Graduate Theses, Dissertations, and Problem Reports

2019

Empirical Asset Pricing with Equity Tail Risk

Jingrui Li
West Virginia University, jgli@mix.wvu.edu

Follow this and additional works at: https://researchrepository.wvu.edu/etd

Part of the Finance and Financial Management Commons

Empirical Asset Pricing with Equity Tail Risk

Jingrui Li

A dissertation submitted to
the John Chambers College of Business and Economics
at West Virginia University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Business Administration in Finance

K. Victor Chow, Ph.D., Chair
Naomi Boyd, Ph.D.
Alexander Kurov, Ph.D.
Eric Olson, Ph.D.
Department of Finance

Morgantown, West Virginia
2019

ABSTRACT

Empirical Asset Pricing with Equity Tail Risk

Jingrui Li

This dissertation comprises three separate chapters on both risk-neutral and physical probability spaced equity tail risk for both the market index and in the cross-section of individual stocks.

The first chapter is titled "Does VIX Truly Measure Return Volatility?" This chapter studies the bias of the VIX index as a volatility measure. Particularly, VIX undervalues (overvalues) volatility when market return is expected to be negatively (positively) skewed. Alternatively, we develop a model-free generalized volatility index (GVIX). This chapter further derives the risk-neutral tail risk estimated from the VIX index.

The second chapter is titled "Decomposing the VIX: Implications for the Predictability of Stock Returns" This chapter studies the tail risk for the market index (S&P 500 index) in both risk-neutral and physical probability space and subsequently quantifies the market tail risk premium. Market tail risk premium also is a driving force of the VIX index, especially during a nervous market condition. The VIX decomposed market tail risk premium possesses significant prediction power for the equity market index (S&P500 index), Fama and French style portfolios, and industry portfolios with a prediction range that varies from one month to 12 months.

The third chapter is titled "The Predictive Power of Tail Risk Premia on Individual Stock Returns" This chapter studies both the risk-neutral and physical probability space tail risk for the cross-section of individual stocks and examines the characteristics of this premium in the cross-section of stock returns. The tail risk premium for individual stocks is statistically and economically priced in the cross-section of individual stock returns. Specifically, the existence of a premium for bearing negative tail risk is significantly associated with negative returns up to one month in the future. In contrast, the premium for bearing positive tail risk has no significant predictive power. This phenomenon cannot be explained by size, book-to-market ratio, market beta, idiosyncratic volatility, momentum, illiquidity, or lottery effect (maximum and minimum monthly returns).

Overall, the results from the three chapters indicate that equity tail risk is an important factor for the market index in both risk-neutral and physical probability spaces, and its premium carries strong return predictability for multiple market-level portfolio assets. Furthermore, equity tail risk and its premium carry significant return prediction power in the cross-section of individual stock returns. This phenomenon is robust to previously documented asset pricing factors.

DEDICATION

To my parents, Li Zhang and Hongyu Li, for their constant patience and support. I would also like to thank my grandparents, Shufang Ding and Qingshan Zhang, for raising and teaching me when I was a little girl. I would also like to thank my aunts, uncles, and cousins for their care and help while I've studied in the USA. I also thank Shangmin Wang for his companionship, understanding, and support throughout my studies in Morgantown, West Virginia.

ACKNOWLEDGMENTS

Table of Contents

Chapter 1   Does VIX Truly Measure Return Volatility?

Chapter Abstract

This article demonstrates theoretically that without imposing any structure on the underlying forcing process, the model-free CBOE volatility index (VIX) does not measure market expectation of volatility but that of a linear moment-combination. Particularly, VIX undervalues (overvalues) volatility when market return is expected to be negatively (positively) skewed. Alternatively, we develop a model-free generalized volatility index (GVIX). With no diffusion assumption, GVIX is formulated directly from the definition of log-return variance, and VIX is a special case of the GVIX. Empirically, VIX generally understates the true volatility, and the estimation errors considerably enlarge during volatile markets. The spread between GVIX and VIX follows a mean-reverting process.

This chapter is coauthored with Victor Chow (West Virginia University) and Wanjun Jiang (Peking University)

1.1 Introduction

One of the most frequently used devices to estimate expectation of market volatility is the Chicago Board of Option Exchanges (CBOE) Volatility Index (VIX). The index, designed to capture the market's aggregate expectation of future volatility over the next 30 days, was originally introduced in 1993 and based on the implied volatility of at-the-money option prices on the S&P100 index. Since the calculation of implied volatility relied heavily on assumptions of option pricing models, in 2003, CBOE made a significant change in the volatility estimating methodology to avoid the problem of modeling specification. The 2003 VIX formulation has intuitive appeal in that option premiums and the underlying asset (S&P 500 index) prices reflect the same information, and the ex-ante volatility of asset's returns can thus be directly extracted from option market data without employing any option pricing model. Simply, a volatility forecast does not require a specific model, only current option prices. Coined with Demeterfi, Derman, Kamal and Zou (1999), this concept of model-free implied volatility has been formalized by Britten-Jones and Neuberger (2000) and further extended by Jiang and Tian (2005, 2007).

Specifically, under diffusion assumptions of the return generating process, the arbitrage-free argument implies that the *fair value* of expected volatility is extractable from the market price of a portfolio composited by all possible out-of-the-money (OTM) call/put options of the underlying index with weight inversely proportional to square value of the strike price. As demonstrated by Jiang and Tian (2005), this model-free implied volatility provides a more efficient forecast for future realized volatility and is more efficient than the historical variance even if the underlying asset price process has jumps. Because of the superiority of the model-free approach to its predecessor, CBOE has recently applied its formulation to a wide range of indices including ones for Nasdaq 100, Dow Jones Industry Average, Russell 2000, interest rate, crude oil, gold/silver, energy, currency (Euro), as well as individual stocks such as Amazon, Apple, Goldman Sachs, Google and IBM.

This chapter argues that the expression of the CBOE volatility index (VIX) is in fact not model-free because its estimates are dependent on specification of a particular underlying stochastic process of returns. Without any diffusion assumption, VIX theoretically characterizes not the fair value of volatility but that of a linear combination of ex-ante distributional return-moments.[1] Particularly, the VIX index constitutes additionally one-third value of the ex-ante third return-moment (skewness). This highlights a potentially significant bias for VIX to serve as a volatility index because skewness continues to occupy a prominent role in financial markets. Intuitively, during a period when asset price-movement experiences downtrend momentum, demand for puts increases relatively to calls as investors seek protection, and put options become relatively more expensive than call options. The value of volatility, which is extractable from the market price of a long option portfolio, increases as put premiums increase. Simultaneously, the price of the third moment (skewness) is replicable by a long position of OTM calls and a short position of OTM puts, on the other hand, becomes largely negative.[2] Consequently, even when the

---

[1] Martin (2011) has shown that VIX^2 is a liner cumulant-combination of log returns but does not highlight the potential bias of VIX due to negative skewness.

[2] Bakshi, Kapadia, and Madan (2003, page 107, Equation [7]) explicitly show that the fair price of the third moment, denoted by W_T, is a linear sum of the out-of-the-money call option premium minus that of the put premium.

VIX (fear gauge) value increases, the index still understates the true value of volatility due to the impact of a negative third moment.

To avoid the impact of high moments on the volatility index as well as to retain the same economic merit as the VIX's conceptualization, we propose an alternative approach, named generalized VIX (GVIX). GVIX is generic because it is based on the direct definition of *log-return variance* with an option formulation extended from Bakshi, Kapadia, and Madan (2003). Neither modeling specification nor diffusion assumption is necessary to be imposed on the underlying asset's return generating process. Further, we prove that VIX is a special case of GVIX.[3]

We empirically investigate estimation errors of VIX and their relationship with ex-ante return-moments of the underlying asset (S&P 500 index) over a period from January 2005 to May 2014. In general, the empirical findings are consistent with our theoretical argument in that VIX statistically and significantly underestimates the true volatility. The VIX-bias dramatically increases as market volatility rises due to the impact of substantially negative third return-moment. In short, the larger the VIX, the higher the downward biases, and these errors could be as much as 559 index basis points.

