Tail risk and asset prices in the short-term

Caio Almeida*[a], Gustavo Freire†[b], René Garcia‡[c], and Rodrigo Hizmeri§[d]

[a]Princeton University
[b]Erasmus University Rotterdam
[c]Université de Montréal
[c]Toulouse School of Economics
[d]University of Liverpool

February 5, 2023

Abstract

We combine high-frequency stock returns with risk-neutralization to extract the daily common component of tail risks perceived by investors in the cross-section of firms. We find that our tail risk measure significantly predicts the equity premium, variance risk premium and realized moments of market returns at short-horizons. Furthermore, a long-short portfolio built by sorting stocks on their recent exposure to tail risk generates abnormal returns with respect to standard factor models and helps explain the momentum anomaly. Incorporating investors' preferences via risk-neutralization is fundamental to our findings.

Keywords: Left tail risk, return predictability, factor models, risk-neutralization, high-frequency data.

JEL Classification: C58, G12, G17.

*E-mail: calmeida@princeton.edu, Department of Economics, Princeton University.

†E-mail: freire@ese.eur.nl, Erasmus School of Economics, Erasmus University Rotterdam.

‡E-mail: rene.garcia@umontreal.ca, Department of Economics, Université de Montréal, and Toulouse School of Economics.

§E-mail: r.hizmeri@liverpool.ac.uk, University of Liverpool Management School, University of Liverpool.

1 Introduction

Left tail risk is a pervasive feature of financial markets. As such, a large body of work has investigated its role in determining asset prices. Taken together, the empirical evidence indicates that compensation required by investors for bearing tail risk is fundamental to explain aggregate market risk premia and the cross-section of stock returns. This evidence is based on a number of different tail measures. In particular, information can be extracted either from stock prices (e.g., Bali et al., 2009; Kelly and Jiang, 2014; Weller, 2019), reflecting risk under the physical or statistical measure under which prices are observed, or from option prices (e.g., Andersen et al., 2015; Bollerslev and Todorov, 2011; Bollerslev et al., 2015), capturing tail risk under the risk-neutral measure incorporating investors' preferences.

In this paper, we propose a new tail measure available at a daily frequency, which allows us to investigate the short-term effects of tail risk on asset prices. We first estimate the common tail risk component of a cross-section of intra-day stock returns on day t, $\lambda_t^{\mathbb{P}}$, using the Hill (1975) power law estimator. This essentially adapts the tail index by Kelly and Jiang (2014) to a high-frequency environment. Then, we introduce a novel version of the Hill estimator, $\lambda_t^{\mathbb{Q}}$, that relies on risk-neutralized returns. More specifically, we apply a nonparametric adjustment to the pooled cross-section of stock returns on day t where "bad" states of nature, represented by negative observations, are overweighted to reflect investors' compensation for risk. The dynamics of the physical and risk-neutral Hill estimators differ substantially as compensation for risk varies over time.

Our approach overcomes two main challenges. First, extreme events are infrequently observed by definition. This limits the information available from the time series of a single asset such as the market index. Second, option maturities are relatively long compared to daily events, which makes it difficult to measure the tail risk specific to day t using option prices. By using high-frequency data on a large cross-section of stock returns, we are able to extract information about the level of tail risk at day t from the individual extreme events experienced by different stocks. Furthermore, our risk-neutralization allows to obtain a tail measure incorporating the economic valuation of tail risks by investors, which otherwise would only be possible using option prices (see, e.g., Aït-Sahalia and Lo, 2000).

Our empirical analysis is conducted considering each of our tail measures ($\lambda_t^{\mathbb{P}}$ and $\lambda_t^{\mathbb{Q}}$) in order to assess the information content of investors' economic valuation of tail risk. In particular, we compute the difference between the two estimators at a given point in time to capture the additional thickness of the left tail coming from the compensation demanded by investors for bearing tail risk. We call this difference the tail risk premium (TRP) and also investigate its implications for asset prices.

A distinctive feature of the tail measures we estimate is that they tend to decrease in periods of market distress. This contrasts with the usual increase that volatility-based risk measures exhibit during crises. In fact, the Hill estimator captures the thickness of the left tail after taking into account the effect of volatility, such that tail risk and volatility can move in different directions. This suggests that crisis periods are more often associated with bursts in volatility rather than more activity in the left tail.[1] Perhaps more surprising, the tail risk premium also decreases during financial crises, indicating that extreme negative events are more painful to investors in calm markets.[2] Intuitively, negative returns contain more information about unfavorable business conditions when they occur in periods of low volatility.

We start by examining the predictive relation between the tail measures and market risk premium in the short-term. We find that the tail risk premium positively predicts excess market returns, consistent with the idea that investors require higher returns to hold the market when compensation for bearing tail risk is higher. This relation is statistically significant and holds out-of-sample for weekly and monthly horizons, while there is no predictability at the daily horizon. Patterns are similar for $\lambda_t^{\mathbb{Q}}$, albeit with insignificant coefficients, whereas $\lambda_t^{\mathbb{P}}$ has no predictive power for the equity premium at the horizons considered. As for the variance risk premium, it is strongly predicted by $\lambda_t^{\mathbb{P}}$, $\lambda_t^{\mathbb{Q}}$ and TRP at all horizons, with a negative relation. This indicates that compensation

[1]This is consistent with the evidence by Christensen et al. (2014), Kelly and Jiang (2014) and Chapman et al. (2018), and the fact that realized kurtosis also decreases during financial crises.

[2]This is in line with contemporaneous evidence from option data in Schreindorfer and Sichert (2022).

for risk in the left tail and compensation for variance risk are inversely related once we control for changes in volatility in computing tail risk.

Tail risk also strongly predicts market risk for all short-term horizons considered. The predictive relation is negative for realized (and option-implied) volatility and realized skewness, while it is positive for realized kurtosis. This holds both for $\lambda_t^{\mathbb{P}}$ and $\lambda_t^{\mathbb{Q}}$, although the predictive power of the risk-neutral tail measure consistently dominates that of the physical one, in- and out-of-sample. In particular, bivariate regressions considering both $\lambda_t^{\mathbb{P}}$ and TRP often render the former insignificant, reaffirming the incremental information afforded by the risk-neutralization. In addition, we show that our tail measures, estimated from a cross-section of U.S. stocks, significantly predict international market realized variances. This suggests important cross-border contagion effects of tail risk.

We then investigate whether short-term tail risk is priced in the cross-section of stocks. To do so, for each of our tail measures, we build a long-short portfolio by sorting stocks each month on their recent exposure to the measure, based on contemporaneous daily regressions. The tail risk factors constructed from $\lambda_t^{\mathbb{Q}}$ and TRP generate statistically significant average returns that cannot be explained by standard factor models, where stocks with high exposure to tail risk (or tail risk premium) have high hedging capacity and are thus highly priced, yielding subsequent low returns. In contrast, the tail factor associated with $\lambda_t^{\mathbb{P}}$ leads to insignificant spreads in returns. In other words, only the short-term exposure to tail risk as perceived by investors (that is, its economic valuation) helps explain differences in expected returns across stocks.

Finally, in light of the fact that momentum remains a puzzle for standard factor models (Fama and French, 2016), we test if including our tail risk factor improves the explanatory power for this anomaly. In fact, Daniel and Moskowitz (2016) show that there is a crash risk component in momentum strategies, the compensation for which can potentially be captured by our factor. We find that the average return of the momentum strategy can be explained by its statistically significant loading on the risk-neutral tail risk factor. The same is not true considering the physical tail risk factor. This suggests that momentum is priced because it captures short-term exposure to tail risk.

The remainder of the paper is organized as follows. After a brief discussion of the related literature, Section 2 describes the methodology to construct our tail measures. Section 3 presents the data and the estimated tail measures, while Section 4 contains our empirical analysis. Section 5 concludes the paper.

1.1 Related literature

Our paper is mainly related to an evolving literature investigating the effects of left tail risk and investors' compensation for such risk on financial markets. Bollerslev and Todorov (2011), Bollerslev et al. (2015) and Andersen et al. (2015, 2017) provide evidence that tail risk is an important determinant of the equity and variance risk premia using option-implied tail measures. Extracting information from observed stock prices, Bali et al. (2009), Kelly and Jiang (2014), Almeida et al. (2017), Weller (2019) and Almeida et al. (2022) show that tail risk strongly predicts future market returns and macroeconomic activity. Computing tail risk at the firm-level, Bali et al. (2014), Chabi-Yo et al. (2018) and Atilgan et al. (2020) document significant cross-sectional relations between tail risk and future stock returns. International evidence on the effects of tail risk beyond the U.S. market is provided by Andersen et al. (2020), Andersen et al. (2021) and Freire (2021). We contribute to this literature by providing a novel method to estimate the tail risk specific to each day t. We document new short-term return predictability for the aggregate market and the cross-section of stocks, with particular focus on the role of incorporating investors' preferences towards tail risk.[3]

The closest work to ours is by Kelly and Jiang (2014), who propose the Hill estimator to estimate the common tail risk component of a cross-section of stocks at a monthly frequency. We adapt their estimator to a daily frequency using intra-day stock returns and put forward a new version of the Hill estimator based on risk-neutralized returns. We use both physical and risk-neutral estimators and their difference to study the relation between tail risk (and its economic valuation) and risk premia at horizons up to a month. In this context, we find that the economic perception of tail risk, as opposed to

[3]For an early contribution on the role of taking economic valuation into account for computing risk measures, see Aït-Sahalia and Lo (2000).

the physical tail risk, is an important determinant of the equity premium, variance risk premium and the cross-section of returns.

Also on a closely related work, Almeida et al. (2022) introduce a daily tail measure based on the expected shortfall of risk-neutralized intra-day market returns.[4] While we also use high-frequency data and risk-neutralization to estimate tail risk at a daily frequency, there are important differences between our approach and theirs. First, we extract information about the tail from extreme events of a cross-section of stocks, while they only consider the market index. Second, the tail measures in the two papers are inherently different, as their measure is closely related to volatility whereas ours is to higher moments such as kurtosis. In fact, their measure predicts excess market returns at the one-day horizon, while we find predictability for weekly and monthly horizons based on our measure. This suggests the tail measures offer complementary information. Third, while Almeida et al. (2022) focus on predicting market risk premia, we also investigate how short-term exposure to tail risk is priced in the cross-section of stocks.

Our paper is also related to the extensive literature identifying factors that are relevant to explain differences in the cross-section of stock returns, including Carhart (1997), Pástor and Stambaugh (2003), Fama and French (2015, 2016), among many others. Using our risk-neutral tail measure, we construct a tradable tail risk factor by sorting stocks based on their recent exposure to tail risk. This factor produces significant spreads in stock returns that cannot be explained by exposures to standard factors. We also show that our tail risk factor significantly explains the average returns of the momentum anomaly (Jegadeesh and Titman, 1993), offering a risk based explanation for momentum that is in line with previous evidence by Daniel and Moskowitz (2016). The risk-neutralization and daily frequency of our tail measure are fundamental to this finding.

[4] Almeida et al. (2022) build on the method by Almeida et al. (2017), who were the first to incorporate risk-neutralization in the estimation of tail risk.

2 Methodology

In this section, we describe the approach we take to estimate left tail risk at a daily frequency. Using a cross-section of intra-day stock returns, we first extract information about the common component of the tail risks of individual firms using the Hill estimator. Then, we introduce a new version of the Hill estimator that relies on risk-neutralized stock returns, thus further incorporating the investors' perception of risk in the estimation of extreme event risk. The difference between the two estimators at a given point in time captures the additional thickness of the left tail distribution that comes from the compensation demanded by investors for bearing tail risk. We call this difference the tail risk premium.

2.1 Hill estimator

Extreme events in financial markets are infrequent, even more so in a high-frequency environment. This makes it challenging to construct an aggregate measure of tail risk relying on a single asset such as the market index, since informative observations for the tail are rare by definition. To overcome this issue, we follow Kelly and Jiang (2014) by adopting a panel estimation approach capturing common tail behavior in the cross-section of individual stock returns. The identifying assumption is that the dynamics of the tail distributions of the firms are similar, so that extreme events in the cross-section allow us to extract the common component of their tail risk at each point in time.

More specifically, we assume that the left tail of the return distribution of asset i follows a power law structure.[5] That is, its day t conditional left tail distribution, defined as the set of extreme returns below some negative threshold u_t, obeys the following relation:

$$P(R_{i,t+1} < r | R_{i,t+1} < u_t \text{ and } \mathcal{F}_t) = \left(\frac{r}{u_t}\right)^{-a_i/\lambda_t}, \tag{1}$$

where $r < u_t < 0$ and $\mathcal{F}_t$ is the conditioning information set.[6] The parameter a_i/λ_t

[5] See Kelly and Jiang (2014) for a detailed motivation of the use of a power law structure to model the left tail distribution of returns. In sum, for a large class of heavy-tailed distributions, the left tail converges to a generalized power law distribution.

[6] $r < u_t < 0$ and $a_i/\lambda_t > 0$ guarantee that the probability $(r/u_t)^{-a_i/\lambda_t}$ is always between zero and

is the tail exponent which determines the shape of the tail distribution of asset i. The constant a_i may be different across assets in the cross-section, implying that they can have different levels of tail risk. However, their dynamics are driven by a common time-varying component, λ_t. The higher the λ_t, the fatter the stock returns' left tails and the higher the probabilities of extreme negative returns in the cross-section. Therefore, we refer to λ_t as our measure of aggregate tail risk.[7]

For each day t in our sample, we estimate the common tail risk component λ_t by applying the standard Hill (1975) power law estimator to our pooled cross-section of intra-day returns:

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

where $R_{k,t}$ is the kth high-frequency return that is below the threshold u_t on day t, K_t is the total number of returns that fall below this threshold within day t and the superscript $\mathbb{P}$ denotes that returns are observed under the physical probability measure.[8] The threshold u_t represents an extreme quantile determining that the observed returns below u_t belong to the left tail and follow the power law structure. We follow Kelly and Jiang (2014) by defining u_t to be the fifth percentile of the return cross-section for each period, which makes the threshold time-varying as the pooled return distribution changes from day to day.[9]

2.2 Risk-neutral Hill estimator

The Hill estimator extracts the common tail risk component from the pooled cross-section of returns observed under the physical probability measure, under which all observations are deemed equally likely to happen. In that sense, $\lambda_t^{\mathbb{P}}$ does not incorporate the true risks that are perceived by investors in financial markets. In particular, if investors are risk averse, then "bad" states of the world, represented by extreme negative

one.

