.

5.5 Benchmark variables

In this section, I test whether the tail risk variables are better in predicting the future stock returns than the benchmark variables used by Welch and Goyal (2007) and the variance risk premium used by Bollerslev et al. (2009). In table 9 and 10 are the results for the predictive regressions and the out-of-sample regressions on the return of the S&P-500 index and of the Russell-2000 index. For predicting the stock-market return in the next month, $LT'(k)$ with a high level of moneyness k is still a good choice; the out-of-sample results for this variable are better than the out-of-sample results for the benchmark variables, though this is not the same for the predictive regressions. For longer forecast horizons, however, the tail-risk related variable $LT(k)$ is not necessarily the best choice. The dividend pay-out ratio and the dividend yield spread have a considerably higher R^2 for the predictive, in-sample regressions. For the out-of-sample forecasts, the variance risk premium, the treasury-bill rate and the term spread have a higher R^2 during at least one forecast horizon than $LT(k)$. However, for the out-of-sample regressions, this result depends on the forecast horizon and the index used. Sometimes $LT(k)$ still gives the best results, i.e. the highest R^2, compared to all the benchmark variables.

The R^2 of the regressions in table 9 and 10 are usually higher than the R^2 found by Kelly and Jiang (2014) when they used the same benchmark variables. There are two possible explanations. First of all, I use a different sample period than Kelly and Jiang (2014), which might influence the results. Since our initial estimation window is ten years, the out-of-sample regressions are performed over the short period 2006-2016 (2006-2011 for the five year returns). Furthermore, I also regress on a different return. Kelly and Jiang used the value-weighted return of CRSP, while I use the return on the S&P-500 index and the Russell-2000 index. Since many benchmark variables (namely the dividend pay-out ratio, the dividend price ratio, the earnings price ratio, the stock variance and the variance risk premium) are constructed with data from the S&P-500 index, it is not surprising that the predictive regressions on the return of the S&P-500 index have a higher R^2 than found by Kelly and Jiang (2014).

6 Robustness check: different performance of λ

The results of section 5.1 indicated that the Hill estimators λ_{SP} and λ_{Rus} of the first model were not that good in predicting future stock returns. In contrast, Kelly and Jiang (2014) found that the Hill estimator λ of their data sample outperformed many benchmark variables in predicting future stock returns. As explained in section 5.1, there are three possible explanations for this difference.

First of all, it might be that the results of this paper and of Kelly and Jiang (2014) are so different because of the different data period; our data period starts in 1996 and ends in April 2016. The data period of Kelly and Jiang (2014) goes from 1963 to 2010. To analyse the difference in the data period, I estimated $\lambda_{original}$; the Hill estimator with the original data sample of Kelly and Jiang (2014), but with the data period from 1996 to 04-2016. The data sample of Kelly and Jiang (2014) contains all the assets from NYSE, AMEX and NASDAQ with share code 10 or 11, and I download this data sample from CRSP. Furthermore, as the return they took the value-weighted returns from CRSP as well.