Finally, we examine time-series properties of VIX, GVIX, and especially, the difference between GVIX and VIX (GV-spread). The individual dynamic process of both VIX and GVIX generally follows a random walk. However, the time variation of GV-spread over our sample period follows a mean-reversion with a break of the stationary trend dated by September 8th, 2008. The coexistence of the stationarity of GV-Spread as well as the non-stationarity of VIX and GVIX implies that the time-series of VIX and that of GVIX are co-integrated. The movement of the GV-spread quantifies the shift of distributional asymmetry of return expectations as well as the change of market sentiments.

Du and Kapadia (2012) is the closest to the current chapter. Du and Kapadia (2012) derive a jump and tail index, JTIX, and uses 22-day moving average of JTIX in their empirical estimation. This chapter differentiates itself from Du and Kapadia (2012) in important ways. First, we provide the discrete approximation procedure following VIX index methodology, this makes the application of our GVIX index more broadly available. Second, in our empirical estimation, we use the daily GVIX and GV-spread (the difference between VIX and the "true" volatility index) estimates instead of moving averages. Third, we examine the mean-reverting property of GV-spread and the possibility in terms of spread trading.

This chapter proceeds as follows. Section 1.2 demonstrates the related literature. Section 1.3 presents an important theoretical relationship between option premiums and an ex-ante moment-combination of returns to the underlying asset. Section 1.4 reviews the derivation of VIX and explicitly proves the equivalence between VIX and a moment-combination. An alternative model-free expression of the ex-ante volatility valuation is also proposed in Section 1.4. Section

---

[3] Martin (2011) developed a volatility index, named SVIX, based on the risk-neutral variance of *holding-period returns*. Since VIX is derived from the *log-return* process, SVIX is not equal to VIX even if the diffusion assumptions of VIX hold and/or the holding-period returns are log-normally distributed. Certainly, the *volatility of log-returns* (GVIX) is different from *that of holding-period returns* (SVIX).

1.5 empirically investigates our theoretical arguments, and Section 1.6 contains brief concluding remarks.

1.2   Related Literature

A pioneer work by Breeden and Litzenberger (1978) set the foundation of research on extracting risk-neutral distribution and model-free volatility from option market prices. The concept of model-fee implied volatility later arose from the research and development of variance-swaps, for example, see Dupire (1994), Neuberger (1994), Carr and Madan (2001), and Baskshi, Kapadia and Madan (2003).

Jiang and Tian (2005) extend Britten-Jones and Neuberger (2000) model-free implied volatility framework to asset prices processes with jumps. They state this model-free implied volatility provides a more efficient forecast than Black-Sholes implied volatility for future realized volatility and is more efficient than the historical variance allowing that the underlying asset price process has jumps.

Car and Wu (2006) describe the major differences between the old and the new volatility indexes of CBOE, they also derive the theoretical underpinnings for the two indexes. They conclude that the switch is due to the fact that the new VIX has a better known and more robust economic interpretation. Additionally, the variance swap underlying the new VIX formulation has a robust replicating portfolio whose option component is static.

Furthermore, in addition to Du and Kapadia (2012) discussed previously, several other papers are close to this chapter. Bakshi, Kapadia and Madan (2003) theoretically derive risk-neutral skewness and kurtosis for asset returns. This chapter adopts similar approach but applies to asset return volatility instead of centralized higher moments of returns and subsequently provides estimation details. Jiang and Tian (2007) examine the mispricing of VIX index and propose a simple smoothing method instead of VIX formula to estimate model-free return volatility. Comparatively, the GVIX volatility measure in this chapter is less restricted to interpolation choices and is based on the straight-forward variance concept. Martin (2011) develops a volatility index, named SVIX, based on the simple variance swaps. However, Martin (2011) does not provide the empirical estimation methodology nor examine the empirical trend of SVIX.

Internationally, Zheng, Jiang and Chen (2017) devise an improved model-free implied variation index (AVIX) based a generalized semi-martingale process with stochastic interest rates and applies it to China 50 ETF option markets. They conclude that AVIX is a better measure of investor sentiment compared to iVIX[4] in Chinese market.

1.3   Ex-ante Moment-Combination and Option Prices

This section demonstrates a generalized relation between option prices and a distributional moments' combination of log-returns without imposing any structure on the underlying focusing

---

[4] In June 2015, Shanghai Stock Exchange (SSE) launched a model-free volatility index for China 50 ETF option market, named iVIX. iVIX is estimated similar to CBOE VIX index based S&P 500 index options.

process. We begin conventionally with a simple holding-period return denoted by a capital R_T and a continuously compounded rate of return (or log return) denoted by a lowercase r_T, respectively. These are written as:

$$R_T = \frac{S_T - S_0}{S_0} \tag{1}$$

$$r_T = [\ln(S_T) - \ln(S_0)] = \ln\left(\frac{S_T}{S_0}\right) \tag{2}$$

where S_0 and S_T are the price of the asset at time 0 and T, accordingly. Consider the Taylor expansion with the remainder of $\ln(S_T)$ centered at S_0, we have

$$\ln(S_T) = \ln(S_0) + \frac{S_T - S_0}{S_0} + \int_{S_0}^{\infty} \frac{-1}{K^2}(S_T - K)^+ dK + \int_0^{S_0} \frac{-1}{K^2}(K - S_T)^+ dK, \tag{3}$$

Equivalently, using our return notations in (1) and (2), equation (3) can be rewritten as:

$$R_T - r_T = \left[\int_{S_0}^{\infty} \frac{1}{K^2}(S_T - K)^+ dK + \int_0^{S_0} \frac{1}{K^2}(K - S_T)^+ dK\right] \geq 0. \tag{4}$$

That is, the return of a long position in a contract that pays a holding-period return (R_T) and a short position in the contract that pays the logarithm of a total return at time T (r_T) can be replicated by payoffs of a long position in $\left(1/_{K^2}\right)$ call option struck at K for all strikes above the current asset price and these of a similar long position in $\left(1/_{K^2}\right)$ put option struck at K for all strikes below the current asset price, where all option contracts have T period of time to expiration. Notably, the non-negative option payoffs ensures that $R_T \geq r_T$.

Consider that under the no-arbitrage condition, a currently fair price of an asset can be extracted from the put-call parity of European options so that $S_0 = C_T(K_A) - P_T(K_A) + K_A e^{-rT}$, where $C_T(K_A)$ and $P_T(K_A)$ are the current premiums of call and put option contracts with an at-the-money (ATM) strike K_A and expiration T, respectively. r is the annualized risk-free rate corresponding to expiration date T. The forward asset price and expected return can then be determined in a risk-neutral framework in that $E(S_T) = e^{rT}[C_T(K_A) - P_T(K_A)] + K_A = S_0 e^{rT}$, and importantly $E(R_T) = \frac{E(S_T) - S_0}{S_0} = e^{rT} - 1$, where $E(\cdot)$ represents the expectation operator. This implies that the fair value of the expected return-difference in (4) can also be drawn from the risk neutral forward value of the replicated option portfolio:

$$E(R_T) - E(r_T) = e^{rT}\left\{\int_{S_0}^{\infty} \frac{1}{K^2} C_T(K)dK + \int_0^{S_0} \frac{1}{K^2} P_T(K)dK\right\}, \tag{5}$$