[7] In extreme value theory, the parameter λ_t is also often called the shape parameter, and its inverse $1/\lambda_t$ the tail index (see, e.g., Danielsson, 2011).

[8] In the Hill formula, returns that fall below threshold u_t are treated as the first K_t entries of R_t. This is without loss of generality since in the pooled cross-section the elements of R_t are exchangeable.

[9] Our empirical results are qualitatively similar if we define u_t to be the first or tenth percentile.

observations in the cross-section, are overweighted to reflect compensation for risk. Since such observations are precisely the ones that matter for estimating tail risk, the economic perception of the left tail of returns may be underestimated by the physical Hill estimator. Moreover, its dynamics can also differ from that captured by $\lambda_t^{\mathbb{P}}$, as compensation for risk demanded by investors may vary over time depending on business conditions.

In order to incorporate investors' compensation for risk in the estimation of left tail risk, we propose a new version of the Hill estimator coupled with a nonparametric risk-neutralization algorithm. The idea is to tilt the physical measure such that risk in the pooled cross-section of stock returns is corrected for. This is possible by identifying a pricing kernel, or stochastic discount factor (SDF), that correctly prices the N pooled intra-day excess returns on a given day:

$$\frac{1}{N}\sum_{n=1}^{N} m_{n,t} R_{n,t} = 0, \tag{3}$$

where we normalize the mean of the SDF to be one ($\frac{1}{N}\sum_{n=1}^{N} m_{n,t} = 1$).[10] The pricing kernel corrects for risk as, after risk-neutralizing the excess returns with $\tilde{R}_{n,t} = m_{n,t} R_{n,t}$, there is no risk premium left. Equivalently, the expectation under the tilted risk-neutral probabilities $m_{n,t}/N$ of the excess returns in the pooled cross-section equals zero. We discuss how to identify the pricing kernel in the next subsection.

The risk-neutral Hill estimator $\lambda_t^{\mathbb{Q}}$ is then obtained by using the risk-neutralized pooled cross-section of returns in equation (2):[11]

$$\lambda_t^{\mathbb{Q}} = \frac{1}{K_t}\sum_{k=1}^{K_t} \ln \frac{\tilde{R}_{k,t}}{u_t}. \tag{4}$$

If investors are risk averse, then "bad" states of the world with negative return observations are associated with high marginal utility, i.e., with values of the pricing kernel above its mean one. Therefore, negative returns get properly overweighted reflecting compensation

[10]This implies an implicit gross risk-free rate of one such that we can treat the net stock returns in the cross-section as excess returns.

[11]That is, we posit that the left tail of the risk-neutral return distribution of each asset i in the cross-section also follows a power law structure.

demanded by investors for higher risk. In other words, the left tail of the risk-neutral distribution is thicker than that of the physical distribution. This naturally implies that $\lambda_t^{\mathbb{Q}} > \lambda_t^{\mathbb{P}}$, where the difference between the two estimators captures the additional tail thickness coming from the risk compensation demanded by investors for extreme negative events. Throughout the paper, we call this difference the tail risk premium ($TRP_t = \lambda_t^{\mathbb{P}} - \lambda_t^{\mathbb{Q}}$).

2.3 Risk-neutralization

The exact distortion of the physical measure, or risk correction of the pooled cross-section of returns, depends on the pricing kernel considered.[12] Besides correctly pricing the returns, there are two important properties that the SDF must satisfy. First, it must be nonnegative in order to be consistent with no-arbitrage. This guarantees that the tilted risk-neutral probabilities $m_{n,t}/N$ constitute a proper probability measure. Second, it should incorporate information about higher moments of the return distribution. This is important for modeling tail risk, since investors' aversion to downside risk is related to negative skewness aversion (see, e.g., Schneider and Trojani, 2015).

We follow the nonparametric approach developed by Almeida and Garcia (2017) to obtain a nonlinear pricing kernel satisfying the properties above. Their method consists in estimating SDFs minimizing a family of discrepancy loss functions (Cressie and Read, 1984) subject to correctly pricing a set of returns. This approach is a generalization of Hansen and Jagannathan (1991), who show how to obtain a minimum variance SDF from data on asset returns. Almeida and Garcia (2017) consider more general loss functions that take into account higher moments and imply nonnegative SDFs. Adapted to our context, the minimum discrepancy problem is given by:

$$\begin{aligned} &\min_{\{m_{1,t},\ldots,m_{N,t}\}} \frac{1}{N}\sum_{n=1}^{N} \frac{m_{n,t}^{\gamma+1}-1}{\gamma(\gamma+1)}, \\ &\text{s.t. } \frac{1}{N}\sum_{n=1}^{N} m_{n,t}R_{n,t} = 0,\ \frac{1}{N}\sum_{n=1}^{N} m_{n,t} = 1,\ m_{n,t} \geq 0\ \forall n, \end{aligned} \tag{5}$$

[12]We consider the realistic case of an incomplete market, where there exists an infinity of pricing kernels that correctly price the stock returns under no-arbitrage.

where the parameter $\gamma \in \mathbb{R}$ indexes the convex loss function in the Cressie and Read (1984) discrepancy family. This family captures as particular cases several loss functions in the literature, such as the Hansen and Jagannathan (1991) quadratic loss function when $\gamma = 1$ and the Kullback Leibler Information Criterion adopted by Stutzer (1995) when $\gamma \to 0$.

Under the assumption of no-arbitrage in the observed sample, Almeida and Garcia (2017) show that solving (5) is equivalent to solving the simpler dual problem below, for $\gamma < 0$:[13]

$$\lambda^*_\gamma = \arg\max_{\lambda \in \Lambda_\gamma} \frac{1}{N} \sum_{n=1}^{N} -\frac{1}{\gamma+1} \left(1 + \gamma\lambda R_{n,t}\right)^{\left(\frac{\gamma+1}{\gamma}\right)}, \tag{6}$$

where $\Lambda_\gamma = \{\lambda \in \mathbb{R} : \text{for } n = 1, ..., N,\ (1 + \gamma\lambda R_{n,t}) > 0\}$. The minimum discrepancy SDF can then be recovered from the first-order condition of (6) with respect to λ, evaluated at λ^*_γ:

$$m^*_{\gamma,n,t} = \left(1 + \gamma\lambda^*_\gamma R_{n,t}\right)^{\frac{1}{\gamma}}, \quad n = 1, ..., N. \tag{7}$$

For each γ, the solution λ^*_γ of the dual problem (6) leads to a different minimum discrepancy SDF. While by construction they all correctly price the pooled cross-section of returns, they do so by representing distinct risk preferences. In particular, Almeida and Freire (2022) show that positive absolute prudence (Kimball, 1990), which is related to aversion to downside risk and a convex marginal utility, is captured by $\gamma < 1$. Moreover, the smaller the γ, the more aversion to downside risk is embedded in the SDF, where the pricing kernel gets more convex putting more weight in extreme negative observations of returns.[14] They also show that, for extreme negative γs (usually below -5), the constrained maximization in the dual problem (6) may not have a solution. In order to successfully identify a pricing kernel capturing aversion to downside risk, we choose the one associated with $\gamma = -3$ to calculate the risk-neutral Hill estimator.[15]

[13]For $\gamma > 0$, the problem is unconstrained with an indicator function in the objective function: $\frac{1}{N}\sum_{n=1}^{N} -\frac{1}{\gamma+1}\left(1+\gamma\lambda R_{n,t}\right)^{\left(\frac{\gamma+1}{\gamma}\right)} I_{\Lambda_\gamma(R_{n,t})}(\lambda)$, where $\Lambda_\gamma(R_{n,t}) = \{\lambda \in \mathbb{R} : (1 + \gamma\lambda R_{n,t}) > 0\}$ and $I_A(x) = 1$ if $x \in A$, and 0 otherwise. For $\gamma = 0$, the problem is unconstrained and the objective function is exponential: $\frac{1}{N}\sum_{n=1}^{N} -e^{\lambda R_{n,t}}$.

[14]Since the mean of the pricing kernel continues to be the same, this means that less weight is given to intermediary return observations.

[15]Considering pricing kernels minimizing loss functions indexed by alternative γs associated with aversion to downside risk (such as -2 or -1) leads to similar conclusions.

3 Data description and implementation details

3.1 Data sources

Our sample consists of 5-minute returns for a panel of 100 stocks that were in the S&P 500 for the entire period between 2000 and 2020. The data is obtained from TickData Inc. Although our approach does not require a balanced panel, it does require liquid assets as otherwise the estimation may be impacted by the presence of zero returns (see, e.g., Bandi et al., 2020, 2017) and liquidity-related microstructure noise (see, e.g., Aït-Sahalia and Yu, 2009; Hansen and Lunde, 2006).[16] Therefore, we consider 100 highly liquid assets that were traded continuously over the sample period and work with a balanced panel for transparency and ease of exposition.

Throughout the paper, we utilize data on market returns, risk factors, and other uncertainty measures. The popular five factors of Fama and French (2015), the momentum factor and the risk-free rate are obtained from Kenneth French's website. The liquidity factor of Pástor and Stambaugh (2003) is available from Lubos Pastor's website. The VIX index and the left tail variation (LTV) proposed by Bollerslev et al. (2015) are respectively obtained from the Chicago Board Options Exchange (CBOE) and from the tailindex website, which is made available by Torben Andersen and Viktor Todorov. Finally, we construct measures of the variance, skewness, and kurtosis of the S&P 500 index using high-frequency returns sampled every 5 minutes also obtained from TickData Inc. In following Andersen et al. (2001, 2003) and Amaya et al. (2015), the realized variance (RV), realized skewness (RSK), and realized kurtosis (RK) take the following form:

[16]To formally exclude the impact of microstructure noise, we have performed the Hausman tests for microstructure noise and first-order serial correlation of Aït-Sahalia and Xiu (2019), for each stock and each day. The tests reject any significant presence of microstructure noise and first-order serial autocorrelation in the returns.

$$RV_t = \sum_{j=1}^{M} R_{t,j}^2, \tag{8}$$

$$RSK_t = \frac{\sqrt{M}\sum_{j=1}^{M} R_{t,j}^3}{RV_t^{3/2}}, \tag{9}$$

$$RK_t = \frac{M\sum_{j=1}^{M} R_{t,j}^4}{RV_t^2}, \tag{10}$$

where $R_{t,j}$ denotes the return on the S&P 500 index over the j-th intra-daily time interval on day t, and M denotes the number of high-frequency return intervals each day.

3.2 Risk-neutral estimates

As described in Section 2, we estimate the SDF for each day t using the pooled cross-section of intra-day stock returns. Sampling returns every 5 minutes based on 6.5 trading hours leads to 78 intra-day observations per stock per trading day, yielding a total of $N = 7,800$ observations by pooling the cross-section of 100 stocks. The estimation of the SDF is subject to the issue of outliers, as in their presence the risk-neutral distribution will allocate most of the probability weight to the (few) largest negative returns. To address this issue and guarantee a smooth risk-neutral measure, we winsorize 0.25% of the lowest and highest returns in the cross-section. We also impose the economic restriction of a 5% lower bound on the annualized equity premium, following Almeida and Freire (2022).[17] While results are similar compared to those where this restriction is not imposed, we keep it because it is economically sound to consider a lower bound on the equity premium (Campbell and Thompson, 2008; Martin, 2017; Pettenuzzo et al., 2014). Note that the restrictions above are only imposed for the estimation of the SDF.[18]

To illustrate, Figure 1 plots the estimated risk-neutral probabilities ($m_{\gamma,n,t}/N$) for various values of γ and the physical probabilities ($1/N$) for a random day in our sample.

[17] More specifically, for each day t, we impose that the average of the pooled returns in the cross-section is at least 5% above the risk-free rate, in annualized terms. That is, we shift the mean of the pooled return distribution to the lower bound when the bound is binding.

[18] That is, to compute the Hill estimator, we do not impose the 5% equity premium lower bound and we use all returns, including the ones that were categorized as outliers. The SDF weight for an outlier is set equal to the weight of the smallest (or highest) return not categorized as outlier.

The observed patterns are representative of other dates. As can be seen, the risk-neutral measures give more probability weight to negative returns and less to positive ones compared to the physical measure. This reflects agents' risk aversion: investors require more compensation (i.e., the SDF is higher) for "bad" states of the world. The relative compensation for risk in the left tail of the returns depends, in turn, on the aversion to downside risk. The smaller the γ, the more averse to downside risk (or, equivalently, the more prudent) is the investor and the greater are the weights to negative returns under the risk-neutral measure. As previously mentioned, we use the SDF associated with $\gamma = -3$ for the estimation of the risk-neutral Hill estimator.

3.3 Tail risk estimates

We estimate the tail risk measures $\lambda_t^{\mathbb{P}}$ and $\lambda_t^{\mathbb{Q}}$ as detailed in Section 2 using the set of intra-day return observations for all stocks for each day t. The upper panels of Figure 2 plot their one-month moving averages, for ease of exposition. The measures share similar dynamics, with a correlation of 83.1%. In particular, both measures tend to decrease in periods of market distress. This is in contrast to the usual increase that standard risk measures based on volatility exhibit during crises. To understand this pattern, the left lower panel of Figure 2 reports the time-varying threshold u_t (in absolute value), that determines where the left tail begins in the Hill estimator. As can be seen, u_t resembles a volatility measure, peaking during financial crises. The tail risk measures $\lambda_t^{\mathbb{P}}$ and $\lambda_t^{\mathbb{Q}}$ can thus be thought of as capturing the thickness of the left tail after taking into account the effect of volatility. In fact, as Kelly and Jiang (2014) note, a fixed percentile is used to define u_t exactly for this reason: if volatility increases but the shape of the return left tail is unchanged, an increase of the threshold (in absolute value) absorbs the effect of volatility changes and leaves estimates of the tail exponent unaffected.[19] Therefore, Figure 2 effectively shows that financial crises are more often associated with bursts in

[19]In unreported tests, we calculate $\lambda_t^{\mathbb{P}}$ and $\lambda_t^{\mathbb{Q}}$ with a constant threshold $u_t = u$ and find that both measures behave like volatility-type measures. This indicates that defining u_t as a fixed percentile of the return cross-section is instrumental to isolate the effects of volatility from the shape of the left tail.

volatility rather than more activity in the left tail.[20]

Even though the tail risk measures $\lambda_t^{\mathbb{P}}$ and $\lambda_t^{\mathbb{Q}}$ display similar dynamics, they are still fundamentally different. The right lower panel of Figure 2 plots the absolute value of the tail risk premium, $|TRP_t|$. By construction, the left tail of the pooled return distribution is always thicker under the risk-neutral measure than under the physical measure. However, the additional thickness of the tail coming from the risk compensation required by investors varies substantially over time. In particular, $|TRP_t|$ tends to be larger during normal times, and smaller during crisis periods. In other words, extreme negative returns are more painful to investors in calm markets. This is in line with contemporaneous evidence from option data by Schreindorfer and Sichert (2022). The intuition is that negative returns contain more information about unfavorable business conditions when they occur in periods of low volatility.