Table 9: **Predictive regression results for the benchmark variables**

Variable	One month			One year			Three years			Five years		
	β	T-stat.	R^2	β	T-stat.	R^2	β	T-stat.	R^2	β	T-stat.	R^2
Return on the S&P-500 index												
Book to market	0.020	0.59	0.14	0.05	1.66	9.47	0.13	1.43	13.11	0.20	1.18	20.58
Default return spread	0.035	1.01	0.43	0.012	0.84	0.44	0.00	1.34	0.39	0.011	2.43	1.54
Default yield spread	-0.044	-1.30	0.70	0.0143	0.41	0.66	0.0205	0.95	2.99	0.0458	2.97	35.65
Dividend pay out ratio	0.00	0.06	0.00	0.0235	0.78	1.81	0.0247	1.40	4.55	0.0465	4.15	**38.07**
Dividend price ratio	-0.00	-0.11	0.00	-0.0565	-2.48	10.37	-0.0416	-4.72	12.91	-0.0196	-1.98	6.97
Earnings price ratio	-0.00	-0.11	0.00	-0.0576	-2.55	10.78	-0.0445	-4.86	14.73	-0.0306	-3.06	16.73
Inflation	0.023	0.67	0.19	-0.0249	-2.15	1.97	-0.008907	-1.63	0.53	-0.00951	-2.33	1.34
Long-term return	0.021	0.62	0.16	0.00424	0.60	0.06	0.00465	1.76	1.43	-0.000803	-0.41	0.00
Long-term yield	-0.012	-0.36	0.05	-0.00693	-0.207	0.14	-0.0192	-0.54	1.69	-0.0384	-0.66	9.75
Net equity expansion	0.073	2.14	1.86	0.0668	1.42	**13.70**	0.0223	-.74	3.48	-0.00548	-0.230	0.48
Stock volatility	-0.077	-2.27	2.10	0.00940	0.36	0.29	0.00832	0.65	0.51	0.0245	3.14	10.32
Term spread	-0.019	-0.55	0.12	0.022	0.71	1.59	0.0586	2.28	**24.92**	0.0428	2.83	31.85
Treasury-bill rate	0.00	0.11	0.00	-0.018	-0.53	1.01	-0.0514	-1.64	16.23	-0.0506	-1.69	30.91
Variance risk premium	0.141	4.28	**7.06**	0.034	1.87	3.70	0.0133	1.01	1.28	0.00616	0.49	0.63
Return on the Russell-2000 index												
Book to market	0.00	0.22	0.02	0.059	1.35	**9.09**	0.037	0.91	12.99	0.021	0.49	5.78
Default return spread	0.035	0.78	0.25	0.021	1.11	1.09	$7.59*10^{-3}$	1.01	0.57	0.00	1.48	0.97
Default yield spread	-0.025	-0.55	0.13	0.048	1.05	6.11	0.043	1.47	18.76	0.056	2.83	**51.90**
Dividend pay out ratio	0.010	0.24	0.02	0.044	1.06	5.16	0.037	1.56	14.68	0.055	3.86	51.93
Dividend price ratio	-0.00	-0.07	0.00	-0.046	-1.59	5.66	-0.019	-1.55	3.75	-0.014	-1.10	3.39
Earnings price ratio	-0.00	-0.12	0.00	-0.053	-1.80	7.72	-0.027	-2.02	7.88	-0.028	-2.18	13.30
Inflation	0.012	0.28	0.03	-0.022	-1.55	1.33	-0.00	-1.41	0.94	-0.014	-2.78	2.91
Long-term return	0.033	0.75	0.23	0.00	0.23	0.01	0.00	1.88	0.40	0.00	0.94	0.08
Long-term yield	-0.021	-0.47	0.09	-0.026	-0.63	1.65	-0.041	-0.91	12.13	-0.041	-0.55	11.09
Net equity expansion	0.044	1.00	0.41	0.054	0.91	7.51	0.00	0.25 0.96	-0.00	-0.31	1.46	
Stock volatility	-0.096	-2.19	1.95	0.021	0.60	1.20	0.022	1.27	4.97	0.036	3.57	21.51
Term spread	0.00	0.18	0.01	0.046	1.15	5.83	0.071	2.08	**52.53**	0.040	2.05	27.16
Treasury-bill rate	-0.018	-0.40	0.07	-0.045	-1.03	5.16	-0.072	-1.72	45.68	-0.049	-1.27	28.52
Variance risk premium	0.169	3.91	**5.97**	0.037	1.55	3.60	0.020	1.12	4.06	0.013	0.81	2.64

All the coefficients β in this table are scaled so that they represent the expected increase (in numeric value) in the annualized future stock returns when the independent variable increases with one standard deviation. In each column, the highest R^2 is blue and bold.