For connecting equation (5) to the expression of moment-combination, we apply Taylor series to an exponential function of r_T as follows:

$$(1 + R_T) = \frac{S_T}{S_0} = exp\left[\ln\left(\frac{S_T}{S_0}\right)\right] = 1 + \sum_{\kappa=1}^{N} \frac{1}{\kappa!}\left[\ln\left(\frac{S_T}{S_0}\right)\right]^\kappa + o\left[\ln\left(\frac{S_T}{S_0}\right)\right]^N \tag{6}$$

Then, taking the expected value of both sides of equation (6) with some re-arrangement, the expected value of return difference between R_T and r_T is equal to a linear combination of distributional moments:

$$E(R_T) - E(r_T) = \frac{1}{2}E(r_T^2) + \frac{1}{6}E(r_T^3) + \frac{1}{24}E(r_T^4) + o[E(r_T^4)] \tag{7}$$

Interestingly, Equation (7) shows that the expected holding-period return is the moment generating function of the log-returns. Consequently, when we incorporate (5) with (7), the forward value of an OTM option portfolio represents an ex-ante moment-combination of log-returns. We demonstrate formally this important result in the following theorem:

Theorem 1. *Let $C_T(K)$ and $P_T(K)$ be the European call and put options with an exercise price of K and expiration date at T, respectively. Without imposing any structure on the underlying focusing process, the fair value of ex-ante moment-combination of log-returns can be extracted from the market price of an OTM option portfolio with weight inversely proportional to square value of the strike price in that*

$$\left[\frac{1}{2}E(r_T^2) + \frac{1}{6}E(r_T^3) + \frac{1}{24}(r_T^4) + o[E(r_T^4)]\right] = e^{rT}\left[\int_{S_0}^{\infty} \frac{1}{K^2}C_T(K)dK + \int_0^{S_0} \frac{1}{K^2}P_T(K)dK\right] \tag{8}$$

Theorem 1 highlights an essential connection between option premiums and market expectations about future return-moments. Specifically, the market price of a $\left(1/_{K^2}\right)$ weighted OTM put/call option portfolio extracts the fair value of a linear combination of ex-ante moments (e.g., variance, skewness, kurtosis, etc.). Next, we note that the VIX index was derived based on the concept of fair value of *future* variance and used *forward* price instead of current price of the underlying asset as a reference point to determine future OTM option payoffs of calls and puts. In addition, since it is almost certain, in practice, no option contract has a strike price exactly equal to the forward price, one could select the first strike price below the forward price as the reference price. In the next corollary, we adjust the fair value of the option portfolio by taking the forward price into consideration.

Corollary 1. *Let S_0 be the current fair price of the underlying asset determined from ATM Put-Call parity, and the forward price (F_0) at time T is then equal to $S_0 e^{rT}$, where r is the annualized risk-free rate of interest. Further, let K_0 be the first strike price below the forward price $(K_0 \le F_0)$. The forward value of the options portfolio in Theorem 1. has the following equality:*

$$e^{rT}\left[\int_{S_0}^{\infty} \frac{1}{K^2}C_T(K)dK + \int_0^{S_0} \frac{1}{K^2}P_T(K)dK\right]$$
$$= \left[(e^{rT} - 1) - \left(\frac{F_0}{K_0} - 1\right) - \ln\left(\frac{K_0}{S_0}\right)\right] + e^{rT}\left[\int_{K_0}^{\infty} \frac{1}{K^2}C_T(K)dK + \int_0^{K_0} \frac{1}{K^2}P_T(K)dK\right] \tag{9}$$

The proof of Corollary 1 is straightforward. By employing Taylor expansion with the remainder of $\ln(S_T)$ about the point K_0, we have

$$ln(S_T) = ln(K_0) + \frac{S_T - K_0}{K_0} + \int_{K_0}^{\infty} \frac{-1}{K^2} (S_T - K)^+ dK + \int_0^{K_0} \frac{-1}{K^2} (K - S_T)^+ dK \tag{10}$$

Substituting (3) into (10), take the risk-neutral expectation, and realize that $E(R_T) = e^{rT} - 1$, we have equation (9). The first term of the right-hand side of equation (9) represents the adjustment between F_0, S_0 and K_0.

In summary, from mathematical applications to the relationship between a holding period return (R_T) and log-return (r_T), we illustrate that without any diffusion assumption and model-specification, a portfolio of OTM options weighted inversely proportional to the square value of the strike price extracts the fair value of a linear combination of ex-ante return-moments. Since the CBOE volatility index is calculated from the price of the identical option portfolio as (9), we explicitly prove, in the next section, that the VIX formulation is actually a moment-combination not a volatility measure in general. A generalized model-free volatility index is also proposed in the next section.

1.4   The VIX and GVIX Indexes

This section reviews briefly the derivation of the CBOE volatility index, illustrates the validity problem of the VIX to serve as a volatility index, and develops a generic method to estimate ex-ante volatility. We first note that although the revamped VIX index on September 2003 no longer relies on any option pricing model, as shown by Demeterfi, Derman, Kamal and Zou (1999), the derivation of the index still assumes that the stochastic process of returns to assets follows a specification of Geometric Brownian Motion (GBM) and Ito's Lemma (IL):

$$\begin{cases} \dfrac{dS_t}{S_t} = \mu dt + \sigma dZ_t, \text{and} & \qquad\qquad (11) \\[2mm] d[\ln(S_t)] = \left(\mu - \dfrac{1}{2}\sigma^2\right) dt + \sigma dZ_t & \qquad\qquad (12) \end{cases}$$

where Z_t is a Wiener process or Brownian motion, μ is a fixed drift, and σ is a constant volatility, accordingly.[5] GBM and IL assume asset returns to be continuous and symmetrically distributed. Thus, higher orders of moments than the variance neither exist nor have impact on the return generating process. Since the log-return over a period of T and its variance can be expressed by $r_T = \int_0^T d[\ln(S_t)]$ and $\int_0^T \sigma^2 \, dt$, respectively, the instantaneous volatility (variance) can then be calculated directly from simultaneous equations (11) and (12) as follows:

---

[5] It is assumed that $dZ_t^2 = dt$, $dt \cdot dZ_t = 0$, and $dt \cdot dt = 0$.

$$\mathbf{v} = \frac{1}{T}\int_0^T \sigma^2 \, dt = \frac{2}{T}\left[\int_0^T \frac{dS_t}{S_t} - r_T\right] \tag{13}$$

As also illustrated by Demeterfi, Derman, Kamal and Zou (1999), equation (13) indicates that the ex-ante volatility is replicated by a portfolio of two positions: (1) a continuous rebalanced position of instantaneously long $\frac{1}{S_t}$ shares of the underlying asset worth one dollar, and (2) a short position in a contract that pays the logarithm of the total return at time T. Under the risk-neutral framework, $E\left(\int_0^T \frac{dS_t}{S_t}\right) = rT$, and $E(R_T) = e^{rT} - 1$, according to (5) and (13), the expected volatility can be expressed as:

$$E(\mathbf{v}) = \sigma^2 = \frac{2}{T}\left\{e^{rT}\left[\int_{S_0}^{\infty} \frac{1}{K^2}C_T(K)dK + \int_0^{S_0}\frac{1}{K^2}P_T(K)dK\right] - [(e^{rT}-1) - rT]\right\}, \tag{14}$$

Consequently, the expected value of return-volatility, under the assumptions (11) and (12), can be extracted from the market price of a portfolio of out-of-the money options weighted inversely proportional to the square value of the strike prices. The third term of equation (14), expressed by $(e^{rT}-1) - rT$, captures the difference between continuously compounding return and single compounding return over a period of T. This difference approaches zero, when T is small.