3.4 Comparison with other risk measures

Table 1 reports the correlation between $\lambda_t^{\mathbb{P}}$, $\lambda_t^{\mathbb{Q}}$, u_t, TRP_t and several risk measures. Both tail risk measures are negatively correlated with volatility measures (RV_t, VIX_t^2, VRP_t and LTV_t), whereas (the absolute value of) u_t is positively correlated with these measures. This is consistent with the fact that the time-varying threshold u_t controls for the effect of volatility in the calculation of $\lambda_t^{\mathbb{P}}$ and $\lambda_t^{\mathbb{Q}}$, as discussed in the previous subsection. In contrast, both tail risk measures are positively related to realized higher-order moments such as skewness and kurtosis. As for the TRP_t, it is negatively related to both $\lambda_t^{\mathbb{P}}$ and $\lambda_t^{\mathbb{Q}}$, indicating that the additional thickness of the tail coming from investors' risk compensation increases (i.e., TRP_t gets more negative) when tail risk increases.

Figure 3 provides further details on the relation between $\lambda_t^{\mathbb{Q}}$ and risk variables (the plots are similar for $\lambda_t^{\mathbb{P}}$). The upper panels make clear that tail risk, as measured as the shape parameter of the left tail of returns, is lower during periods of high volatility. On the other hand, the lower panels show that realized skewness and kurtosis co-move considerably with $\lambda_t^{\mathbb{Q}}$. This is especially true for kurtosis, which is often regarded as a

[20]This is in line with previous findings in the literature (see, e.g., Chapman et al., 2018; Christensen et al., 2014; Kelly and Jiang, 2014).

measure of tail thickness. As can be seen, like $\lambda_t^{\mathbb{Q}}$, realized kurtosis tends to be lower during periods of market distress. This suggests that measures of tail thickness, such as the ones we propose here, are more closely related to higher-order return moments like kurtosis than to second moments such as volatility.

4 Empirical results

This section provides empirical evidence of the information content of our tail risk measures for asset prices in the short-term. We document the predictive power of the tail measures in forecasting the equity premium, variance risk premium, U.S. and international market variances, as well as various risk measures. In particular, we perform daily ($h = 1$), weekly ($h = 5$), and monthly ($h = 22$) in- and out-of-sample forecasting exercises. The out-of-sample exercise is based on a rolling window of 500 days. In addition, we investigate how tail risk is priced in the cross-section of stocks. To do so, we construct monthly long-short portfolios by sorting stocks on their recent exposure to the tail risk measures.

4.1 Predicting excess market returns

Table 2 contains the forecasting results for the excess market returns, i.e., the equity premium. We report the coefficients and t-statistics that are computed using Newey-West robust standard errors with a lag length equal to the forecasting horizon h. All reported coefficients are scaled to be interpreted as the effect of a one standard deviation increase in the regressor on future excess market returns. The last two columns of each subpanel convey the out-of-sample R^2 for an unconstrained and equity constrained forecast. The equity constraint sets to zero all the forecasted values that are non-positive (see, e.g., Campbell and Thompson, 2008; Pettenuzzo et al., 2014). This constraint can be seen as a mild restriction, since it is difficult to consider an equilibrium framework where risk averse investors would hold the market if the expected compensation was negative.[21] The out-of-

[21]The use of economic constraints to sharpen market return forecasts is standard in the literature. While previous studies have also considered no-arbitrage restrictions (e.g., Ang and Piazzesi, 2003) and

sample R^2 is defined as $R^2_{oos} = 1 - \sum_t \left(r_{t+1:t+h} - \widehat{r}_{t+1:t+h|t}\right)^2 \Big/ \sum_t \left(r_{t+1:t+h} - \overline{r}_{t+1:t+h|t}\right)^2$, where $\widehat{r}_{t+1:t+h|t}$ is the out-of-sample forecast of the market return based solely on data through t, and $\overline{r}_{t+1:t+h|t}$ is the historical average return estimated through period t.

From the in-sample predictive regressions, we find a negative relation between the physical tail risk measure $\lambda_t^{\mathbb{P}}$ and the equity premium irrespective of the forecasting horizon. By contrast, the risk-neutral tail measure $\lambda_t^{\mathbb{Q}}$ positively predicts weekly and monthly excess market returns, consistent with a premium for bearing tail risk. While coefficients are insignificant in both cases, the discrepancies in their signs for $\lambda_t^{\mathbb{P}}$ and $\lambda_t^{\mathbb{Q}}$ suggest that only the economic perception of tail risk carries a premium. The third sub-panel, which reports the bivariate regression based on $\lambda_t^{\mathbb{P}}$ and the tail risk premium (TRP), corroborates these findings. Although both coefficients are negative, it is important to note that the TRP_t is negative by construction. Thus, as the wedge between the risk-neutral and physical tail risk increases (i.e., TRP_t becomes more negative), so does the future market return. In other words, investors require higher returns to hold the market at weekly and monthly horizons when compensation for bearing tail risk is higher. This relation is both economically and statistically significant.[22]

To further assess the information content of the tail risk premium, we compare the in-sample R^2 of the univariate and bivariate regressions. As can be seen, after adding the TRP_t onto the univariate regression of the physical tail risk, the weekly and monthly R^2 increase substantially, from 0.012% and 0.003% to 0.102% and 0.379%, respectively. In addition, the univariate (bivariate) regression shows a consistent decrease (increase) in performance as the forecasting horizon lengthens. The increasing predictive power of the tail risk premium for longer horizons is in line with the previous literature (see, e.g., Andersen et al., 2015, 2020; Bollerslev et al., 2015), and is supported by the fact that at shorter horizons the unpredictable and noisy component of returns dominates, whereas the predictable component emerges as the holding period increases (see. e.g., Andersen et al., 2020; Stambaugh, 1999).

slope coefficient constraints (e.g., Campbell and Thompson, 2008), we use the equity constraint because it is economically sound and easy to implement.

[22] These results are robust to controlling for several alternative predictors of excess market returns as shown in Table 8 in Appendix B.

We now focus on the out-of-sample performance, which is assessed by means of the R^2_{oos} and its constrained counterpart, $R^2_{oos,EC}$.[23] Beginning with the R^2_{oos}, we find that the physical tail risk has no predictive power for future excess market returns, regardless of the forecasting horizon. In contrast, the risk-neutral tail measures are able to beat the benchmark historical mean at the monthly horizon. For instance, the monthly R^2_{oos} of $\lambda_t^{\mathbb{P}}$, $\lambda_t^{\mathbb{Q}}$, and the bivariate regression $(\lambda_t^{\mathbb{P}} + TRP_t)$ are respectively -0.018%, 0.272%, and 0.354%. In other words, the inclusion of the TRP_t results in an increase of the R^2_{oos} of 37.20 basis points with respect to the univariate case with $\lambda_t^{\mathbb{P}}$. Turning to the equity constrained results, we see that imposing this economic restriction increases the out-of-sample R-squared for all predictors. Nevertheless, this improvement does not alter the ranking of the forecasts, as $\lambda_t^{\mathbb{Q}}$ and $(\lambda_t^{\mathbb{P}} + TRP_t)$ generate the highest $R^2_{oos,EC}$ for weekly and monthly forecasts, respectively.

In summary, our results indicate that the tail risk premium is an important determinant of the equity premium at weekly and monthly horizons. In particular, accounting for investors' aversion to downside risk in computing tail risk affords incremental information about future market returns above and beyond that contained in the physical tail risk.

4.2 Predicting variance and alternative risk measures

The temporal variation of volatility is usually associated with time-varying economic uncertainty.[24] In addition, the variance risk premium (VRP) captures investors' compensation for variance risk and is often regarded as a proxy for aggregate risk aversion (see, e.g., Bekaert et al., 2013; Campbell and Cochrane, 1999). Given that our tail risk measures capture extreme events associated with "bad" states of the economy, and our risk-neutralized measures further incorporate investors' aversion to downside risk, it is natural to assess their information content for predicting risk and compensation for risk.

Table 3 reports the forecasting results for the VRP and various risk measures of

[23]The use of the R^2_{oos} to evaluate the out-of-sample performance of candidate models is a popular choice in the literature (see, e.g., Campbell and Thompson, 2008; Kelly and Jiang, 2014; Pettenuzzo et al., 2014). Negative values indicate that the predictor performs worse than forecasts given by the historical mean.

[24]While we use the volatility and variance terms interchangeably, the results are based on predicting variances.

the S&P 500 index, such as the squared VIX (VIX^2), realized variance (RV), left tail variation (LTV), realized skewness (RSK), and realized kurtosis (RK).[25] As it is customary in the forecasting volatility literature (see, e.g., Andersen et al., 2007, 2003), we model the logarithmic variances.[26] Panels A, B, and C report the coefficients and t-statistics at the one-day ($h = 1$), one-week ($h = 5$), and one-month ($h = 22$) forecasting horizon, respectively. Coefficients are scaled to be interpreted as the effect of a one standard deviation increase on future market risk. The last two columns of each sub-panel report the in- and out-of-sample R^2.

The univariate predictive regressions in Table 3 indicate that both the physical and risk-neutral tail risk are strong predictors of variance risk and compensation for such risk, for all horizons considered. Even so, the predictive power of $\lambda_t^{\mathbb{Q}}$ consistently dominates that of $\lambda_t^{\mathbb{P}}$. In line with the discussion in Section 3.4, the coefficients are negative and strongly significant, suggesting an inverse relation between activity in the left tail and volatility. Considering the bivariate regressions, we further find that adding the TRP substantially increases the R_{is}^2 across all forecasting horizons. For instance, focusing on the results for the VIX^2, the R_{is}^2 of $\lambda_t^{\mathbb{P}}$ increases from 6% to 18% and from 3.8% to 14.6% for one-day and one-month horizons, respectively. Moreover, the inclusion of TRP_t renders $\lambda_t^{\mathbb{P}}$ insignificant for predicting monthly VRP and LTV, reaffirming the incremental information provided by the risk-neutralization. Results are similar out-of-sample, where the $\lambda_t^{\mathbb{Q}}$ and TRP_t are always superior to $\lambda_t^{\mathbb{P}}$ in terms of R_{oos}^2.

The results above highlight how the tail measures we construct are fundamentally different from existing ones in the literature. For instance, Bollerslev et al. (2015) demonstrate how the LTV is an important component of the VRP, where both are positively related. Furthermore, the tail measure of Almeida et al. (2022) positively predicts the VRP. In contrast, our measures, which capture the thickness of the left tail disentangled from the effects of volatility, strongly negatively predict the VRP. In this sense, our results suggest that compensation for risk in the left tail and compensation for variance

[25] The VRP is defined as the difference between the VIX^2 (adjusted to a monthly scale) and the realized variance over the previous 30 calendar days.

[26] Results for predictive regressions estimated on levels are qualitatively similar.

risk are inversely related once we control for changes in volatility in computing tail risk.

Turning to the higher-order moments, it can be seen that both physical and risk-neutral tail risk are important predictors of skewness and kurtosis, across all horizons. In particular, an increase in tail risk decreases future skewness: periods with more activity in the left tail are followed by market return distributions that are more skewed to the left. Moreover, tail risk positively predicts kurtosis: thicker left tails, as captured by our tail measures, are indicative of thicker tails in the market return distribution in subsequent periods. More specifically, a one-standard deviation increase in $\lambda_t^{\mathbb{P}}$ and $\lambda_t^{\mathbb{Q}}$ increases kurtosis at the monthly horizon by 2.93 and 3.68, respectively. The bigger magnitude of the $\lambda_t^{\mathbb{Q}}$'s coefficient is observed across all forecasting horizons, reflecting the incremental information afforded by incorporating the risk compensation required by investors. This also translates into both stronger coefficient significance and higher R-squared. For instance, the t-statistics and R_{is}^2 for the monthly forecasts of kurtosis are 4.35 and 6.34, and 3.41% and 5.38%, for $\lambda_t^{\mathbb{P}}$ and $\lambda_t^{\mathbb{Q}}$, respectively. The bivariate regressions further indicate that the TRP_t provides additional relevant information for future skewness and kurtosis, as its inclusion substantially increases the in- and out-of-sample performance relative to the physical tail risk.

Finally, we investigate whether our tail measures, based on a cross-section of U.S. stock returns, can also be useful to forecast the variances of international stock indices.[27] Table 4 conveys the results for predicting realized variances of the FTSE (United Kingdom), DAX (Germany), CAC-40 (France), STOXX (Eurozone), IBEX (Spain), Nikkei (Japan), and SSEC (Shanghai). The results reveal that both physical and risk-neutral U.S. tail risk are strong predictors of international market variances. This suggests our measures capture an important common tail risk component reflecting cross-border contagion and transmission of bad states of the economy across different countries.[28] Importantly, these relations are much stronger when incorporating investors' aversion to downside risk with the risk-neutralization. The R_{is}^2 of $\lambda_t^{\mathbb{Q}}$ is generally twice as large as that of its physical

[27]The realized variances of the international markets, obtained from the Oxford-Man Institute's Realized Library, are also estimated based on 5-minute returns.