Table 10: **Out-of-sample results for the benchmark variables**

Variable	One month		One year		Three years		Five years	
π	1.03		0.93		0.73		0.53	
Crit. value[1]	1.584		1.584		1.584		1.079	
	R^2	ENC-New	R^2	ENC-New	R^2	ENC-New	R^2	ENC-New
For the return on the S&P-500 index								
Book to market	-1.54	-0.81	-0.38	11.42	10.89	12.86	27.51	17.03
Default return spread	-5.14	-2.05	-2.77	-0.59	-2.69	-0.26	-4.72	-1.17
Default yield spread	-4.72	6.53	-11.64	4.14	5.44	5.35	58.98	79.56
Dividend pay out ratio	-7.47	1.58	1.65	7.06	2.48	31.35	41.03	89.43
Dividend price ratio	-1.10	-0.48	-2.65	4.53	-16.98	1.85	-13.01	-1.27
Earnings price ratio	-1.29	-0.62	-0.89	8.79	-13.58	8.96	-0.70	5.98
Inflation	-1.01	-0.34	2.86	3.21	0.54	0.47	0.31	0.26
Long-term return	-1.42	-0.80	-1.32	-0.70	-1.33	-0.48	-4.36	-1.27
Long-term yield	-1.80	-0.59	-4.11	1.42	-8.60	2.96	13.07	6.41
Net equity expansion	0.29	3.61	-13.56	42.92	-20.00	25.76	-14.50	1.52
Stock volatility	-3.23	2.45	-4.78	0.07	-4.79	-1.43	19.08	14.94
Term spread	-0.49	-0.11	**5.26**	4.04	**56.67**	89.08	46.56	35.62
Treasury-bill rate	-1.53	-0.36	2.37	2.18	45.57	53.01	**51.21**	43.80
Variance risk premium	**9.31**	13.45	-12.95	1.28	-3.93	0.12	-5.95	-1.61
For the return on the Russell-2000 index								
Book to market	-1.49	-0.81	-8.77	5.12	7.37	8.14	6.94	3.39
Default return spread	-4.92	-1.95	-3.56	0.10	-2.90	-0.42	-4.43	-1.12
Default yield spread	-4.74	0.47	-3.64	3.03	-4.03	2.66	-4.72	2.51
Dividend pay out ratio	-6.24	-0.58	-5.16	0.86	-5.92	50.78	46.87	132.81
Dividend price ratio	-1.00	-0.48	0.06	0.93	-2.61	1.31	-1.54	0.81
Earnings price ratio	-1.24	-0.63	0.26	0.61	0.36	0.57	0.57	0.61
Inflation	-1.85	-0.69	-0.68	0.87	1.20	0.90	4.04	1.71
Long-term return	-0.84	-1.60	0.10	1.37	-1.56	-0.39	-4.20	-1.20
Long-term yield	-1.73	-0.81	-0.87	2.80	12.80	14.32	16.72	8.27
Net equity expansion	-2.13	-0.48	-3.77	-1.21	-7.32	15.31	-4.19	3.66
Stock volatility	-3.55	1.84	-1.79	3.70	4.02	14.18	13.18	32.46
Term spread	-0.74	-0.42	0.70	0.70	**62.04**	127.85	38.91	26.67
Treasury-bill rate	-1.37	-0.72	0.51	1.36	51.27	101.05	**42.77**	33.98
Variance risk premium	**10.50**	10.48	**12.76**	12.08	-8.75	1.20	-17.07	-2.67

1. The approximate critical value comes from Clark and McCracken (2001) at a 5% significance level
In each column, the highest R^2 is blue and bold.

In table 11 and 12 are the results, and $\lambda_{original}$ is indeed worse in predicting the stock returns than the variable of Kelly and Jiang (2014), especially for the out-of-sample regressions. Therefore, the difference between the results might be due to the different data period.

To test if the size of the data sample influences the results in this paper, I also take a random sample of 500 respectively 2000 assets of the original data sample used by Kelly and Jiang (2014) each month and use this random sample to estimate λ_{500} and λ_{2000}. The original data sample contains around 2000 assets in 1963 to roughly 6000 assets around 2010. For the sample size of 500 assets, I indeed see that λ_{500} performs usually a bit worse than $\lambda_{original}$ in the predictive regressions (thought better in the out-of-sample regressions). So for the S&P-500 index, the smaller sample size might partly explain why our Hill estimator λ_{SP} is worse in predicting the stock returns than the estimator of Kelly and Jiang (2014). One objection, however, is that by randomly selecting a number of companies each month, most companies are only in the data sample for one month. Consequently, there are less observations per asset for the Fama-French regression, which makes this regression less reliable.

I also estimated ξ for the original data sample of Kelly and Jiang (2014), with all the assets from NYSE, AMEX or NASDAQ with share code 10 or 11, and with a random selection of 500 or 2000 assets each month from this original data sample. If I predict the returns in the next five years, the maximum likelihood estimator ξ still performs better in the out-of-sample regressions than the mean return, which was the same for ξ_{SP} and ξ_{Rus}. Furthermore, for this time horizon $\xi_{original}$ performs better than ξ_{2000}, and ξ_{2000} in turn performs better than ξ_{500}. However, in contrast with ξ_{SP} and ξ_{Rus}, $\xi_{original}$ and ξ_{2000} are not always better in predicting the stock returns in the next three years than the mean return, while ξ_{500} does perform better than the mean return.