By applying the results of Corollary 1 to equation (14), we have the VIX formulation, which is identical to that in Demeterfi, Derman, Kamal and Zou (1999), as follows:

$$VIX^2 = E(\mathbf{v}) = \frac{2}{T}\left\{rT - \left(\frac{F_0}{K_0} - 1\right) - ln\left(\frac{K_0}{S_0}\right) + e^{rT}\left[\int_{K_0}^{\infty} \frac{1}{K^2}C_T(K)dK + \int_0^{K_0}\frac{1}{K^2}P_T(K)dK\right]\right\} \tag{15}$$

Therefore, independent of any option pricing model, the implied volatility is calculated explicitly from a set of out-of-the money option prices, current asset price, and risk-free rate of interest. Furthermore, applying Taylor's expansion of log function and ignoring terms higher than the second order, the terms preceding the integrals in (15) can be restated as

$$\left[rT - \left(\frac{F_0}{K_0} - 1\right) - ln\left(\frac{K_0}{S_0}\right)\right] = \left[ln\left(\frac{F_0}{K_0}\right) - \left(\frac{F_0}{K_0} - 1\right)\right] \approx -\frac{1}{2}\left(\frac{F_0}{K_0} - 1\right)^2 \tag{16}$$

Consequently, identical to that shown in CBOE's VIX whitepaper, the volatility index can be calculated under the discrete framework as follows:

$$\widehat{VIX}^2 = \frac{2e^{rT}}{T}\sum_i \frac{1}{K_i^2}Q(K_i)\Delta K_i - \frac{1}{T}\left(\frac{F_0}{K_0} - 1\right)^2 \tag{17}$$

where $Q(K_i)$ is the midpoint of the bid-ask spread for the option (a call if $K_i > K_0$, a put if $K_i < K_0$, both call and put if $K_i = K_0$), and $\Delta K_i = \frac{1}{2}(K_{i+1} - K_{i-1})$.[6] The forward price is calculated from

---

[6] For the minimum (maximum) strike, ΔK_i is simply the distance to the next strike above (below).

at-the-money options using put-call parity so that $F_0 = K_A + e^{rT}[C_T(K_A) - P_T(K_A)]$. Furthermore, CBOE calculates the VIX from an interpolation of two volatility indexes with respect to two expiration dates: the near-term (T_1) and next-term (T_2), respectively. Finally, by taking a weighted average of these two VIX measures, one obtains an annualized index (denoted as $\widehat{VIX}_a$) as follows:

$$\widehat{VIX}_a = \sqrt{\left\{\left[\frac{N(T_2) - N(30)}{N(T_2) - N(T_1)}\right](T_1 \cdot \widehat{VIX}_1^2) + \left[\frac{N(30) - N(T_1)}{N(T_2) - N(T_1)}\right](T_2 \cdot \widehat{VIX}_2^2)\right\}\left(\frac{N(365)}{N(30)}\right)} \qquad (18)$$

where $N(T_1)$ and $N(T_2)$ are the number of time-intervals (e.g. minutes, days) of the near term and next term options. $N(30)$ and $N(365)$ are the number of time-intervals in 30 days and in a 365-day year, correspondingly.

The CBOE volatility index is derived based on the model-free volatility expectation shown in (14). The key assumption required to derive VIX is that the stochastic process for the underlying asset price is continuous and follows the GBM and IL. When there are relatively small jumps in the stock price process, Jiang and Tian (2005) and Carr and Wu (2006, 2008) show that this model-free approach is an excellent approximation of the risk-neutral, expected quadratic variation of the logarithm of the asset price. However, if there is an appreciable risk of a large jump, then the approximation error can be very significant. In the next theorem, we show that without any diffusion assumption of the underlying asset return generating process, the model-free VIX expression is in fact a moment-combination of log-returns to its underlying portfolio.

Theorem 2. *let $V_T = E(r_T^2)$, $W_T = E(r_T^3)$, and $X_T = E(r_T^4)$ be the second, third and fourth ex-ante return-moments to an asset. Without any modeling specification about the stochastic process of returns, the formulation of COBE volatility index* (VIX) *is equal to a linear combination of return-moments such that:*

$$VIX = \frac{1}{\sqrt{T}}\sqrt{\left[V_T + \frac{W_T}{3} + \frac{X_T}{12} + o(X_T)\right] - 2[(e^{rT} - 1) - rT]},$$

and
$$\lim_{T \to 0}[(e^{rT} - 1) - rT] = 0$$

$$(19)$$

Equation (19) can be directly derived from Theorem 1, Corollary 1 as well as equations (14) and (15). Theorem 2 identifies the invalidity of VIX as a volatility index unless high moments of the return distribution neither exist nor have impact on the return generating process. Therefore, any approximation error of VIX to the true volatility could intuitively be due to the impact of high return-moments.

Bakshi, Kapadia, and Madan (2003) show that without imposing any structure on the underlying forcing process, the distributional moments, V_T, W_T, and X_T can be extracted from prices of OTM option portfolios. This indicates that the expected value of ex-ante volatility can also be estimated without any specification of the underlying stochastic process. We extend the Bakshi, Kapadia, and Madan (2003) approach and propose a direct estimation of ex-ante return-volatility in the following theorem.

Theorem 3. *Let $C_T(K)$ and $P_T(K)$ be the European call and put options with an exercise price of K and expiration date T, respectively. S_0 is the current asset value under no-arbitrage condition that determined from ATM Put-Call parity, $S_0 = C_T(K_A) - P_T(K_A) + K_A e^{-rT}$, and the forward prices is $F_0 = E(S_T) = S_0 e^{rT}$. Further, let K_0 be the first strike price y below the forward price (F_0). Assume that the finite first and second moments of log-return distribution (denoted by μ_T and V_T, respectively) exist. Without any specification about the form of distribution and/or stochastic process of returns, a generalized volatility index (GVIX), can be directly formulated by the definition of log-return standard deviation as follows:*

$$GVIX = \frac{1}{\sqrt{T}}\sqrt{V_T - (\mu_T)^2} \tag{20}$$

where

$$\mu_T = ln\left(\frac{K_0}{S_0}\right) + \left(\frac{F_0}{K_0} - 1\right) - e^{rT}\left[\int_{K_0}^{\infty}\frac{1}{K^2}C_T(K)dK + \int_0^{K_0}\frac{1}{K^2}P_T(K)dK\right], \tag{21}$$

and

$$V_T = ln^2\left(\frac{K_0}{S_0}\right) + 2ln\left(\frac{K_0}{S_0}\right)\left(\frac{F_0}{K_0} - 1\right)$$
$$+ 2e^{rT}\left[\int_{K_0}^{\infty}\frac{\left[1 - ln\left(\frac{K}{S_0}\right)\right]}{K^2}C_T(K)dK + \int_0^{K_0}\frac{\left[1 + ln\left(\frac{S_0}{K}\right)\right]}{K^2}P_T(K)dK\right] \tag{22}$$

The ex-ante mean of log-returns, $\mu_T = E(r_T)$, can be calculated directly from (5) and (9), and the second-moment, $V_T = E(r_T^2)$, can be derived based on Taylor expansion. Note that Taylor's expansion with remainder for any continuously twice differentiable function can be expressed as: $f(S_T) = f(a) + f'(a)(S_T - a) + \int_a^{\infty} f''(K)(S_T - K)^+ dK + \int_0^a f''(K)(K - S_T)^+ dK$, where $f(a)$ and $f''(K)$ are the first and second derivatives of the payoff with respect to S evaluated at a and K, respectively. To derive V_T, define $f(S_T) = ln^2\left(\frac{S_T}{S_0}\right)$, and let $a = K_0$. Then,

$$ln^2\left(\frac{S_T}{S_0}\right) = ln^2\left(\frac{K_0}{S_0}\right) + 2ln\left(\frac{K_0}{S_0}\right)\left(\frac{S_T - K_0}{K_0}\right) + 2\int_{K_0}^{\infty}\frac{\left[1 - ln\left(\frac{K}{S_0}\right)\right]}{K^2}(S_T - K)^+ dK + 2\int_0^{K_0}\frac{\left[1 + ln\left(\frac{S_0}{K}\right)\right]}{K^2}(K - S_T)^+ dK \tag{23}$$

Further, by taking the risk-neutral expectation for both sides of equation (23), we have equation (22). Our derivation of (22) provides similar results as shown in Bakshi, Kapadia, and Madan (2003). Theorem 3 formulates a generalized mechanism to extract fair value of ex-ante volatility from OTM option portfolios. In the next theorem, we show that the CBOE's VIX is just a special case of the proposed volatility index (GVIX).