[28]The high degree of association between the volatility of the S&P 500 and the international volatilities, with an average correlation of 72.9%, confirms that volatilities are driven by common components.

counterpart, across all forecasting horizons. Even more striking, the R^2_{is} of the bivariate regressions including TRP are often three times as large as those of the univariate regressions. In particular, the TRP_t is strongly significant and, for $h = 22$, subsumes the information contained in $\lambda_t^{\mathbb{P}}$ for all the indices with the exception of the DAX.

In summary, we document that our tail risk measures possess strong predictive power for the variance risk premium and several risk measures for the S&P 500 index,[29] as well as for a number of international market variances. This holds for daily, weekly and monthly forecasting horizons. Predictive performance is always stronger when the economic valuation of tail risk is taken into account with our risk-neutralized tail measures. In particular, the incremental information of the TRP_t generally subsumes the information content of the physical tail risk at monthly horizons.

4.3 Tail risk and the cross-section of stock returns

Hitherto, we have shown that the economic perception of tail risk by investors is an important determinant of aggregate market risk and risk premia at short-horizons. This section investigates whether recent exposure to tail risk is priced in the cross-section of stock returns through portfolio sorts. To do so, at the end of each month in our sample, we measure the insurance value of our 100 individual stocks with daily regressions over the previous 7 months, i.e., we estimate contemporaneous betas with respect to our tail measures: $R_{i,t} = \mu_i + \beta_i TR_{i,t}$, where $TR_{i,t} \in \{\lambda_t^{\mathbb{P}},\, \lambda_t^{\mathbb{Q}},\, TRP_t\}$.[30] Then, we form equally-weighted portfolios over the next month by sorting the 100 stocks into portfolios using quintile breakpoints calculated based on the given sorting variable.

The first three panels of Table 5 report the results for $\lambda_t^{\mathbb{P}}$, $\lambda_t^{\mathbb{Q}}$ and TRP_t, respectively. In addition to the average returns of the quintile portfolios, we also report the portfolios alphas (i.e., intercepts) from regressions of portfolio excess returns on the Fama-French three and five factors as well as extended models controlling for momentum (Carhart, 1997) and liquidity (Pástor and Stambaugh, 2003) factors. The last two columns report

[29] Tables 9, 10, 11 and 12 in Appendix B show that such predictive power is robust to controlling for lagged risk variables and several alternative predictors.

[30] Our results are robust to different estimation windows for the betas, such as 3, 5, 9, 12 and 24 months. These results are available upon request.

the average returns and alphas of the high minus low zero net investment portfolio and associated t-statistics, which are estimated using Newey-West robust standard errors. Panel D presents the p-values from various tests of the monotonicity (Patton and Timmermann, 2010) of average returns across the five quintile portfolios reported in Panels A–C. All tests have a null hypothesis of a flat pattern (no relation). While the MR Up and MR Down tests have alternative hypotheses of an increasing and decreasing pattern, the MR test alternative hypothesis is unrestricted. The tests are estimated using 10,000 bootstrap replications and a block length equal to 10 months.

Several conclusions can be drawn from these results. First, stocks that are more positively related to tail risk in the short-term earn lower returns. This is economically sound, as stocks with high β_i provide hedging opportunities against tail risk and are thus highly priced, yielding subsequent low returns. This relation is monotonic across quintile portfolios for $\lambda_t^{\mathbb{Q}}$, which is formally confirmed by the rejection of the flat pattern using the MR and MR Down tests. In contrast, a flat pattern cannot be rejected for $\lambda_t^{\mathbb{P}}$. For the TRP_t, there is a monotonic increasing pattern across quintile portfolios, which is due to its negative sign as $\lambda_t^{\mathbb{P}}$ is always smaller than $\lambda_t^{\mathbb{Q}}$. Stocks with more negative β_i with respect to TRP_t have higher insurance value for tail risk premium and are highly priced, with subsequent low returns. This relation is confirmed by the rejection of the MR and MR Up tests.

Second, exposure to physical tail risk generates insignificant average returns for the high minus low portfolio. On the other hand, the return spreads associated with exposures to $\lambda_t^{\mathbb{Q}}$ and TRP_t are both statistically and economically significant, where the corresponding high minus low strategies earn an average monthly return of -0.67% and 0.87%, respectively. This shows that recent exposure to tail risk as perceived by investors is strongly priced in the cross-section. To further illustrate, Figure 4 plots the cumulative returns of the quintile portfolios based on TRP_t. As can be seen, while the high portfolio performs reasonably well on its own, the robust profitability of the high minus low strategy is mainly driven by selling the stocks with higher hedging capacity for tail risk.

Third, the average high minus low returns in Table 5 are generally larger (in absolute value) after controlling for standard factor models in the literature. For instance, controlling for the Fama-French 3 factors results in an average monthly excess return of −1.11% (t-statistic −3.88) and 1.42% (t-statistic 4.44) for $\lambda_t^{\mathbb{Q}}$ and TRP_t, respectively. The reason is that the tail factors are negatively correlated with the market, size and value factors. This can be seen from Table 6, which provides further details on the regressions of our high minus low portfolio returns on factors models. Again, only the tail measures incorporating investors' preferences are able to generate statistically significant alphas. In particular, the large alphas of the high minus low $\lambda_t^{\mathbb{Q}}$ and TRP_t portfolios hold with significant factor exposure and high adjusted R^2. This suggests that our risk-neutral tail factors capture risk premium that is not reflected in firms' exposures to the market, size, value, profitability and momentum factors. By contrast, the tail factor based on $\lambda_t^{\mathbb{P}}$ holds no relation with standard factors, as it is only significantly exposed to the market factor and the adjusted R^2 of the regressions are low.

The results above unambiguously show that only the tail measures incorporating investors' preferences drive risk premium in the cross-section of stocks. To illustrate the differences between the physical and risk-neutral tail measures, Figure 5 plots, for each measure, the time series of the average β_i within each quintile portfolio and its difference between the high and low portfolios. During financial crises (e.g., the dot-com bubble, the global financial crisis and the Covid-19 pandemic), stocks' exposures to risk-neutral tail risk and tail risk premium generally increase, as would be expected.[31] In contrast, the β_is with respect to physical tail risk either decrease or fail to increase by the same magnitude. This suggests that the additional information content of the economic valuation of tail risk for the cross-section of returns is especially relevant during periods of market distress.

In sum, we find that the investors' perception of tail risk and the compensation for such risk in the short-term is strongly priced in the cross-section of stocks. High minus low portfolios based on the recent exposure to tail risk generate statistically and economically significant average returns, which are even larger after controlling for standard factor models. The information content of risk-neutralization beyond that contained in physical

[31] Note that the TRP_t is negative by construction and therefore a more negative β_i implies a higher sensitivity to tail risk premium.

tail risk is especially relevant during financial crises.

4.4 Tail risk factor and the momentum anomaly

Since Jegadeesh and Titman (1993), momentum has been one of the most widely studied anomalies in the cross-section of returns. Even so, there is still no consensus on how to explain it. As documented by Fama and French (2016), momentum remains one of the few anomalies for which predominant factor models such as Fama and French (2015) hold no explanatory power. More recently, Kelly et al. (2021) show that a sizable fraction of momentum can be explained by conditional risk exposure, as stocks' past performance can be seen as a noisy proxy for their time-varying loadings to priced factors. In this section, we alternatively investigate whether the momentum strategy remains profitable after controlling for its static exposure to our tail risk factors. Our motivation comes from the fact that there is a crash risk component in momentum strategies as they experience large negative returns during financial crises (Daniel and Moskowitz, 2016), such that the compensation for such risk can potentially be captured by our factors.

Table 7 conveys the regression results of the momentum high minus low returns on the Fama-French five factor model plus the liquidity factor of Pástor and Stambaugh (2003), as well as extended models including the tail risk factors. The first column shows that, in our sample, momentum generates a positive alpha over the Fama-French and liquidity factors, which is statistically significant at the 10% level. Further controlling for the physical tail risk factor does not help in explaining momentum, as noted by a larger alpha, an insignificant loading on the tail factor and the decrease in the adjusted R-squared. In contrast, after adding the tail risk factor based on $\lambda_t^{\mathbb{Q}}$ or TRP_t, the alpha of the momentum strategy becomes insignificant and even negative, and the adjusted R-squared increases substantially. In other words, the significant exposure of the momentum strategy to risk-neutral tail risk helps explain the spreads in returns it generates, which is aligned with our initial motivation.

In sum, we find that the risk premium associated with the momentum anomaly is in large part coming from the significant exposure of this strategy to risk captured by our tail factors. That is, short-term tail risk helps explain momentum. Importantly, this holds true only when investors' preferences are incorporated in the tail measures.

5 Conclusion

In this paper, we introduce a new tail risk measure at a daily frequency by combining high-frequency returns of a cross-section of stocks with a risk-neutralization algorithm. We use our measure to shed light on the effects of tail risk on asset prices at short-horizons and investigate to what extent these effects depend on information coming from the physical measure, under which asset prices are observed, and the risk-neutral measure, which incorporates investors' preferences.

We find that the compensation required by investors for bearing tail risk is an important determinant of the equity premium and the variance risk premium at horizons up to a month. Our tail measure also strongly predicts market risk, both for the U.S. and for international markets. In addition, tail risk is priced in the cross-section of stocks. A tradable tail factor built by sorting stocks on their recent exposure to tail risk produces significant spreads in stock returns that cannot be explained by standard factor models. Using our tail factor, we show that exposure of momentum strategies to tail risk helps explain the momentum anomaly. Incorporating investors' preferences in the estimation of tail risk is fundamental to our findings.

References

Aït-Sahalia, Y. and Lo, A. W. (2000). Nonparametric risk management and implied risk aversion. *Journal of Econometrics*, 94(1-2):9–51.

Aït-Sahalia, Y. and Xiu, D. (2019). A Hausman test for the presence of market microstructure noise in high frequency data. *Journal of Econometrics*, 211(1):176–205.

Aït-Sahalia, Y. and Yu, J. (2009). High frequency market microstructure noise estimates and liquidity measures. *The Annals of Applied Statistics*, pages 422–457.

Almeida, C., Ardison, K., Garcia, R., and Orlowski, P. (2022). High-frequency tail risk premium and stock return predictability. *Working Paper*.

Almeida, C., Ardison, K., Garcia, R., and Vicente, J. (2017). Nonparametric tail risk, stock returns, and the macroeconomy. *Journal of Financial Econometrics*, 15(3):333–376.

Almeida, C. and Freire, G. (2022). Pricing of index options in incomplete markets. *Journal of Financial Economics*, 144(1):174–205.

Almeida, C. and Garcia, R. (2017). Economic implications of nonlinear pricing kernels. *Management Science*, 63(10):3361–3380.

Amaya, D., Christoffersen, P., Jacobs, K., and Vasquez, A. (2015). Does realized skewness predict the cross-section of equity returns? *Journal of Financial Economics*, 118(1):135–167.

Andersen, T. G., Bollerslev, T., and Diebold, F. X. (2007). Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *The Review of Economics and Statistics*, 89(4):701–720.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Ebens, H. (2001). The distribution of realized stock return volatility. *Journal of Financial Economics*, 61(1):43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71(2):579–625.

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

Andersen, T. G., Fusari, N., and Todorov, V. (2017). Short-term market risks implied by weekly options. *The Journal of Finance*, 72(3):1335–1386.

Andersen, T. G., Fusari, N., and Todorov, V. (2020). The pricing of tail risk and the equity premium: evidence from international option markets. *Journal of Business & Economic Statistics*, 38(3):662–678.

Andersen, T. G., Todorov, V., and Ubukata, M. (2021). Tail risk and return predictability for the japanese equity market. *Journal of Econometrics*, 222(1):344–363.

Ang, A. and Piazzesi, M. (2003). A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics*, 50(4):745–787.

Atilgan, Y., Bali, T. G., Demirtas, K. O., and Gunaydin, A. D. (2020). Left-tail momentum: Underreaction to bad news, costly arbitrage and equity returns. *Journal of Financial Economics*, 135(3):725–753.

Bali, T. G., Cakici, N., and 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.

Bali, T. G., Demirtas, K. O., and Levy, H. (2009). Is there an intertemporal relation between downside risk and expected returns? *Journal of Financial and Quantitative Analysis*, 44(4):883–909.

Bandi, F. M., Kolokolov, A., Pirino, D., and Renò, R. (2020). Zeros. *Management Science*, 66(8):3466–3479.

Bandi, F. M., Pirino, D., and Renò, R. (2017). Excess idle time. *Econometrica*, 85(6):1793–1846.

Barndorff-Nielsen, O. E. and Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2(1):1–37.

Bekaert, G., Hoerova, M., and Duca, M. L. (2013). Risk, uncertainty and monetary policy. *Journal of Monetary Economics*, 60(7):771–788.

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

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

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

Campbell, J. Y. and Cochrane, J. H. (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of political Economy*, 107(2):205–251.

Campbell, J. Y. and Thompson, S. B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies*, 21(4):1509–1531.

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

Chabi-Yo, F., Ruenzi, S., and Weigert, F. (2018). Crash sensitivity and the cross section of expected stock returns. *Journal of Financial and Quantitative Analysis*, 53(3):1059–1100.

Chapman, D. A., Gallmeyer, M. F., and Martin, J. S. (2018). Aggregate tail risk and expected returns. *The Review of Asset Pricing Studies*, 8(1):36–76.

Christensen, K., Oomen, R. C., and Podolskij, M. (2014). Fact or friction: Jumps at ultra high frequency. *Journal of Financial Economics*, 114(3):576–599.

Cressie, N. and Read, T. R. (1984). Multinomial goodness-of-fit tests. *Journal of the Royal Statistical Society: Series B (Methodological)*, 46(3):440–464.

Daniel, K. and Moskowitz, T. J. (2016). Momentum crashes. *Journal of Financial Economics*, 122(2):221–247.

Danielsson, J. (2011). *Financial risk forecasting: the theory and practice of forecasting market risk with implementation in R and Matlab*. John Wiley & Sons.

Fama, E. F. and French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1):1–22.