Table 11: **Predictive regression results for λ and ξ**

Variable	One month			One year			Three years			Five years		
	β	T-stat.	R^2	β	T-stat.	R^2	β	T-stat.	R^2	β	T-stat.	R^2
$\lambda^1_{original}$	0.025	0.69	0.198	0.050	1.45	7.35	0.057	2.39	23.11	0.012	0.93	2.14
λ^2_{500}	-0.031	0.87	0.31	0.017	0.88	0.89	0.148	1.18	1.94	-0.061	-0.25	0.09
λ^2_{2000}	+0.00	0.01	0.00	0.040	3.30	4.52	0.481	7.01	19.27	0.308	2.03	2.22
$\xi^1_{original}$	0.015	0.42	0.07	0.01	0.19	0.11	-0.00	-0.01	0.00	-0.03	-1.25	14.81
ξ^2_{500}	0.016	0.44	0.08	0.00	-0.58	0.19	-0.016	-1.46	1.97	-0.016	-1.59	4.48
ξ^2_{2000}	0.00	0.09	0.00	0.00	0.07	0.01	-0.00	-0.12	0.07	-0.025	-1.09	10.22

1. Those variables are made with all the assets from NYSE, AMEX and NASDAQ with share code 10 or 11 and the dependent variable is the value-weighted return of CRSP
2. Those variables are made with a random selection of 500/2000 assets each month from all the assets from NYSE, AMEX and NASDAQ with share code 10 or 11 and the dependent variable is the value-weighted return of CRSP
All the coefficients β in this table are scaled so that they represent the expected increase (in numeric value) in the annualized future stock returns when the independent variable increases with one standard deviation. For the whole table, the data period is from 1996 to 04-2016.

Table 12: **Out-of-sample results for λ and ξ**

Variable	One month		One year		Three years		Five years	
π	1.03		0.93		0.73		0.53	
Crit. value[1]	1.584		1.584		1.584		1.079	
	R^2	ENC-New	R^2	ENC-New	R^2	ENC-New	R^2	ENC-New
$\lambda^2_{original}$	-0.42	0.01	-12.17	-0.91	-22.35	-0.17	-12.69	-3.07
λ^3_{500}	0.75	0.71	0.78	0.64	-0.16	0.04	-0.76	-0.14
λ^3_{2000}	-0.69	-0.35	-14.68	-2.74	-19.95	0.02	-8.88	-2.05
$\xi^2_{original}$	-0.44	-0.24	-14.65	-4.94	-19.40	-6.35	23.33	11.95
ξ^3_{500}	-0.53	-0.28	0.28	0.81	6.32	4.86	1.74	1.66
ξ^3_{2000}	-0.48	-0.29	-11.16	-3.81	-10.59	-3.56	12.79	6.76

1.The approximate critical value comes from Clark and McCracken (2001)
1. Those variables are made with all the assets from NYSE, AMEX and NASDAQ with share code 10 or 11 and the dependent variable is the value-weighted return of CRSP
2. Those variables are made with a random selection of 500/2000 assets each month from all the assets from NYSE, AMEX and NASDAQ with share code 10 or 11 and the dependent variable is the value-weighted return of CRSP. For the whole table, the data period is from 1996 to 04-2016.

7 Conclusion and discussion

The research question in the beginning of the paper was:

Which tail risk variable is the best predictor of returns on the stock market?

From the three tail risk variables, $LT(k)$ is the best in forecasting stock returns in both data samples. This tail risk variable is calculated with the option prices of out-of-the-money put options and the smallest time-to-maturity available in the data sample. The level of moneyness of the option is denoted by k in the brackets. In particular, for both large and small companies, $LT(0.95)$, $LT(0.85)$ and $LT(0.75)$ are the best in forecasting the returns one year or longer ahead. $LT'(0.95)$ is the best in forecasting the stock returns in the next month. This variable is, just like $LT(k)$, calculated with the option prices of out-of-the-money put options, but then with a time-to-maturity of 30 days. The results for all the tail risk variables were quite similar in both data samples, which makes the results more reliable. While $LT(k)$ was better in predicting the stock returns of the small companies than the ones of the large companies, $LT'(k)$ was better in predicting the returns (one month forward) for the large companies than for the small companies.