Theorem 4. *CBOE's VIX is a special case of GVIX in that they are equal only if the stochastic process of asset's return follows Geometric Brownian Motion and Ito's Lemma as shown in (11) and (12).*

The proof of the above theorem is straight-forward. Based on the assumption of Geometric Brownian Motion and Ito's Lemma, $E\left[ln^2\left(\frac{S_T}{S_0}\right)\right] = \left[\left(\mu - \frac{1}{2}\sigma^2\right)T\right]^2 + \sigma^2 T$, and $E\left[ln\left(\frac{S_T}{S_0}\right)\right] = \left(\mu - \frac{1}{2}\sigma^2\right)T$.

From (13), (14), and (19), $\text{GVIX}^2 = \frac{1}{T}\left\{E\left[\ln^2\left(\frac{S_T}{S_0}\right)\right] - \left\{E\left[\ln\left(\frac{S_T}{S_0}\right)\right]\right\}^2\right\} = \sigma^2 = \frac{1}{T}\int_0^T \sigma^2\,dt = E(V) = \text{VIX}^2$. Thus, if the diffusion assumptions of (11) and (12) hold, then $\text{GVIX} \equiv \text{VIX}$. (Q.E.D.)

Finally, according to equations (19) and (20), it is clear that the deviation of VIX^2 from the ex-ante variance of log-returns (GVIX^2) could be primarily driven by the high moments of returns in that

$$VIX^2 - GVIX^2 \approx \frac{1}{T}\left(\mu_T^2 + \frac{W_T}{3} + \frac{X_T}{12}\right). \tag{24}$$

The equality of (24) indicates that the spread of GVIX and VIX (GV-Spread) is determined by the ex-ante mean return of the underlying asset as well as the ex-ante high moments of the return distribution.

In practice, the calculation of our generalized volatility index is as simple as that of VIX (equation [17]). No additional information is necessary, and one is able to compute the GVIX based on three aspects: option prices (Q_K), arbitrage-free forward price of the underlying asset ($F_0 = S_0 e^{rT}$) and risk-free rate of interest (r) such that:

$$\hat{\mu}_T = \left[\ln\left(\frac{K_0}{S_0}\right) + \left(\frac{F_0}{K_0} - 1\right)\right] - e^{rT}\sum_i \frac{1}{K_i^2}Q(K_i)\Delta K_i, \tag{25}$$

$$\hat{V}_T = \left[\ln^2\left(\frac{K_0}{S_0}\right) + 2\ln\left(\frac{K_0}{S_0}\right)\left(\frac{F_0}{K_0} - 1\right)\right] + 2e^{rT}\sum_i \frac{1}{K_i^2}\left[1 + \ln\left(\frac{S_0}{K_i}\right)\right]Q(K_i)\Delta K_i, \tag{26}$$

and

$$\widehat{GVIX} = \frac{1}{\sqrt{T}}\sqrt{\hat{V}_T - (\hat{\mu}_T)^2} \tag{27}$$

Furthermore, to be consistent with CBOE's VIX estimation, we adjust the GVIX index on a 30-day basis. Following the same interpolation procedures as VIX estimation (see equation [18]), we first identify option contracts of the near-term and next-term, denoted by T_1 and T_2, accordingly. Then, one calculates two GVIX indexes according to T_1 and T_2, respectively. Finally, by taking a weighted average, one could calculate the annualized GVIX index based on 30-day market returns as follows:

$$\widehat{GVIX}_a = \sqrt{\left\{\left[\frac{N(T_2) - N(30)}{N(T_2) - N(T_1)}\right](T_1 \cdot \widehat{GVIX}_1^2) + \left[\frac{N(30) - N(T_1)}{N(T_2) - N(T_1)}\right](T_2 \cdot \widehat{GVIX}_2^2)\right\}\left(\frac{N(365)}{N(30)}\right)} \tag{28}$$

Having mathematically explained the bias of the VIX formulation and proposed an alternative volatility measure in theory, we now pursue our empirical investigation in the next section.

1.5   Empirical Investigation

We empirically examine our theoretical arguments using the S&P500 index (SPX) options data during the 2362-daily period between January 2005 and May 2014.[7] To ensure the empirical calculation is consistent with CBOE's estimation, our daily VIX estimates, according to equation (18), are recalculated following the estimation procedures specified in the CBOE's VIX whitepaper.

[Insert Table 1 here]

Table 1 compares our VIX computation (denoted by $\widehat{VIX}_a$) and daily VIX closing price (denoted as $\widehat{VIX}_{CBOE}$). Panel 1 numerically reports the difference between $\widehat{VIX}_{CBOE}$ and $\widehat{VIX}_a$ according to the empirical quantile of $\widehat{VIX}_{CBOE}$.   The difference ranges from -0.05 to 0.16. We note that this difference could be due to the time frequency of data observation in which we use daily prices versus the minute basis adopted by CBOE. Panel 2 shows the bootstrapping test for the mean difference of $\widehat{VIX}_{CBOE}$ and $\widehat{VIX}_a$. By selecting randomly 2000 data from 2362 daily observations, we calculate the mean and the associated Z-statistic. This procedure is then repeated 10,000 times for computing the average of means and that of Z-statistics. The mean difference is statistically insignificance indicating the estimates of $\widehat{VIX}_{CBOE}$ and $\widehat{VIX}_a$ are consistent with each other. Furthermore, to examine the equality of probability distributions between $\widehat{VIX}_{CBOE}$ and $\widehat{VIX}_a$, we employ a two-sample Kolmogorov-Smirnov (*K-S*) Test. Panel 3 reports the *K-S* test statistic ($D = 0.0085$) with an associated *p-value*. It appears that the test fails to reject the equality of the two distributions. From the analytical results in Table 1, we conclude that the diversity of our VIX estimation from CBOE's VIX is almost undistinguishable.

After the above robustness check for our estimation procedures, we now inspect the bias of VIX in relation to distributional moments. For formulating the value of ex-ante third and fourth return-moments (W_T and X_T), we follow Bakshi, Kapadia, and Madan (2003) with some modifications.

$$W_T = \ln^3\left(\frac{K_0}{S_0}\right) + 3\ln^2\left(\frac{K_0}{S_0}\right)\left(\frac{F_0}{K_0} - 1\right)$$
$$+ 3e^{rT}\left[\int_{K_0}^{\infty} \frac{\left[2\ln\left(\frac{K}{S_0}\right) - \ln^2\left(\frac{K}{S_0}\right)\right]}{K^2} C_T(K)dK - \int_0^{K_0} \frac{\left[2\ln\left(\frac{S_0}{K}\right) + \ln^2\left(\frac{S_0}{K}\right)\right]}{K^2} P_T(K)dK\right] \quad (29)$$

and

$$X_T = \ln^4\left(\frac{K_0}{S_0}\right) + 4\ln^3\left(\frac{K_0}{S_0}\right)\left(\frac{F_0}{K_0} - 1\right)$$
$$+ 4e^{rT}\left[\int_{K_0}^{\infty} \frac{\left[3\ln^2\left(\frac{K}{S_0}\right) - \ln^3\left(\frac{K}{S_0}\right)\right]}{K^2} C_T(K)dK + \int_0^{K_0} \frac{\left[3\ln^2\left(\frac{S_0}{K}\right) + \ln^3\left(\frac{S_0}{K}\right)\right]}{K^2} P_T(K)dK\right] \quad (30)$$