Fama, E. F. and French, K. R. (2016). Dissecting anomalies with a five-factor model. *The Review of Financial Studies*, 29(1):69–103.

Freire, G. (2021). Tail risk and investors' concerns: Evidence from brazil. *The North American Journal of Economics and Finance*, 58:101519.

Hansen, L. P. and Jagannathan, R. (1991). Implications of security market data for models of dynamic economies. *Journal of Political Economy*, 99(2):225–262.

Hansen, P. R. and Lunde, A. (2006). Realized variance and market microstructure noise. *Journal of Business & Economic Statistics*, 24(2):127–161.

Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, pages 1163–1174.

Jegadeesh, N. and Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, 48(1):65–91.

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

Kelly, B. T., Moskowitz, T. J., and Pruitt, S. (2021). Understanding momentum and reversal. *Journal of Financial Economics*, 140(3):726–743.

Kimball, M. S. (1990). Precautionary saving in the small and in the large. *Econometrica*, pages 53–73.

Martin, I. (2017). What is the expected return on the market? *The Quarterly Journal of Economics*, 132(1):367–433.

Pástor, L. and Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy*, 111(3):642–685.

Patton, A. J. and Timmermann, A. (2010). Monotonicity in asset returns: New tests with applications to the term structure, the capm, and portfolio sorts. *Journal of Financial Economics*, 98(3):605–625.

Pettenuzzo, D., Timmermann, A., and Valkanov, R. (2014). Forecasting stock returns under economic constraints. *Journal of Financial Economics*, 114(3):517–553.

Schneider, P. and Trojani, F. (2015). Fear trading. *Swiss Finance Institute Research Paper*, (15-03).

Schreindorfer, D. and Sichert, T. (2022). Volatility and the pricing kernel. *Swedish House of Finance Research Paper*.

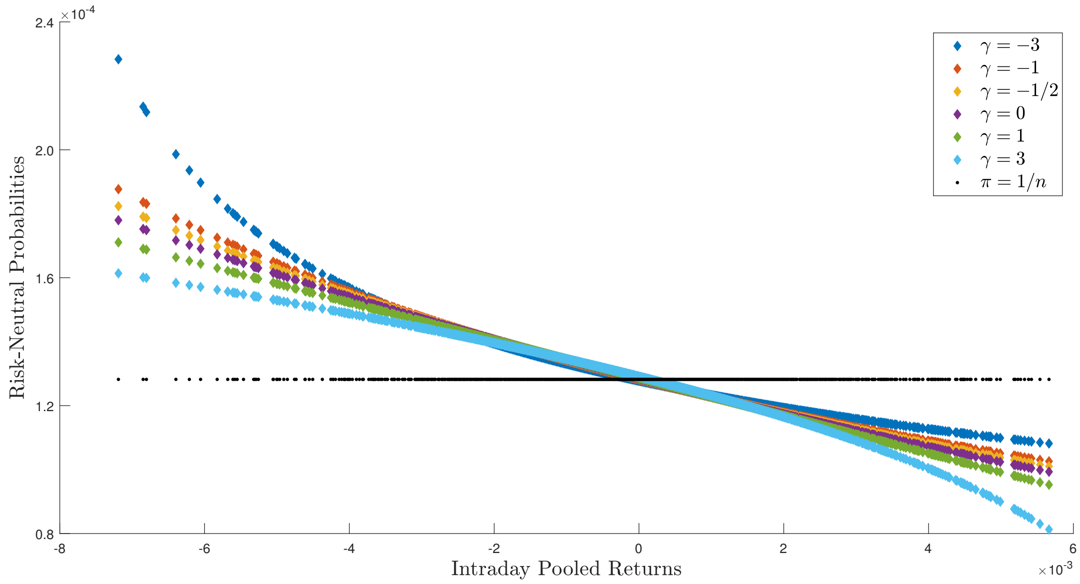
Stambaugh, R. F. (1999). Predictive regressions. *Journal of Financial Economics*, 54(3):375–421.

Stutzer, M. (1995). A Bayesian approach to diagnosis of asset pricing models. *Journal of Econometrics*, 68(2):367–397.

Weller, B. M. (2019). Measuring tail risks at high frequency. *The Review of Financial Studies*, 32(9):3571–3616.

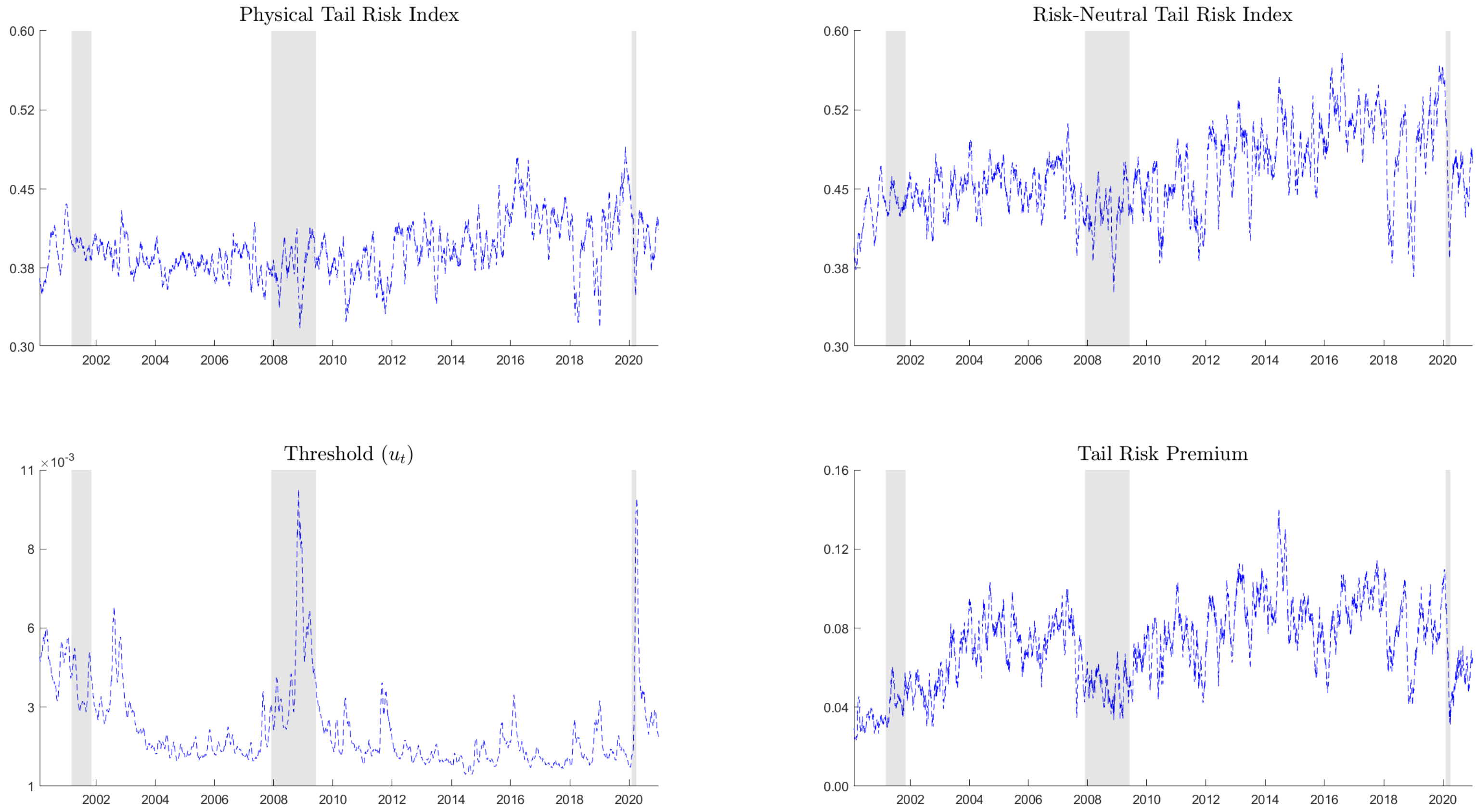
A Figures and Tables

Figure 1: Minimum dispersion risk-neutral probabilities



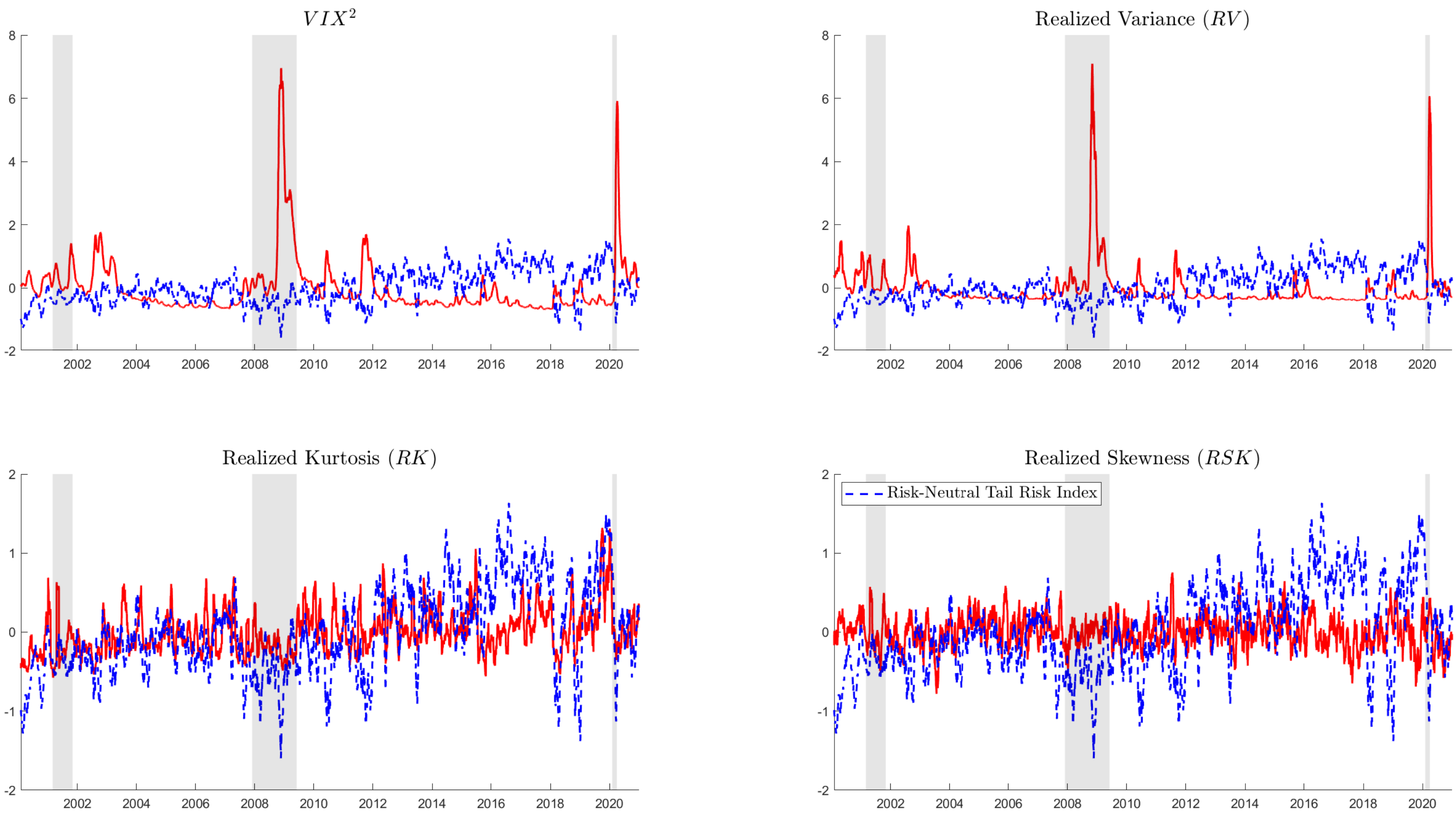
Note: The figure depicts the minimum dispersion risk-neutral probabilities for various values of γ and the physical measure ($\pi = 1/N$) for 78,000 intradaily returns sampled every 5 minutes for a random day in our sample.

Figure 2: Tail risk index measures



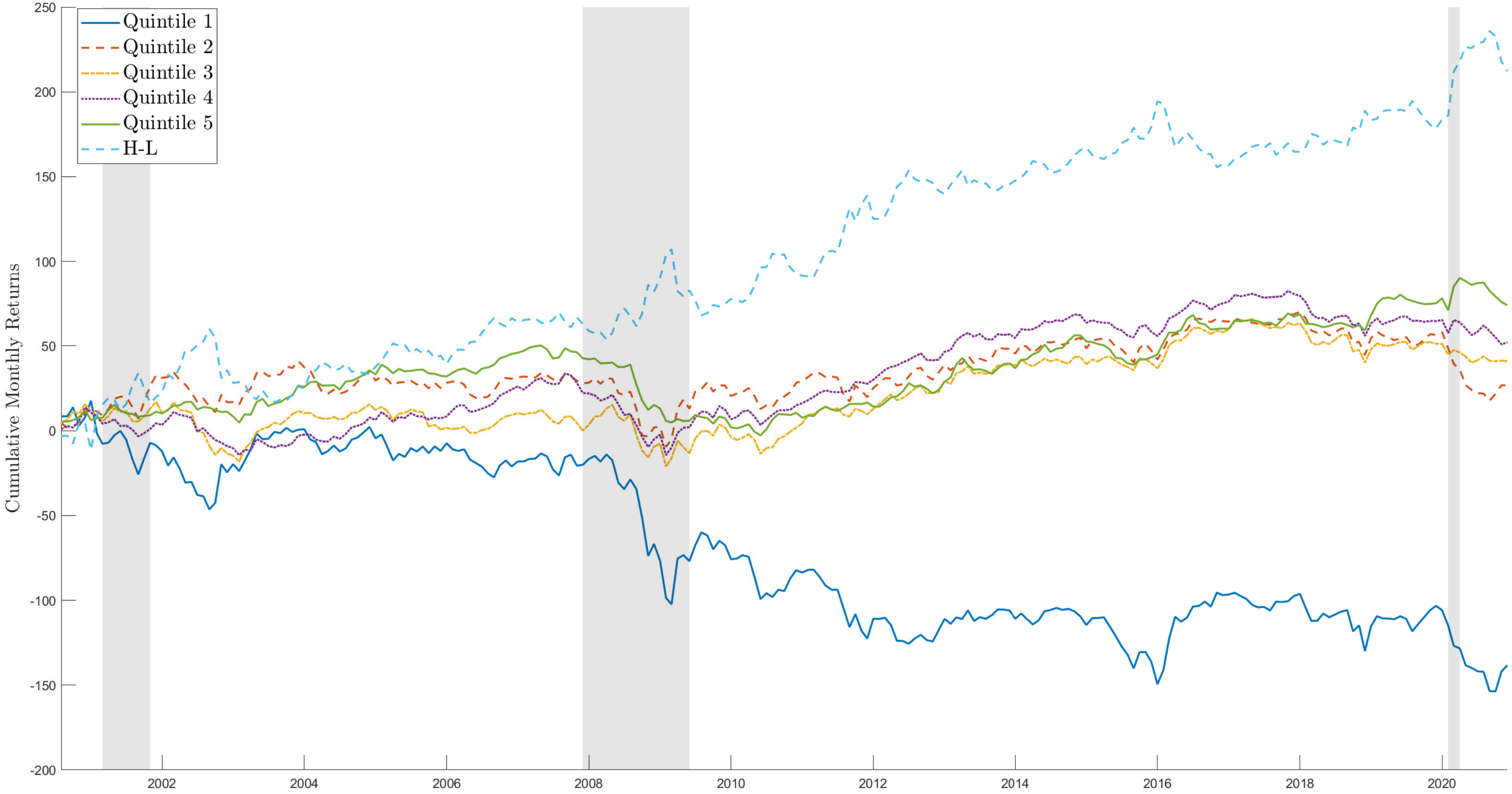
Note: The figure plots, in the upper panels, the 1-month moving average of the physical and risk-neutral tail risk indices. In the bottom panels, the corresponding moving averages for the threshold and the tail risk premium are depicted. For illustration purposes, the bottom panels plot the absolute value of both the threshold and the tail risk premium. Shaded areas depict NBER recession dates. The sample ranges from January, 2000 to December, 2020.