I also compared the tail risk variables to the benchmark variables of Welch and Goyal (2007) and the variance risk premium used by Bollerslev et al. (2009). $LT'(0.95)$ is still a good choice when predicting the stock returns in the next month. For the out-of-sample regressions, the R^2 is higher than the R^2 of the benchmark variables. For the prediction of the returns more than one month forward, the results are less clear. Sometimes $LT(k)$ is still the best in the out-of-sample predictions, but sometimes a benchmark variable has a higher R^2. For the in-sample, predictive regressions, there is always a benchmark variable that performs better than $LT(k)$ (i.e. has a higher R^2).

There are several reasons why I should be careful to generalize the results. First of all, the sample period covers only 20 years, 244 months in total. For the out-of-sample regressions, I used 120 months of those 244 months for the initialization period, leaving us with an even smaller sample. And when I predict the stock returns one year or longer ahead, the data sample becomes even smaller. Such a small sample makes the results less reliable. In future research, when more data are available, one could (partly) repeat this research to see if the results still holds. Future research can also test if the results holds in other (not necessarily longer) data samples.

Another reason to be careful to generalize the results are the different results found by Kelly and Jiang (2014). In their data sample, the variable λ was quite good in forecasting stock returns and outperformed many of the benchmark variables I used as well. That our results are so different indicates that the performance of λ is very sensitive to the data sample and period. In general, this could be an indication that the ability of variables to forecast stock returns is sensitive to the data sample and the sample period used, and that one cannot just generalize the results to other data samples. Moreover, I rejected the hypothesis that a higher tail risk forecasts a higher return in the next month for most variables. Since most other articles do not reject this hypothesis, the result is quite strange and it might also be an indication that you cannot just generalize the results to other situations. This is not the case for $LT(k)_{Rus}$ and $LT'(k)_{Rus}$ (the tail risk variable derived from option prices, calculated with the Russell-2000 index) and the higher levels of moneyness k.

Future research could also look at the ability of other variables to forecast stock returns. This article only looked at three tail risk variables and some standard benchmark variables. It would be interesting to compare those variables with all the other tail risk variables found in the literature, to see which one is really the best in forecasting the stock returns. Furthermore, this paper only looked at the linear regression from the return on a constant and one variable. Other articles often include two variables in the regressions, which would be interesting to evaluate as well.

Appendix

Appendix A: λ estimated for 1963-2010

When I estimated the predictive regression of the return of the S&P-500 index in the next month on λ_{SP}, I found a significantly negative coefficient. In other words, this result suggests that if the tail risk is higher, the return in the next month will be on average lower. This is not in line with the results of Kelly and Jiang (2014) and with the hypothesis in the introduction. Therefore, this part of the appendix will analyse whether the same result holds if I estimate λ_{SP} for the whole data period used by Kelly and Jiang (2014), namely from 1963 to 2010. Also, I will replicate the results of Kelly and Jiang (2014) to test whether the coefficients are positive. In figure 7 is a plot of the threshold u_{SP} used to calculate λ_{SP} and the threshold $u_{original}$. I used $u_{original}$ to estimate $\lambda_{original}$ for the period 1963-2010 with all assets from NYSE, AMEX and NASDAQ with share code 10 or 11, i.e. to estimate the Hill estimator λ for exactly the same data sample and data period as used by Kelly and Jiang (2014). From figure 7 it appears that the threshold of λ_{SP} is always higher than the threshold of $\lambda_{original}$, which is probably due to the size of the firms in both samples. It also appears that although $u_{original}$ is lower than u_{SP}, they both follow the same pattern.

Figure 8 is a plot of $\lambda_{original}$ and λ_{SP} from 1963 to 2010. λ_{SP} appears to be more "spiky" than $\lambda_{original}$, and follows a less clear pattern than $\lambda_{original}$.