Similar to those in (17) and (18), estimates of W_T and X_T as well as their annualization can be calculated from option data:

---

[7] We obtained options data from Ivolatility.com

$$\widehat{W}_T = \left[\ln^3\left(\frac{K_0}{S_0}\right) + 3\ln^2\left(\frac{K_0}{S_0}\right)\left(\frac{F_0}{K_0} - 1\right)\right] + 3e^{rT}\sum_i\frac{1}{K_i^2}\left[2\ln\left(\frac{K_i}{S_0}\right) - \ln^2\left(\frac{K_i}{S_0}\right)\right]Q_{K_i}\Delta_{K_i}, \qquad (31.1)$$

$$\widehat{W}_a = \left\{\left[\frac{N(T_2) - N(30)}{N(T_2) - N(T_1)}\right](T_1 \cdot \widehat{W}_1) + \left[\frac{N(30) - N(T_1)}{N(T_2) - N(T_1)}\right](T_2 \cdot \widehat{W}_2)\right\}\left(\frac{N(365)}{N(30)}\right), \qquad (31.2)$$

$$\hat{X}_T = \left[\ln^4\left(\frac{K_0}{S_0}\right) + 4\ln^3\left(\frac{K_0}{S_0}\right)\left(\frac{F_0}{K_0} - 1\right)\right] + 4e^{rT}\sum_i\frac{1}{K_i^2}\left[3\,ln^2\left(\frac{K_i}{S_0}\right) - ln^3\left(\frac{K_i}{S_0}\right)\right]Q_{K_i}\Delta_{K_i}, \qquad (32.1)$$

and

$$\hat{X}_a = \left\{\left[\frac{N(T_2) - N(30)}{N(T_2) - N(T_1)}\right](T_1 \cdot \hat{X}_1) + \left[\frac{N(30) - N(T_1)}{N(T_2) - N(T_1)}\right](T_2 \cdot \hat{X}_2)\right\}\left(\frac{N(365)}{N(30)}\right), \qquad (32.2)$$

[Insert Plot 1 here]

Plot 1 illustrates VIX's estimation bias of volatility and how this bias is related to distributional moments. Expressed by equations (20) and (27), since it directly estimates *ex-ante standard deviation* of log-returns, GVIX could serve as a measure of the *true return volatility*. Therefore, we adopt the difference between GVIX and VIX as the deviation of VIX from the true volatility and refer to this deviation as VIX-bias. Plot 1A depicts the daily VIX-bias over our sample period. The daily VIX-bias, in a range from -0.03 to 5.58, is generally positive indicating that VIX understated the true volatility during January 2005 to May 2014. This volatility undervaluation was extremely large during the financial crisis of 2009. Plot 1B plots the divergence of VIX from the second moment (V_a). A similarly reversed pattern as Plot 1A appears in Plot 1B. This provides further evidence of the downward bias of VIX. Notably, the movement of daily market expectation of third-moment estimates ($\widehat{W}_a$), pictured in Plot 1C, moves almost perfectly and adversely with the VIX-bias. The third-moment estimates were generally non-positive during our sample period and had a value ranging from -3585 to 19.51. Since the composition of VIX includes one-third of the third-moment according to Theorem 2, the negative value of $\widehat{W}_a$ appears to be the main cause for VIX's underestimation of the true volatility. Plot 1D displays the impact of the fourth-moment (measured by one-twelfth of $\hat{X}_a$) over time. Again, the movement of VIX-bias seems to be highly correlated with that of $\hat{X}_a$. In addition, the value of $\hat{X}_a$ also dramatically increased when the market experienced high volatility.

[Insert Table 2 here]

To test VIX-bias statistically, we again employ a bootstrapping approach. In addition to the overall sample, three subsample periods are defined by 2005-2007, 2008-2011, and 2012-2014, respectively. For each iteration, we select randomly 700 (out of 754), 900 (out of 1008), 500 (out of 600) and 2000 (out of 2362) daily observations from the three sub-samples (2205-2007, 2008-2011 and 2012-2014) as well as the overall sample, respectively. The sample mean of the VIX-bias and its associated Z-statistic are then calculated. We repeat this procedure 10,000 times and compute the average of means and that of Z-statistics. Table 2 reports testing results of two measures: the mean difference between GVIX and VIX ($\widehat{GVIX}_a - \widehat{VIX}_a$) as well as that between VIX and the second return-moment ($\widehat{VIX}_a - \sqrt{\hat{V}_a}$). It appears that VIX statistically understates the return-volatility on average. The sample mean of ($\widehat{GVIX}_a - \widehat{VIX}_a$) ranges from 22 to 83 index basis-points

13

(BP), and that of $(\widehat{VIX}_a - \sqrt{\widehat{V}_a})$ has a value from -86 to -26 index BP. The Z-value of all estimates is considerably large indicating that the mean deviation of VIX from the return-volatility is statistically and strongly significant. We note that the VIX-bias during the period of highly volatile markets (2008-2011) was much higher than that during other time periods. In addition, since from equation (24), the VIX-base is theoretically driven by high moments, we also reports, in Table 2, the average value of $\frac{1}{3}\widehat{W}_a$ and that of $\frac{1}{12}\widehat{X}_a$ as well as their test statistics. It shows that all test statistics are highly significant indicating all estimates of the high moments are different from zero. The third-moment was largely negative and reached an extreme value of -82.76 during the period of (2008-2011). Although the fourth-moment provides some contribution to the VIX-bias, the impact was much smaller than that of the third-moment. In summary, Table 2 shows that the negative third-moment could be the main reason for the downward bias of VIX in estimating ex-ante volatility.

[Insert Table 3 here]

Corresponding to Plot 1, Table 3 reports correlations of VIX-bias to the return-moments. The VIX-bias appears to be highly correlated with all three return-moments: $\widehat{V}_a$, $\widehat{W}_a$, and $\widehat{X}_a$. The negative correlation between VIX-bias and $\widehat{W}_a$ is particularly large (about 97%) and is consistent across all sample periods. The second return-moment ($\widehat{V}_a$) is also negatively and almost perfectively correlated with the third return-moment ($\widehat{W}_a$). This indicates that the non-positive skewness of returns becomes more negative as market volatility rises. Note that the VIX index itself is highly correlated with the market volatility as well. Therefore, one would expect that the downward VIX-bias could increase as the value of VIX rises. We illustrate this phenomenon in Plot 2.

[Insert Plot 2 here]

To show the descending bias of VIX, we plot the VIX deviations from the true volatility (GVIX) on the y-axis and the corresponding VIX values (sorted from the minimum value of 9.87 to the maximum value of 80.70) on the x-axis. Plot 2 shows that the *downward bias* of VIX is quite significant, and the VIX undervaluation accelerates as VIX values increase. Importantly, the magnitude of this acceleration dramatically enlarges as the VIX values reach 30 and above. This accentuates a major drawback of the VIX index that frequently serves as an "investor *fear gauge*". It has been widely viewed that VIX values greater than 30 are generally associated with a large amount of volatility as a result of investor fear, while values below 20 generally correspond to less stressful times in the markets. Corresponding to the appearance of Plot 2, we report concisely the distribution of the VIX-bias according to three levels of VIX values: VIX < 30, 30 ≤ VIX < 50, and VIX > 50, respectively. The median bias increases from 28 to 126 index BP when VIX index value rises from below 30 to above 30. The median bias for VIX values above fifty even reaches 338 index BP. Furthermore, to examine the change of the bias sensitivity to the level of VIX, we let $I_1 = 1$, if $\widehat{VIX}_a < 30$, and 0 otherwise as well as $I_2 = 1$, if $\widehat{VIX}_a \geq 30$, and 0 otherwise. To this end, we estimate an ordinary least squares regression:

$$\left(\widehat{VIX}_a - \widehat{GVIX}_a\right) = \alpha + \beta_1(\widehat{VIX}_a \times I_1) + \beta_2(\widehat{VIX}_a \times I_2) + \epsilon \tag{33}$$

We report the regression outcomes in a sub-table under Plot 2. It appears that the intercept ($\alpha = 0.38$), and the two beta coefficients ($\beta_1 = -0.039$; $\beta_2 = -0.053$) are all statistically different from zero. The R^2 of the regression is 86%. Again, further decrease of the beta after VIX index passes 30 suggests that the elasticity of VIX-bias increases as the market become more volatile. In summary, the COBE volatility index (VIX) significantly underestimates the true market volatility, and the degree of bias considerably increases as the VIX index value rises above 30. Consequently, the CBOE volatility index could fail to serve as an accurate indicator for investor fear, and financial products using VIX as the underlying index could be undervalued as well.