Figure 3: Tail risk index and risk measures



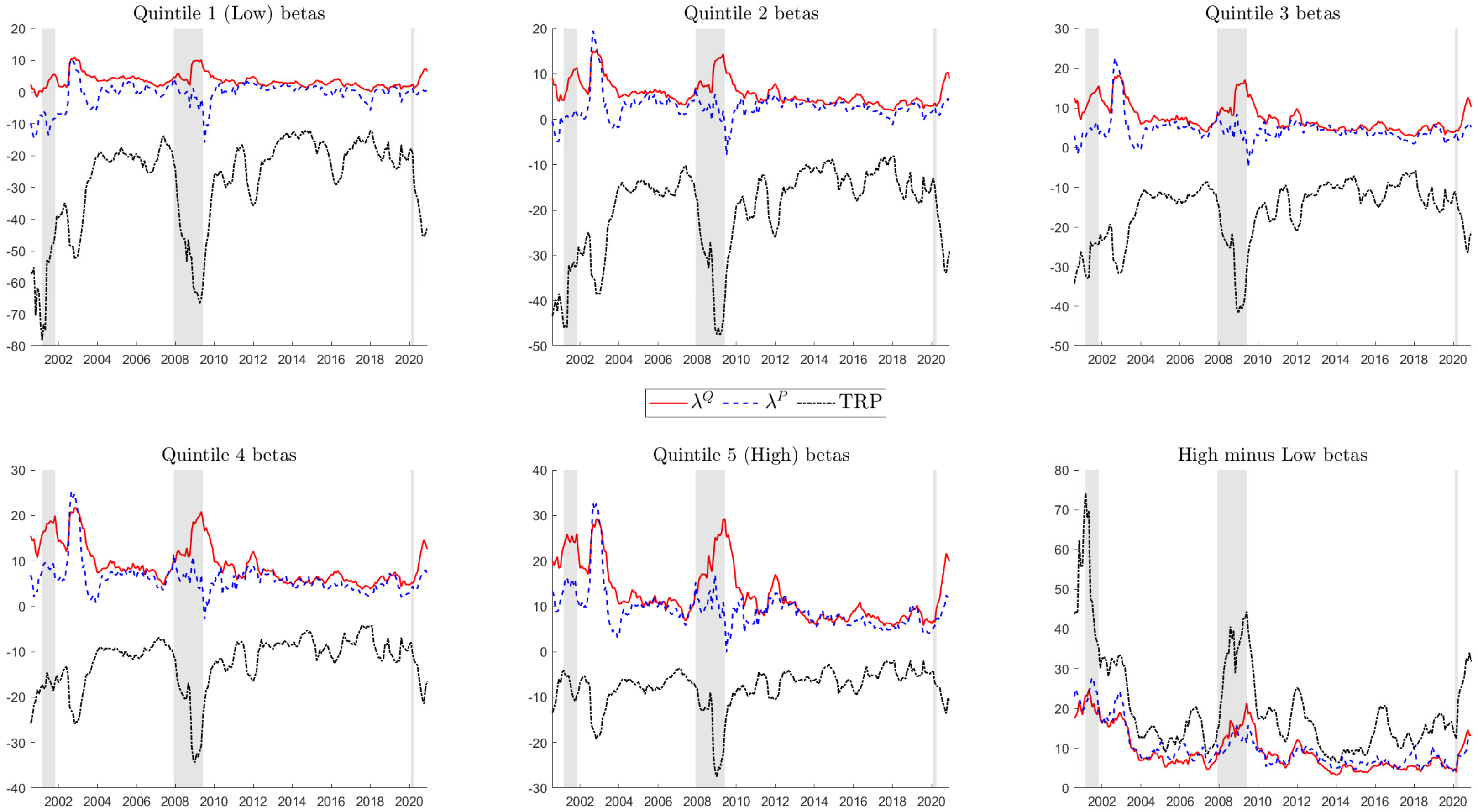
Note: The figure plots, in the upper panels, the 1-month standardized moving average of the squared VIX and realized variance of the S&P 500 index. Similarly, the bottom panels depict the corresponding stardandized moving averages for the realized kurtosis and realized skewness. For comparison, we also plot the 1-month standardized moving average of the risk-neutral tail risk index (blue dotted line). Shaded areas depict NBER recession dates. The sample ranges from January, 2000 to December, 2020.

Figure 4: Cumulative monthly quintile portfolio returns formed by sorting on TRP



Note: The figure depicts the cumulative monthly returns for each quintile portfolio and the high minus low zero net investment portfolio formed by sorting on the tail risk premium (TRP). Shaded areas depict NBER recession dates. The sample ranges from January, 2000 to December, 2020.

Figure 5: Portfolio quintile βs



Note: The figure depicts the time series average sensitivity to tail risk for all stocks within each quintile portfolio and the high minus low zero net investment portfolio. Shaded areas depict NBER recession dates. The sample ranges from January, 2000 to December, 2020.

Table 1: Correlation and AR(1) coefficients

	$\lambda_t^{\mathbb{P}}$	$\lambda_t^{\mathbb{Q}}$	TRP_t	$\lvert u_t \rvert$	RV_t	VIX_t^2	VRP_t	RSK_t	RK_t	LTV_t
$\lambda_t^{\mathbb{P}}$	0.398	0.831	−0.316	−0.202	−0.157	−0.215	−0.184	0.019	0.205	−0.167
$\lambda_t^{\mathbb{Q}}$		0.334	−0.790	−0.416	−0.228	−0.302	−0.252	0.266	0.229	−0.209
TRP_t			0.210	0.488	0.215	0.278	0.227	−0.433	−0.166	0.173
$\lvert u_t \rvert$				0.913	0.791	0.809	0.414	−0.081	−0.095	0.517
RV_t					0.691	0.756	0.398	0.044	0.080	0.539
VIX_t^2						0.971	0.606	−0.012	−0.084	0.723
VRP_t							0.588	−0.076	−0.062	0.520
RSK_t								−0.042	0.140	−0.018
RK_t									0.052	−0.069
LTV_t										0.932

Note: The table reports in the off-diagonal the correlation of each pair of variables, and in the main diagonal the AR(1) coefficient. The sample ranges from January, 2000 to December, 2020.

Table 2: Excess market return predictability

	One-day ($h = 1$)			One-week ($h = 5$)			One-month ($h = 22$)		
$\lambda_t^{\mathbb{P}}$	−0.021		−0.023	−0.022		−0.042	−0.024		−0.108
t-stat	−1.542		−1.587	−0.553		−1.023	−0.226		−0.844
$\lambda_t^{\mathbb{Q}}$		−0.014			0.015			0.123	
t-stat		−1.080			0.397			1.272	
TRP_t			−0.006			−0.064			−0.268
t-stat			−0.423			−1.695*			−2.524**
R^2_{is}	0.050	0.023	0.054	0.012	0.006	0.102	0.003	0.087	0.379
R^2_{oos}	−**0.246**	−0.249	−0.612	−0.221	**−0.139**	−0.522	−0.018	0.272	**0.354**
$R^2_{oos,EC}$	**0.033**	−0.025	−0.050	0.358	**0.361**	0.288	1.313	1.528	**1.854**

Note: The table reports the results for univariate and bivariate predictive regressions for the market excess returns over one-day ($h = 1$), one-week ($h = 5$), and one-month ($h = 22$). For forecasting horizons larger than 1 day, we consider the excess market returns from $t + 1$ to $t + h$. We compute the t-statistics using Newey-West robust standard errors with a lag length equal to h. The R^2_{is} is the OLS R-squared, and the last two rows report the out-of-sample R^2 for an unconstrained and equity constrained forecast. The equity constraint sets to zero all the forecasted values that are non-positive. The out-of-sample exercise is based on a rolling window of 500 days and $R^2_{oos} = 1 - \sum_t \left(r_{t+1:t+h} - \widehat{r}_{t+1:t+h|t}\right)^2 \Big/ \sum_t \left(r_{t+1:t+h} - \overline{r}_{t+1:t+h|t}\right)^2$, where $\widehat{r}_{t+1:t+h|t}$ is the out-of-sample forecast of the excess return based on only data through t, and $\overline{r}_{t+1:t+h|t}$ is the historical average estimated through period t. A negative R^2_{oos} implies that the prediction performs worse than setting the forecast equal to the sample mean. We denote with an EC (in the subscript) the out-of-sample R^2 corresponding to the equity constrained forecasts. *, ** and *** indicate statistical significance at respectively the 10%, 5% and 1% level. Bold numbers highlight the best out-of-sample performance in terms of the R^2_{oos} and $R^2_{oos,EC}$ for a given horizon. The sample ranges from January, 2000 to December, 2020.

Table 3: Risk predictability

	$\lambda_t^{\mathbb{P}}$	t-stat	R^2_{is}	R^2_{oos}	$\lambda_t^{\mathbb{Q}}$	t-stat	R^2_{is}	R^2_{oos}	$\lambda_t^{\mathbb{P}}$	t-stat	TRP_t	t-stat	R^2_{is}	R^2_{oos}
Panel A: One-day ($h = 1$)														
VIX^2	−0.185	−6.830***	6.020	1.900	−0.300	−13.675***	15.755	6.185	−0.099	−3.696***	0.275	14.413***	18.007	**6.210**
RV	−0.358	−19.117***	9.926	−0.636	−0.533	−31.525***	22.141	1.862	−0.217	−11.893***	0.447	22.314***	23.958	**2.097**
VRP	−0.117	−7.955***	2.302	2.165	−0.166	−10.904***	4.649	4.217	−0.076	−5.301***	0.130	9.028***	4.885	**5.863**
LTV	−0.108	−7.033***	2.536	3.112	−0.137	−9.407***	4.113	5.318	−0.079	−5.182***	0.090	7.227***	4.139	**5.640**
RSK	−0.044	−4.118***	0.301	−0.039	−0.062	−6.009***	0.604	**0.402**	−0.028	−2.508**	0.048	4.489***	0.633	0.194
RK	0.224	6.593***	0.782	0.623	0.267	8.064***	1.113	**0.641**	0.175	4.866***	−0.153	−4.375***	1.113	0.524
Panel B: One-Week ($h = 5$)														
VIX^2	−0.173	−6.014***	5.347	1.672	−0.287	−12.095***	14.663	5.975	−0.088	−3.122***	0.270	13.463***	17.053	**6.011**
RV	−0.298	−12.356***	7.951	−0.802	−0.468	−22.184***	19.739	2.126	−0.166	−7.242***	0.418	19.402***	22.142	**2.557**
VRP	−0.460	−5.633***	1.735	1.543	−0.686	−7.849***	3.855	3.429	−0.280	−3.428***	0.574	6.242***	4.166	**5.083**
LTV	−0.468	−4.693***	2.112	2.771	−0.589	−6.332***	3.355	4.487	−0.348	−3.523***	0.380	5.586***	3.369	**4.682**
RSK	−0.160	−4.909***	0.916	0.330	−0.181	−6.068***	1.175	**0.868**	−0.131	−3.811***	0.091	3.243***	1.186	0.695
RK	0.803	5.859***	1.664	2.113	1.058	8.724***	2.895	2.317	0.569	3.922***	−0.740	−6.661***	2.942	**2.532**
Panel C: One-Month ($h = 22$)														
VIX^2	−0.142	−4.567***	3.790	1.042	−0.252	−9.398***	11.897	4.561	−0.063	−2.050**	0.253	11.252***	14.572	**5.432**
RV	−0.218	−6.155***	4.865	−1.442	−0.376	−12.042***	14.481	0.921	−0.102	−2.982***	0.368	12.931***	17.381	**1.971**
VRP	−0.938	−2.400**	0.610	1.059	−1.624	−4.185***	1.827	1.367	−0.438	−1.064	1.594	3.846***	2.198	**3.969**
LTV	−1.221	−2.173**	0.929	1.482	−1.812	−3.476***	2.050	2.627	−0.744	−1.319	0.090	3.793***	2.211	**3.657**
RSK	−0.450	−3.482***	1.627	0.462	−0.466	−4.174***	1.748	1.095	−0.395	−2.881***	0.176	1.792*	1.851	**1.687**
RK	2.929	4.350***	3.413	4.853	3.676	6.375***	5.384	4.608	2.187	3.114***	−2.354	−5.549***	5.404	**5.719**

Note: The table reports the results for univariate and bivariate predictive regressions for the measures of risk of the S&P 500 index over one-day ($h = 1$), one-week ($h = 5$), and one-month ($h = 22$). We consider the squared VIX (VIX^2), the variance risk-premium (VRP), the left tail variation (LTV), the realized variance (RV), realized skewness (RSK), and realized kurtosis (RK). The latter three realized measures are constructed from high-frequency data sampled every 5-min. For forecasting horizons larger than 1 day, we aggregate the risk measures from $t+1$ to $t+h$. The t-statistics are computed using Newey-West robust standard errors with a lag length equal to h. The R^2_{is} is the OLS R-squared and the last column of each sub-panel reports the out-of-sample R^2. The out-of-sample exercise is based on a rolling window of 500 days and $R^2_{oos} = 1 - \sum_t \left(F_{t+1:t+h} - \widehat{F}_{t+1:t+h|t}\right)^2 \Big/ \sum_t \left(F_{t+1:t+h} - \overline{F}_{t+1:t+h|t}\right)^2$, where $\widehat{F}_{t+1:t+h|t}$ is the out-of-sample forecast of the risk measure based on only data through t, and $\overline{F}_{t+1:t+h|t}$ is the historical mean estimated through period t. A negative R^2_{oos} implies that the prediction performs worse than setting a forecast equal to the sample mean. *, ** and *** indicate statistical significance at respectively the 10%, 5% and 1% level. Bold numbers highlight the best out-of-sample performance in terms of the R^2_{oos} for each horizon. The sample ranges from January, 2000 to December, 2020.