Figure 7: Thresholds for $\lambda_{KellyandJian}$ and λ_{SP} of the whole data period (1963 to 2010)

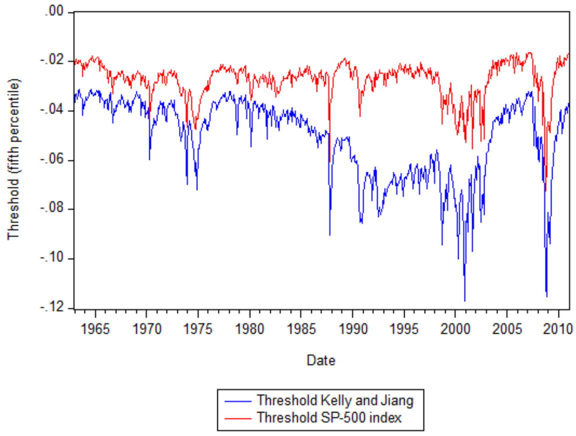
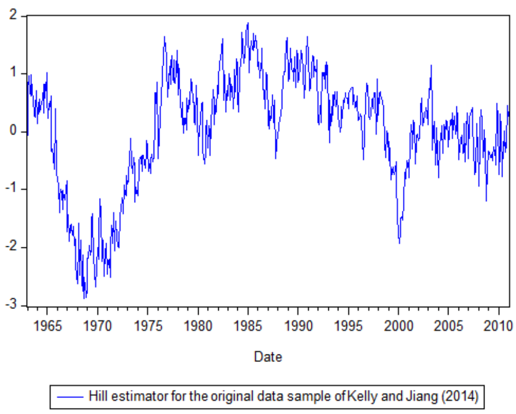
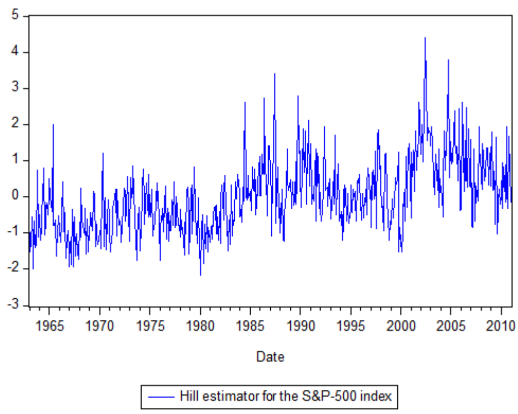


Figure 8: $\lambda_{original}$ and λ_{SP} from 1963 to 2010, both scaled



(a) $\lambda_{original}$



(b) λ_{SP}

In table 13 are the results of the predictive regression of λ_{SP} on the return of the S&P-500 index and on the value-weighted return of CRSP. To emphasize the comparison with the results of Kelly and Jiang (2014), the coefficients are now scaled to be the *percentage* increase in the annualized future stock returns instead of the numeric increase in the annualized future stock returns. In other words, they are a factor 100 larger than the coefficients in the previous tables. There is still a negative coefficient in the regressions on the return in the next month, but the result is not significantly different from zero any more. Furthermore, the R^2 is in both regressions quite low compared to the R^2 of the other variables in this article and to the R^2 of $\lambda_{original}$. This means that λ_{SP} is probably not a good variable to use to predict future stock returns, even if I extend the data period.

I also estimated the coefficients of the original regressions of Kelly and Jiang (2014) in table 13. For the return one month ahead, the coefficient and the R^2 are exactly the same and moreover, positive. This means that the negative coefficient of the regression with λ_{SP} is due to the data sample and not the data period. Furthermore, it indicates that $\lambda_{original}$ is (nearly) exactly the same as the λ found by Kelly and Jiang (2014). For the return of one year ahead or more, both the coefficients and the R^2 are higher than found by Kelly and Jiang (2014). A possible explanation is that I might have calculated the returns more than one year ahead differently. I took the simple returns of the value-weighted index including dividends, and calculated the return of more than one month ahead by the standard rule for calculating multi-period returns. Kelly and Jiang (2014) might have used the log of the returns instead, the value-weighted return without dividends or they might have directly calculated the return from the level of value-weighted index.

Table 13: **Result for $\boldsymbol{\lambda_{SP}}$ and λ for 1963-2010**

Variable	One month			One year			Three years			Five years		
	β	T-stat.	R^2	β	T-stat.	R^2	β	T-stat.	R^2	β	T-stat.	R^2
On the return of the S&P-500 index												
λ_{SP}	-1.52	-0.66	0.08	0.27	0.15	0.02	0.67	0.47	0.35	0.36	0.24	0.11
On the value-weighted return of CRSP												
λ_{SP}	-2.01	-0.84	0.12	-0.11	-0.06	0.00	0.29	0.19	0.06	-0.12	-0.08	0.01
$\lambda_{original}$	4.54	2.09	0.7	5.00	2.40	8.21	5.86	2.85	29.00	5.91	3.04	32.04

All the coefficients β in this table are scaled so that they represent the expected increase in *percentages* in the annualized future stock returns when the independent variable increases with one standard deviation. For the whole table, the data period is from 1963 to 2010.