The last part of empirical investigation focuses on the time-series property of the VIX, GVIX and the spread between GIX and VIX (GV-spread). We employ three different unit-root tests including Dickey and Fuller (1979), Phillips and Perron (1988) as well as Zivot and Andrew (1992). The augmented Dickey-Fuller (ADF) investigates a unit root present in an autoregressive progress. The Phillips–Perron (PP) test makes a non-parametric correction for any serial correlation and heteroskedasticity in the regressive errors by modifying the ADF test statistics. Zivot-Andrews (ZA) test examines unit-root of time series with endogenous structural break. Specifically, ZA test is a sequential test which utilizes the full sample and uses a different dummy variable for each possible break date. Identifying potential break of stationary trend is important in that our sample period involves market shocks. As Table 4 presents, VIX and GVIX series are non-stationary in that we fail to reject the null hypothesis of unit-root, according to ADF and ZA tests. However, the unit-root hypothesis of GV-spread is strongly rejected indicating that the time-series of VIX and GVIX are co-integrated. According to the PP test, the null hypothesis of unit-root is rejected for all three series.

[Insert Table 4 here]

Plot 3 depicts the time plots of ZA daily *t*-statistics over the entire sample period from January 2005 to May 2014. The null hypothesis of ZA test is a unit-root process without any exogenous structural breaks, and the alternative hypothesis is a trend-stationary process with possible structural change occurring at an unknown point in time. Any negative daily *t*-statistic that has a value less than the critical value of -4.80 is statistically significant and the null hypothesis of unit-root is rejected. Since all test-statistics of VIX and GVIX are negative but greater than the critical value, we fail to reject the unit-root hypothesis. Nevertheless, for the GV-spread, most of the daily t-statistics are located below the critical value line indicating that the unit-root hypothesis is strongly rejected in favor of stationarity. In Plot 3, the break date of the trend is indicated by the minimum test statistic and marked by a vertical "dot-line". Therefore, the endogenous structural breakpoint date for the GV-spread is on September 8th, 2008. This implies that the GV-spread follows a stationary trend, and the beak of the GV-spread trend could predict the shock or crisis.[8] Intuitively, according to equation (24), the GV-spread is characterized by the log-return moments. Particularly, the spread enlarges as the expectation of future returns skews negatively. On the other hand, the spread reduces, if the returns are expected to be symmetrically distributed. In his study of "the crash of '87", Bates (1991) shows the market sentiment can be reflected by the distributional asymmetry of ex-ante returns, and the asymmetry moves back to the symmetry as markets calm down. Therefore, the mean-reverting process of GV-spread, quantified by the

---

[8] The mean-reversion of the GV-Spread suggests that analogous to VIX, GVIX is not just an index of ex-ante return volatility but a tradable financial product.

asymmetric parameters of return-distribution, thus represents the movement of market sentiment over time.

[Insert Plot 2 here]

1.6   Conclusion

The central contributions of this chapter are threefold. First, we identify that without any diffusion assumption, the model-free formulation of the CBOE volatility index (VIX) is actually a linear combination of ex-ante return-moments, not the expected volatility. Precisely, VIX constitutes additionally one-third value of the ex-ante third return-moment. This indicates that VIX could considerably understate the true volatility when the market expectation of returns is negatively skewed. Notably, since skewness continues to occupy a prominent role in financial markets, VIX's potential bias of volatility estimation due to skewness is not insignificant.

Second, we present an alternative model-free methodology, extended from Bakshi, Kapadia, and Madan (2003), for measuring ex-ante volatility. This new method, named generalized VIX (GVIX), is generic because it is based on the direct formulation of variance, and there is no stochastic assumption on the return generating process of the underlying asset. Therefore, GVIX serves as a proxy for the true ex-ante volatility. We show that VIX is indeed a special case of GVIX, and the calculation of GVIX is as simple as that of VIX.

Third, we find statistically that the time-series of the spread between GVIX and VIX (named GV-Spread) is stationary and mean-reverting. However, the individual time-series of VIX and GVIX appear to be non-stationary or random walk. This indicates the existence of a co-integration system between VIX and GVIX. Intuitively, the inequality between GVIX and VIX characterizes the asymmetric (skewed) expectation of future returns, and the mean-reversion of the GV-Spread captures the back-and-forth movement of the market sentiment over time.

In addition, we empirically investigate our theoretical arguments using S&P500 option data over a period from January 2005 to May 2014. The analytical results support our theories that VIX did statistically understate the true volatility, and the estimation errors were strongly and negatively correlated with the third return-moment or skewness. Importantly, the volatility underestimation of VIX is significantly enlarged as market volatility increased, and the magnitude of the undervaluation could reach to as much as 559 index basis points. Consequently, theoretical and empirical evidence suggests that the VIX formulation should be applied cautiously when estimating ex-ante return volatility as well as for developing volatility based financial products.

References

Bakshi, G., Kapadia, N. and Madan, D., 2003. Stock return characteristics, skew laws, and the differential pricing of individual equity options. The Review of Financial Studies, 16(1), pp.101-143.

Bates, D.S., 1991. The Crash of '87: Was It Expected? The Evidence from Options Markets. The journal of finance, 46(3), pp.1009-1044.

Black, F. and Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of political economy, 81(3), pp.637-654.

Breeden, D.T. and Litzenberger, R.H., 1978. Prices of state-contingent claims implicit in option prices. Journal of business, pp.621-651.

Britten‐Jones, M. and Neuberger, A., 2000. Option prices, implied price processes, and stochastic volatility. The Journal of Finance, 55(2), pp.839-866.

Carr, P. and Madan, D., 2001. Towards a theory of volatility trading. Option Pricing, Interest Rates and Risk Management, Handbooks in Mathematical Finance, pp.458-476.

Carr, P., and Lee, R., 2009. Volatility derivatives. Annu. Rev. Financ. Econ., 1(1), pp.319-339.

Carr, P., and Liuren W., 2006. A tale of two indices. Journal of Derivatives 13, 13–29.

Carr, P., and Wu, L., 2008. Variance risk premiums. The Review of Financial Studies, 22(3), pp.1311-1341.

Demeterfi, K., Derman E., Michael K., and Joseph Z., 1999, More than you ever wanted to know about volatility swaps, Quantitative Strategies Research Notes, Goldman Sachs.

Dickey, D., and Wayne F., 1979, Distribution of Estimators for Autoregressive time series with a unit root, Journal of the American Statistical Association 74, 427-431.

Du, J. and Kapadia, N., 2012. Tail and volatility indices from option prices. Unpublished working paper. University of Massachusetts, Amhurst.

Dupire, B., 1994. Pricing with a smile. Risk, 7(1), pp.18-20.

Exchange, C.B.O., 2009. The CBOE volatility index-VIX. White Paper, pp.1-23.