Table 4: International realized variance predictability

Ticker	Country	$\lambda_t^{\mathbb{P}}$	t-stat	R^2_{is}	R^2_{oos}	$\lambda_t^{\mathbb{Q}}$	t-stat	R^2_{is}	R^2_{oos}	$\lambda_t^{\mathbb{P}}$	t-stat	TRP_t	t-stat	R^2_{is}	R^2_{oos}
Panel A: One-day ($h = 1$)															
FTSE	U.K.	−0.213	−12.642***	4.483	−1.048	−0.337	−21.373***	11.192	0.057	−0.118	−6.884***	0.303	16.603***	12.605	**0.421**
DAX	Germany	−0.227	−13.463***	4.782	0.254	−0.373	−23.645***	12.961	1.608	−0.117	−6.880***	0.349	18.952***	15.007	**2.031**
CAC-40	France	−0.212	−12.830***	4.702	−0.898	−0.352	−23.365***	12.983	1.264	−0.107	−6.433***	0.333	19.611***	15.168	**1.930**
STOXX	Eurozone	−0.216	−12.401***	4.275	−1.268	−0.336	−20.767***	10.314	−0.125	−0.123	−6.916***	0.296	16.163***	11.463	**0.238**
IBEX	Spain	−0.114	−7.049***	1.451	−2.033	−0.227	−15.428***	5.710	0.030	−0.036	−2.143**	0.250	15.888***	7.671	**0.792**
Nikkei	Japan	−0.194	−11.910***	4.295	0.251	−0.310	−20.468***	10.919	2.097	−0.106	−6.393***	0.281	17.430***	12.371	**2.492**
SSEC	Shanghai	−0.192	−12.165***	4.132	−0.508	−0.303	−20.518***	10.290	1.146	−0.107	−6.636***	0.271	17.311***	11.558	**1.634**
Panel B: One-Week ($h = 5$)															
FTSE	U.K.	−0.177	−8.258***	3.757	−1.092	−0.313	−16.242***	11.651	1.196	−0.079	−3.802***	0.314	16.753***	14.249	**2.213**
DAX	Germany	−0.187	−8.411***	3.810	1.302	−0.327	−16.253***	11.711	2.811	−0.085	−3.899***	0.325	16.627***	14.228	**3.606**
CAC-40	France	−0.170	−7.968***	3.554	−0.808	−0.301	−15.998***	11.151	1.449	−0.074	−3.553***	0.303	16.742***	13.689	**2.424**
STOXX	Eurozone	−0.177	−7.899***	3.709	−0.870	−0.296	−14.904***	10.368	0.252	−0.088	−3.971***	0.281	15.040***	12.160	**1.835**
IBEX	Spain	−0.079	−3.792***	0.807	−2.699	−0.187	−10.316***	4.461	−0.423	−0.007	−0.341	0.229	13.291***	6.821	**0.817**
Nikkei	Japan	−0.156	−7.218***	3.376	1.229	−0.276	−13.960***	10.487	3.910	−0.070	−3.282***	0.277	15.756***	12.823	**4.661**
SSEC	Shanghai	−0.152	−7.410***	3.298	0.382	−0.264	−14.243***	9.966	2.782	−0.070	−3.475***	0.261	15.916***	12.034	**3.722**
Panel C: One-Month ($h = 22$)															
FTSE	U.K.	−0.123	−3.937***	2.042	−1.536	−0.247	−8.766***	8.203	0.419	−0.037	−1.225	0.274	11.648***	11.117	**2.152**
DAX	Germany	−0.145	−4.462***	2.609	2.347	−0.276	−9.240***	9.408	3.334	−0.053	−1.698*	0.293	11.299***	12.171	**5.011**
CAC-40	France	−0.121	−3.888***	2.024	−0.750	−0.241	−8.750***	8.025	0.801	−0.037	−1.215	0.266	11.484***	10.809	**2.614**
STOXX	Eurozone	−0.123	−3.768***	2.082	−0.870	−0.232	−7.894***	7.379	0.252	−0.046	−1.425	0.245	9.857***	9.498	**1.835**
IBEX	Spain	−0.042	−1.333	0.251	−3.264	−0.135	−5.079***	2.633	−1.723	0.018	0.574	0.191	8.600***	4.948	**0.214**
Nikkei	Japan	−0.105	−3.228***	1.809	1.281	−0.214	−7.134***	7.514	3.519	−0.030	−0.933	0.241	10.027***	10.307	**5.306**
SSEC	Shanghai	−0.107	−3.514***	2.028	1.092	−0.208	−7.448***	7.576	3.173	−0.037	−1.242	0.223	9.963***	9.946	**5.123**

Note: The table reports the results for univariate and bivariate predictive regressions for the realized variance (RV) of international market indices over one-day ($h = 1$), one-week ($h = 5$), and one-month ($h = 22$). The RVs are constructed from high-frequency data sampled every 5-min. For forecasting horizons larger than 1 day, we aggregate the realized variance from $t + 1$ to $t + h$. The t-statistics are computed using Newey-West robust standard errors with a lag length equal to h. The R^2_{is} is the OLS R-squared and the last column of each sub-panel reports the out-of-sample R^2. The out-of-sample exercise is based on a rolling window of 500 days and $R^2_{oos} = 1 - \sum_t \left(F_{t+1:t+h} - \widehat{F}_{t+1:t+h|t}\right)^2 \Big/ \sum_t \left(F_{t+1:t+h} - \overline{F}_{t+1:t+h|t}\right)^2$, where $\widehat{F}_{t+1:t+h|t}$ is the out-of-sample forecast of the realized measure based on only data through t, and $\overline{F}_{t+1:t+h|t}$ is the historical mean estimated through period t. A negative R^2_{oos} implies that the prediction performs worse than setting a forecast equal to the sample mean. *, ** and *** indicate statistical significance at respectively the 10%, 5% and 1% level. Bold numbers highlight the best out-of-sample performance in terms of the R^2_{oos} for each horizon. The sample ranges from January, 2000 to December, 2020.

Table 5: Monthly sorted portfolios

	Low	2	3	4	High	High−Low	t-stat
Panel A: $\lambda_t^{\mathbb{P}}$							
Average Return	0.038	0.188	0.021	0.245	−0.266	−0.304	−1.135
CAPM alpha	−0.394	−0.194	−0.408	−0.214	−0.840	−0.446	−1.361
FF3 alpha	−0.423	−0.195	−0.405	−0.222	−0.868	−0.445	−1.492
FF5 alpha	−0.492	−0.363	−0.459	−0.278	−0.840	−0.348	−1.238
FF5 + Mom alpha	−0.488	−0.363	−0.455	−0.276	−0.836	−0.349	−1.239
FF5 + Mom + Liq alpha	−0.467	−0.334	−0.428	−0.259	−0.833	−0.367	−1.336
Panel B: $\lambda_t^{\mathbb{Q}}$							
Average Return	0.230	0.178	0.173	0.083	−0.438	−0.667	−2.179*
CAPM alpha	−0.092	−0.168	−0.258	−0.416	−1.117	−1.025	−3.397*
FF3 alpha	−0.074	−0.155	−0.260	−0.443	−1.181	−1.107	−3.875*
FF5 alpha	−0.276	−0.380	−0.383	−0.484	−0.914	−0.638	−2.226*
FF5 + Mom alpha	−0.278	−0.379	−0.377	−0.479	−0.906	−0.628	−2.104*
FF5 + Mom + Liq alpha	−0.258	−0.352	−0.349	−0.464	−0.897	−0.639	−2.143*
Panel C: TRP_t							
Average Return	−0.564	0.109	0.167	0.213	0.302	0.866	2.423*
CAPM alpha	−1.291	−0.401	−0.253	−0.135	0.029	1.320	3.931*
FF3 alpha	−1.361	−0.436	−0.254	−0.119	0.058	1.419	4.435*
FF5 alpha	−1.075	−0.444	−0.407	−0.295	−0.212	0.863	2.805*
FF5 + Mom alpha	−1.065	−0.439	−0.406	−0.295	−0.213	0.852	2.715*
FF5 + Mom + Liq alpha	−1.048	−0.431	−0.383	−0.266	−0.192	0.856	2.722*

Panel D: Monotonic Relation (MR) Test

	MR	MR Up	MR Down
$\lambda_t^{\mathbb{P}}$	0.781	0.213	0.056
$\lambda_t^{\mathbb{Q}}$	**0.031**	0.964	**0.036**
TRP_t	**0.005**	**0.011**	0.974

Note: The table reports the results of univariate portfolio analyses of the relation between the tail risk measures and the cross-section of returns. Monthly portfolios are formed by sorting the 100 stocks into portfolios using quintile breakpoints calculated based on the given sort variable using the 100 stocks. The table also reports portfolios alphas from regressions of portfolio excess returns using the Fama-French three and five factors as well as extended models controlling for momentum (Carhart, 1997) and liquidity (Pástor and Stambaugh, 2003) factors. Returns and alphas are in percentage terms. The last two columns report the high minus low zero net investment portfolio and associated t-statistics, which are estimated using Newey-West robust standard errors with a lag length equal to 5. The star (*) besides the t-statistic denotes statistical significance at the 5% or better. Panel D presents the p-values from various tests of the monotonicity (Patton and Timmermann, 2010) of average returns across the 5 quintile portfolios reported in Panels A–C. All tests have a null hypothesis of a flat pattern (no relation). While the MR Up and MR Down tests have alternative hypotheses of an increasing and decreasing pattern, the MR test is unrestricted. Bold p-values indicate significance at the 5% or better. The tests are estimated using 10,000 bootstrap replications and a block length equal to 10 months. The sample ranges from August, 2000 to December, 2020.

Table 6: High minus low tail risk factor regressions

	$\lambda^{\mathbb{P}}$			$\lambda^{\mathbb{Q}}$			TRP_t		
α	−0.445	−0.348	−0.367	−1.107	−0.637	−0.639	1.419	0.863	0.856
t-stat	−1.492	−1.237	−1.336	−3.875⋆	−2.226⋆	−2.143⋆	4.435⋆	2.805⋆	2.722⋆
MKT	0.245	0.207	0.192	0.506	0.327	0.222	−0.647	−0.434	−0.323
t-stat	2.950	2.412	2.077	5.867	3.495	2.525	−7.386	−4.826	−3.925
SMB	−0.042	−0.066	−0.111	0.505	0.373	0.351	−0.610	−0.464	−0.460
t-stat	−0.423	−0.629	−1.068	4.272	3.069	3.363	−5.416	−3.980	−4.264
HML	0.234	0.307	0.321	0.114	0.421	0.294	−0.162	−0.558	−0.407
t-stat	1.656	1.799	1.735	0.797	3.370	2.281	−0.913	−3.396	−2.467
RMW		−0.122	−0.154		−0.667	−0.520		0.736	0.552
t-stat		−0.695	−0.826		−3.400	−2.992		3.447	3.312
CMA		−0.109	−0.097		−0.338	−0.225		0.525	0.401
t-stat		−0.423	−0.355		−1.405	−1.094		1.728	1.586
Mom			0.014			−0.292			0.341
t-stat			0.153			−3.602			5.152
Liq			10.207			6.148			−2.166
t-stat			1.213			0.725			−0.271
R^2_{adj}	8.592	8.329	8.306	32.515	39.040	44.285	40.726	47.860	53.484

Note: The table reports the regression results of the high minus low tail factor on the Fama-French three and five factor models, as well as extended models controlling for momentum (Carhart, 1997) and liquidity (Pástor and Stambaugh, 2003) factors. The t-statistics are estimated using Newey-West robust standard errors with a lag length equal to 5, and a star (⋆) besides the α's t-statistic denotes statistical significance at the 5% or better. The sample ranges from August, 2000 to December, 2020.