Appendix B: Results for the value-weighted return of CRSP

In the paper, I found that the coefficient in the predictive regressions of the future monthly return on a tail risk variable is often negative. In this section, I test whether the negative coefficients are also negative when I regress the tail risk variables on the value-weighted return of CRSP, as used by Kelly and Jiang (2014). I did not include $LT(K)_{Rus}$ and $LT'(k)_{Rus}$, since those variables usually had positive coefficients, also for the regression on the return in the next month. The results of this regression are in table 14. However, for the regression on the return in the next month, the coefficients are still negative. The regression for the return in the next year gives (nearly always) a positive coefficient, which is in line with the results in the paper.

Table 14: **Result for the predictive regression with the value-weighted return of CRSP**

Variable	One month			One year		
	β	T-stat.	R^2	β	T-stat.	R^2
Hill's estimator λ						
λ_{SP}	-0.09	-2.43	2.39	-0.88	-0.40	0.22
λ_{Rus}	-0.035	-1.00	0.41	-0.029	-1.21	2.31
Maximum likelihood estimator ξ						
ξ_{SP}	-0.040	-1.24	0.64	-0.001	-0.04	0.00
ξ_{Rus}	-0.024	-0.67	0.19	-0.033	-0.73	3.36
Option price estimator LT(k) with a minimum time-to-maturity						
$LT(0.6)_{SP}$	-0.024	-0.68	0.19	0.045	1.61	6.10
$LT(0.65)_{SP}$	-0.024	-0.68	0.19	0.041	1.50	5.10
$LT(0.7)_{SP}$	-0.028	-0.79	0.25	0.038	1.43	4.34
$LT(0.75)_{SP}$	-0.028	-0.78	0.25	0.036	1.40	4.06
$LT(0.8)_{SP}$	-0.027	-0.76	0.24	0.037	1.39	4.07
$LT(0.85)_{SP}$	-0.024	-0.68	0.19	0.037	1.41	4.22
$LT(0.9)_{SP}$	-0.016	-0.44	0.08	0.037	1.38	4.12
$LT(0.95)_{SP}$	-0.002	-0.05	0.00	0.030	1.10	2.78
$LT(\bar{k})_{SP}$	-0.015	-0.42	0.07	0.034	1.25	3.64
$LT(k)^2_{SP}$	-0.024	-0.66	0.18	0.036	1.37	4.21
Option price estimator LT′(k) with a time-to-maturity of 30 days						
$LT'(0.6)_{SP}$	-0.057	-1.63	1.10	0.034	1.30	3.56
$LT'(0.65)_{SP}$	-0.046	-1.32	0.7	0.033	1.26	3.28
$LT'(0.7)_{SP}$	-0.038	-1.06	0.47	0.031	1.21	2.94
$LT'(0.75)_{SP}$	-0.30	-0.85	0.30	0.033	1.29	3.28
$LT'(0.8)_{SP}$	-0.027	-0.77	0.25	0.034	1.33	3.48
$LT'(0.85)_{SP}$	-0.020	-0.58	0.14	0.034	1.30	3.46
$LT'(0.9)_{SP}$	-0.017	-0.47	0.09	0.032	1.20	3.08
$LT'(0.95)_{SP}$	-0.011	-0.32	0.04	0.026	0.95	2.06
$LT'(\bar{k})_{SP}$	-0.021	-0.58	0.14	0.031	1.17	2.95

1. $LT(\bar{k})/LT'(\bar{k})$ is the mean of $LT(k)/LT'(k)$ for all levels of moneyness k

3. This $LT(k)$ is made with the mean of $LT(k)$ for all levels of moneyness k, except k=0.95, based on the in-sample performance.

All the coefficients β in this table are scaled so that they represent the expected increase (in numeric value) in the annualized future stock returns when the independent variable increases with one standard deviation. For the whole table, the data period is from 1996 to 04-2016.

References

Almeida, C., Ardison, K., Garcia, R., & Vicente, J. (2017). Nonparametric tail risk, stock returns, and the macroeconomy. *Journal of Financial Econometrics*, nbx007.