Table 7: Momentum anomaly and the tail risk factor

	FF5 + Liq	FF5 + Liq+$\lambda^{\mathbb{P}}$	FF5 + Liq+$\lambda^{\mathbb{Q}}$	FF5 + Liq+TRP
α	0.022	0.028	−0.183	−0.270
t-stat	0.078	0.099	−0.551	−0.810
MKT	−0.328	−0.331	−0.227	−0.181
t-stat	−3.533*	−3.394*	−2.707*	−2.114*
SMB	−0.012	−0.010	0.100	0.145
t-stat	−0.081	−0.069	0.783	1.139
HML	−0.443	−0.449	−0.309	−0.255
t-stat	−3.088*	−2.901*	−2.241*	−1.980*
RMW	0.540	0.543	0.325	0.291
t-stat	2.441*	2.476*	1.522	1.396
CMA	0.365	0.367	0.260	0.188
t-stat	1.124	1.120	0.952	0.770
Liq	6.530	6.356	7.876	6.510
t-stat	0.685	0.689	0.817	0.716
Tail Risk Factor		0.017	−0.317	0.338
t-stat		0.154	−2.784*	3.305*
R^2_{adj}	25.339	25.042	31.965	33.671

Note: The table reports the regression results of the momentum factor on the Fama-French five factor models plus the liquidity factor (Pástor and Stambaugh, 2003), as well as extended models controlling for the tail risk factor. A star (*) denotes statistical significance at the 5% or better. The t-statistics are estimated using Newey-West robust standard errors with a lag length equal to 5. The sample ranges from August, 2000 to December, 2020.

B Robustness Results

Table 8: Market excess return predictability with control variables

	One-day ($h = 1$)			One-week ($h = 5$)			One-month ($h = 22$)		
$\lambda^{\mathbb{P}}$	−0.026		−0.026	−0.036		−0.057	−0.099		−0.159
t-stat	−1.855⋆		−1.769⋆	−0.898		−1.349	−0.812		−1.201
$\lambda^{\mathbb{Q}}$		−0.023			0.008			0.028	
t-stat		−1.585			0.183			0.238	
TRP			0.000			−0.082			−0.240
t-stat			−0.006			−1.862⋆			−1.777⋆
RV	−0.007	−0.009	−0.007	0.028	0.033	0.041	−0.311	−0.297	−0.274
t-stat	−0.252	−0.297	−0.248	0.353	0.412	0.502	−1.550	−1.499	−1.413
JV	−0.028	−0.029	−0.028	−0.033	−0.032	−0.030	0.022	0.025	0.032
t-stat	−1.055	−1.079	−1.056	−0.658	−0.642	−0.597	0.219	0.239	0.309
LTV	0.092	0.092	0.092	0.315	0.320	0.322	0.712	0.726	0.732
t-stat	4.089⋆⋆⋆	4.052⋆⋆⋆	4.062⋆⋆⋆	4.340⋆⋆⋆	4.397⋆⋆⋆	4.439⋆⋆⋆	3.134⋆⋆⋆	3.138⋆⋆⋆	3.209⋆⋆⋆
RSK	−0.002	0.004	−0.002	−0.008	−0.010	−0.043	−0.058	−0.063	−0.159
t-stat	−0.117	0.341	−0.104	−0.329	−0.361	−1.355	−1.160	−1.068	−1.989⋆⋆
RK	0.021	0.021	0.021	−0.031	−0.040	−0.037	0.020	−0.006	0.002
t-stat	1.452	1.413	1.446	−0.915	−1.205	−1.125	0.243	−0.073	0.025
VRP	−0.046	−0.045	−0.046	−0.218	−0.216	−0.219	−0.373	−0.369	−0.378
t-stat	−2.054⋆⋆	−2.027⋆⋆	−2.052⋆⋆	−2.808⋆⋆⋆	−2.793⋆⋆⋆	−2.841⋆⋆⋆	−1.727⋆	−1.713⋆	−1.749⋆
R^2_{is}	0.912	0.890	0.912	2.233	2.202	2.354	3.079	3.020	3.350

Note: The table reports the multivariate predictive regressions for the market excess returns over one-day ($h = 1$), one-week ($h = 5$), and one-month ($h = 22$). For forecasting horizons larger than 1 day, we consider the excess market returns from $t + 1$ to $t + h$. We compute the t statistics using Newey-West robust standard errors with a lag length equal to h. The control variables are the realized variance (RV), the jump variation (JV), the left tail variation (LTV) of Bollerslev et al. (2015), realized skewness (RSK), realized kurtosis (RK), and the variance risk premium (VRP) (Bollerslev et al., 2009). JV is estimated as $JV = (RV - BV)^+$, and BV is the bipower variation measure proposed by Barndorff-Nielsen and Shephard (2004). ⋆, ⋆⋆ and ⋆⋆⋆ indicate statistical significance at respectively the 10%, 5% and 1% level. The sample ranges from January, 2000 to December, 2020.

Table 9: Predicting risk measures with $\lambda_t^{\mathbb{P}}$ and control variables

	F_t	t-stat	JV_t	t-stat	$\lambda_t^{\mathbb{P}}$	t-stat	R^2_{is}
Panel A: One-day ($h = 1$)							
VIX^2	0.606	7.726***	−0.011	−0.496	−0.066	−4.678***	67.298
RV	0.610	10.760***	−0.064	−1.198	−0.269	−16.814***	35.249
LTV	0.623	85.683***	0.012	2.597***	−0.003	−0.746	86.822
RSK	−0.034	−2.913***	0.020	1.732*	−0.042	−3.981***	0.517
RK	0.143	3.609***	−0.096	−1.881*	0.192	5.565***	1.087
Panel B: One-week ($h = 5$)							
VIX^2	0.593	7.424***	−0.005	−0.228	−0.056	−3.803***	65.379
RV	0.573	8.478***	−0.055	−1.104	−0.215	−11.082***	34.066
LTV	2.776	40.028***	0.090	3.404***	0.000	0.012	76.605
RSK	−0.083	−3.799***	0.039	2.038***	−0.157	−4.801***	1.192
RK	0.619	4.941***	−0.590	−3.996***	0.658	4.921***	2.880
Panel C: One-month ($h = 22$)							
VIX^2	0.547	6.724***	0.003	0.141	−0.034	−1.949*	58.042
RV	0.492	6.459***	−0.034	−0.680	−0.146	−5.116***	27.385
LTV	9.669	21.064***	0.806	2.038**	0.420	2.230**	60.921
RSK	−0.008	−0.164	0.115	2.038**	−0.446	−3.456***	1.732
RK	1.990	4.818***	−2.055	−4.290***	2.458	3.845***	5.515

Note: The table reports the multivariate predictive regressions for the measures of risk of the S&P 500 index over one-day ($h = 1$), one-week ($h = 5$), and one-month ($h = 22$). We consider the squared VIX (VIX^2), realized variance (RV), left tail variation (LTV), realized skewness (RSK), and realized kurtosis (RK). For forecasting horizons larger than 1 day, we aggregate the risk measures from $t + 1$ to $t + h$. The t-statistics are computed using Newey-West robust standard errors with a lag length equal to h. F_t, in the first column, denotes the lagged value of the variable of interest. JV is estimated as $JV = (RV - BV)^+$, and BV is the bipower variation measure proposed by Barndorff-Nielsen and Shephard (2004). *, ** and *** indicate statistical significance at respectively the 10%, 5% and 1% level. The sample ranges from January, 2000 to December, 2020.

Table 10: Predicting risk measures with $\lambda_t^{\mathbb{Q}}$ and control variables

	F_t	t-stat	JV_t	t-stat	$\lambda_t^{\mathbb{Q}}$	t-stat	R^2_{is}
Panel A: One-day ($h = 1$)							
VIX^2	0.581	7.581***	−0.008	−0.383	−0.126	−8.483***	69.109
RV	0.546	10.577***	−0.045	−0.918	−0.413	−25.479***	42.252
LTV	0.622	86.019***	0.012	2.572**	−0.006	−1.516	86.828
RSK	−0.019	−1.570	0.016	1.405	−0.056	−5.038***	0.686
RK	0.123	3.132***	−0.081	−1.657*	0.233	6.916***	1.333
Panel B: One-week ($h = 5$)							
VIX^2	0.568	7.279***	−0.002	−0.099	−0.116	−7.356***	67.041
RV	0.515	8.413***	−0.037	−0.823	−0.354	−18.952***	40.675
LTV	2.775	40.111***	0.090	3.402***	−0.002	−0.087	76.605
RSK	−0.040	−1.664*	0.029	1.602	−0.169	−5.260***	1.248
RK	0.522	4.349***	−0.524	−3.990***	0.906	7.816***	3.787
Panel C: One-month ($h = 22$)							
VIX^2	0.525	6.572***	0.006	0.334	−0.094	−5.102***	59.346
RV	0.442	6.437***	−0.018	−0.406	−0.277	−10.329***	32.659
LTV	9.650	21.140***	0.814	2.029***	0.253	1.636	60.852
RSK	0.117	1.759*	0.084	1.656*	−0.492	−4.057***	1.928
RK	1.689	4.337***	−1.841	−4.503***	3.175	5.911***	6.959

Note: The table reports the multivariate predictive regressions for the measures of risk of the S&P 500 index over one-day ($h = 1$), one-week ($h = 5$), and one-month ($h = 22$). We consider the squared VIX (VIX^2), realized variance (RV), left tail variation (LTV), realized skewness (RSK), and realized kurtosis (RK). For forecasting horizons larger than 1 day, we aggregate the risk measures from $t + 1$ to $t + h$. The t-statistics are computed using Newey-West robust standard errors with a lag length equal to h. F_t, in the first column, denotes the lagged value of the variable of interest. JV is estimated as $JV = (RV - BV)^+$, and BV is the bipower variation measure proposed by Barndorff-Nielsen and Shephard (2004). *, ** and *** indicate statistical significance at respectively the 10%, 5% and 1% level. The sample ranges from January, 2000 to December, 2020.

Table 11: Predicting risk measures with TRP_t and control variables

	F_t	t-stat	JV_t	t-stat	$\lambda_t^{\mathbb{P}}$	t-stat	TRP_t	t-stat	R^2_{is}
Panel A: One-day ($h = 1$)									
VIX^2	0.575	7.676***	−0.008	−0.402	−0.031	−2.291**	0.129	8.229***	69.766
RV	0.536	10.557***	−0.044	−0.913	−0.171	−11.360***	0.345	19.631***	43.291
LTV	0.622	86.063***	0.012	2.538***	−0.001	−0.246	0.007	1.769*	86.830
RSK	−0.016	−1.227	0.015	1.343	−0.030	−2.672***	0.040	3.142***	0.693
RK	0.123	3.128***	−0.082	−1.659*	0.154	4.240***	−0.134	−3.817***	1.333
Panel B: One-week ($h = 5$)									
VIX^2	0.561	7.388***	−0.002	−0.112	−0.022	−1.558	0.126	7.718***	67.784
RV	0.505	8.419***	−0.036	−0.813	−0.123	−6.816***	0.322	17.580***	42.153
LTV	2.775	40.106***	0.090	3.406***	0.002	0.059	0.005	0.178	76.605
RSK	−0.056	−2.120**	0.032	1.738*	−0.138	−3.979***	0.063	1.904*	1.293
RK	0.524	4.373***	−0.520	−4.007***	0.475	3.347***	−0.646	−6.037***	3.831
Panel C: One-month ($h = 22$)									
VIX^2	0.518	6.697***	0.006	0.321	−0.002	−0.110	0.120	6.472***	60.318
RV	0.431	6.545***	−0.017	−0.389	−0.065	−2.344***	0.285	11.849***	34.597
LTV	9.648	21.124***	0.801	2.051**	0.473	2.163**	0.177	0.868	60.938
RSK	0.082	1.134	0.092	1.795*	−0.383	−2.801***	0.208	1.770*	1.980
RK	1.692	4.358***	−1.835	−4.509***	1.882	2.815***	−2.039	−5.014***	6.974

Note: The table reports the multivariate predictive regressions for the measures of risk of the S&P 500 index over one-day ($h = 1$), one-week ($h = 5$), and one-month ($h = 22$). We consider the squared VIX (VIX^2), realized variance (RV), left tail variation (LTV), realized skewness (RSK), and realized kurtosis (RK). For forecasting horizons larger than 1 day, we aggregate the risk measures from $t + 1$ to $t + h$. The t-statistics are computed using Newey-West robust standard errors with a lag length equal to h. F_t, in the first column, denotes the lagged value of the variable of interest. JV is estimated as $JV = (RV - BV)^+$, and BV is the bipower variation measure proposed by Barndorff-Nielsen and Shephard (2004). *, ** and *** indicate statistical significance at respectively the 10%, 5% and 1% level. The sample ranges from January, 2000 to December, 2020.

Table 12: Predicting variance risk premium (VRP) with control variables

	One-day ($h = 1$)			One-week ($h = 5$)			One-month ($h = 22$)		
RV	0.081	0.072	0.071	−0.021	−0.061	−0.069	−1.353	−1.449	−1.480
t-stat	2.289***	2.030**	1.980**	−0.090	−0.256	−0.284	−1.876*	−2.010**	−2.051**
JV	0.003	0.006	0.006	0.018	0.031	0.031	0.609	0.645	0.644
t-stat	0.090	0.172	0.171	0.187	0.326	0.325	1.720*	1.805*	1.812*
LTV	0.251	0.248	0.249	1.312	1.297	1.300	5.878	5.835	5.846
t-stat	8.620***	8.580***	8.632***	6.888***	6.840***	6.898***	5.893***	5.843***	5.901***
$\lambda^{\mathbb{P}}$	−0.045		−0.029	−0.200		−0.122	−0.264		−0.056
t-stat	−4.688***		−2.894***	−3.509***		−2.056**	−0.843		−0.165
$\lambda^{\mathbb{Q}}$		−0.070			−0.320			−0.629	
t-stat		−7.330***			−5.244***			−2.186**	
TRP			0.058			0.279			0.744
t-stat			6.061***			4.075***			2.504**
R^2_{is}	24.825	25.455	25.548	22.435	23.123	23.250	27.271	27.547	27.711

Note: The table reports the multivariate predictive regressions for the variance risk premium (VRP) (Bollerslev et al., 2009) over one-day ($h = 1$), one-week ($h = 5$), and one-month ($h = 22$). For forecasting horizons larger than 1 day, we aggregate the risk measures from $t+1$ to $t+h$. The t-statistics are computed using Newey-West robust standard errors with a lag length equal to h. The control variables are the realized variance (RV), jump variation (JV), and left tail variation (LTV) of Bollerslev et al. (2015). JV is estimated as $JV = (RV - BV)^+$, and BV is the bipower variation measure proposed by Barndorff-Nielsen and Shephard (2004). *, ** and *** indicate statistical significance at respectively the 10%, 5% and 1% level. The sample ranges from January, 2000 to December, 2020.