Andersen, T. G., Fusari, N., & Todorov, V. (2015). The risk premia embedded in index options. *Journal of Financial Economics*, *117*(3), 558–584.

Bali, T. G., Cakici, N., & Whitelaw, R. F. (2014). Hybrid tail risk and expected stock returns: when does the tail wag the dog? *The Review of Asset Pricing Studies*, *4*(2), 206–246.

Barro, R. J. (2006). Rare disasters and asset markets in the twentieth century. *The Quarterly Journal of Economics*, *121*(3), 823–866.

Bates, D. S. (1996). Jumps and stochastic volatility: exchange rate processes implicit in deutsche mark options. *The Review of Financial Studies*, *9*(1), 69–107.

Bates, D. S. (2000). Post-'87 crash fears in the s&p 500 futures option market. *Journal of Econometrics*, *94*(1), 181–238.

Bollerslev, T., Tauchen, G., & Zhou, H. (2009). Expected stock returns and variance risk premia. *The Review of Financial Studies*, *22*(11), 4463–4492.

Bollerslev, T. & Todorov, V. (2011). Tails, fears, and risk premia. *The Journal of Finance*, *66*(6), 2165–2211.

Bollerslev, T., Todorov, V., & Xu, L. (2015). Tail risk premia and return predictability. *Journal of Financial Economics*, *118*(1), 113–134.

Brodin, E. & Klüppelberg, C. (2008). Extreme value theory in finance. *Encyclopedia of quantitative risk analysis and assessment.*

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

Clark, T. E. & McCracken, M. W. (2001). Tests of equal forecast accuracy and encompassing for nested models. *Journal of econometrics*, *105*(1), 85–110.

Eraker, B. (2004). Do stock prices and volatility jump? reconciling evidence from spot and option prices. *The Journal of Finance*, *59*(3), 1367–1403.

Fama, E. F. & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, *33*(1), 3–56.

Gabaix, X. (2012). Variable rare disasters: an exactly solved framework for ten puzzles in macro-finance. *The Quarterly journal of economics*, *127*(2), 645–700.

Gao, G., Gao, P., & Song, Z. (2016). Do hedge funds exploit rare disaster concerns?

Gilli, M. & Këllizi, E. (2006). An application of extreme value theory for measuring financial risk. *Computational Economics*, *27*(2), 207–228.

Gopikrishnan, P., Plerou, V., Amaral, L. A. N., Meyer, M., & Stanley, H. E. (1999). Scaling of the distribution of fluctuations of financial market indices. *Physical Review E*, *60*(5), 5305.

Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *The annals of statistics*, *3*(5), 1163–1174.

Hodrick, R. J. (1992). Dividend yields and expected stock returns: alternative procedures for inference and measurement. *The Review of Financial Studies*, *5*(3), 357–386.

Huisman, R., Koedijk, K. G., Kool, C. J. M., & Palm, F. (2001). Tail-index estimates in small samples. *Journal of Business & Economic Statistics*, *19*(2), 208–216.

Jondeau, E. & Rockinger, M. (2003). Testing for differences in the tails of stock-market returns. *Journal of Empirical Finance*, *10*(5), 559–581.

Kearns, P. & Pagan, A. (1997). Estimating the density tail index for financial time series. *The Review of Economics and Statistics*, *79*(2), 171–175.

Këllezi, E. & Gilli, M. (2003). An application of extreme value theory for measuring risk. *Computational Economics*, *27*, 1–23.

Kelly, B. & Jiang, H. (2014). Tail risk and asset prices. *The Review of Financial Studies*, *27*(10), 2841–2871.

Longin, F. M. (2000). From value at risk to stress testing: the extreme value approach. *Journal of Banking & Finance*, *24*(7), 1097–1130.

McNeil, A. J. & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of empirical finance*, *7*(3), 271–300.

Pan, J. (2002). The jump-risk premia implicit in options: evidence from an integrated time-series study. *Journal of financial economics*, *63*(1), 3–50.

Rietz, T. A. (1988). The equity risk premium a solution. *Journal of monetary Economics*, *22*(1), 117–131.

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

Wachter, J. A. (2013). Can time-varying risk of rare disasters explain aggregate stock market volatility? *The Journal of Finance*, *68*(3), 987–1035.

Welch, I. & Goyal, A. (2007). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies*, *21*(4), 1455–1508.