Jiang, G.J. and Tian, Y.S., 2005. The model-free implied volatility and its information content. The Review of Financial Studies, 18(4), pp.1305-1342.

Jiang, G.J. and Tian, Y.S., 2007. Extracting model-free volatility from option prices: An examination of the VIX index. Journal of Derivatives, 14(3).

Martin, I., 2011. Simple variance swaps (No. w16884). National Bureau of Economic Research.

Neuberger, A., 1994. oThe Log Contract: A New Instrument to Hedge Volatility. p Journal of.

Phillips, P.C. and Perron, P., 1988. Testing for a unit root in time series regression. Biometrika, 75(2), pp.335-346.

Zheng, Z., Jiang, Z. and Chen, R., 2017. AVIX: An Improved VIX Based on Stochastic Interest Rates and an Adaptive Screening Mechanism. Journal of Futures Markets, 37(4), pp.374-410.

Zivot, E. and Andrews, D.W.K., 2002. Further evidence on the great crash, the oil-price shock and the unit-root hypothesis. Journal of business & economic statistics, 20(1), pp.25-44.

Tables and Plots

<div align="center">

Table 1
Consistency between VIX estimate and CBOE VIX index
</div>

Panel A reports the quantile estimates of observed daily closed CBOE's VIX index (VIX_{CBOE}) and our calculated daily VIX (VIX_a) during a period from January 2005 to May 2014. Panel B reports the bootstrapped means and their associated Z-statistics of the difference between VIX_{CBOE} and VIX_a. By selecting randomly 1000 data from 2362 daily observations, we calculate the mean and the associated Z-statistic. This procedure is then repeated 10,000 times for computing the average of means and that of Z-statistics. Panel C provides results from a K-S nonparametric test for the equivalence of probability distributions between VIX_{CBOE} and VIX_a.

Panel A: Quantile estimates

	Quintile				
	Minimum	*25%*	*50%*	*75%*	*Maximum*
$\widehat{VIX}_{CBOE}$	9.89	13.51	17.50	23.68	80.86
Corresponding $\widehat{VIX}_a$	9.87	13.56	17.50	23.66	80.70
Difference	0.02	-0.05	0.00	0.02	0.16

Panel B: Bootstrapping test for mean difference between VIX_{CBOE} and VIX_a

Average mean difference	*Average Z-statistics*
0.0015	0.31

Panel C: Two-sample Kolmogorov-Smirnov Test

*D-Statistic**	*p-value*
0.0085	1.0

* Critical values of 1%, 5% and 10% are 0.0474. 0.0396, 0.0355, respectively.
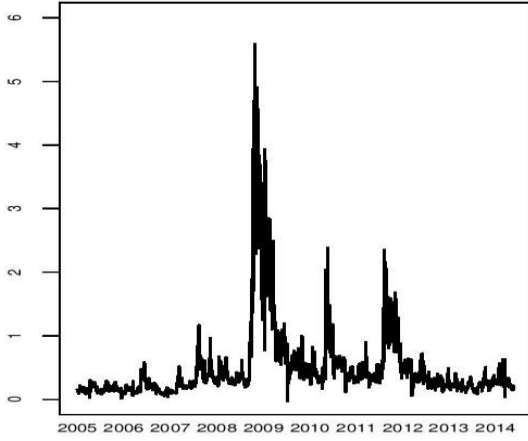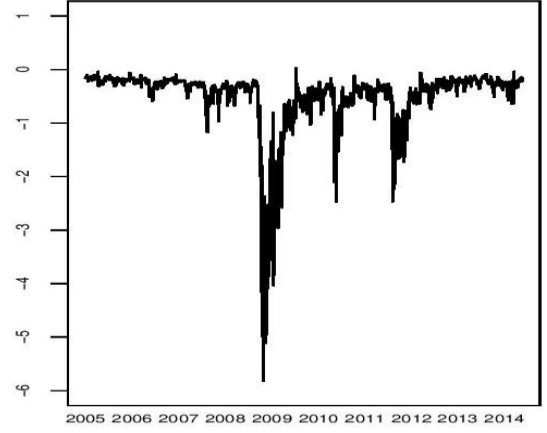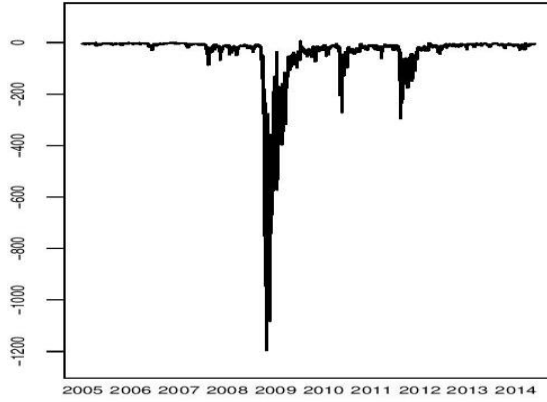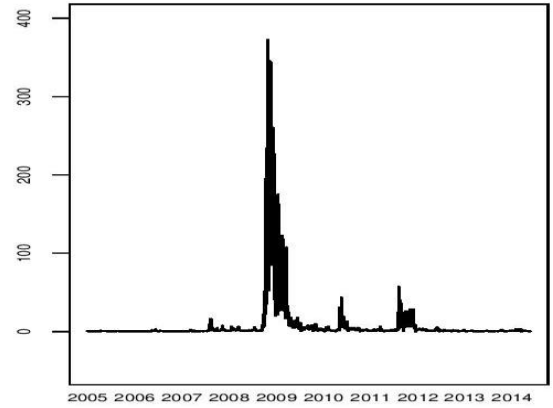
Plot 1. VIX Bias, Ex-ante Volatility and High Moments (2005-2014). These Plots present daily $\widehat{GVIX}_a$, $\widehat{VIX}_a$, ex-ante (annualized) moment measures and their differences from January 2005 to May 2014, where subscript letter "a" represents annualization. Plot 1A depicts the difference between GVIX and VIX. Since GVIX is a direct measure of ex-ante volatility of log-return as shown in equations (20) and (27), $(\widehat{GVIX}_a - \widehat{VIX}_a)$ represents the deviation of VIX from the true volatility. We refer this deviation as VIX-bias. Plot 1B plots the divergence of VIX from the second moment (V_a). Plot 1C pictures the movement of the *third moment*, measured by one-third of $\widehat{W}_a$. Plot 1D displays the changes of the *fourth moment*, measured by one-twelfth of $\widehat{X}_a$. Notably, from Theorem 2, $\text{VIX}_a \approx \sqrt{V_a + \frac{1}{3}W_a + \frac{1}{12}X_a}$.

Table 2
The average VIX-Biases and ex-ante high moments

This table reports the bootstrapped *means* (*Z-statistics*) of the difference between annualized (extrapolated) VIX calculated from daily option premiums and the annualized GVIX as well as the second return-moment (V_a), respectively. In addition, the estimates of ex-ante third and fourth-moment ($\widehat{W}_a$, and $\widehat{X}_a$) are presented in this table. Note that $\widehat{GVIX}_a$ measures the ex-ante standard deviation of log-return (the true volatility) as shown in equations (20) and (27). Also, from (24), $(VIX^2 - GVIX^2) \approx \frac{1}{T}\left(\mu_T^2 + \frac{W_T}{3} + \frac{X_T}{12}\right)$. For each iteration, we select randomly 700 (out of 754), 900 (out of 1008), 500 (out of 600) and 2000 (out of 2362) daily observations from the three sub-periods, 2205-2007, 2008-2011 and 2012-2014 as well as the overall period, respectively. The sample mean and its associated Z-statistic are then calculated. We repeat this procedure 10,000 times and compute the average of means and that of Z-statistics. The numbers in parenthesis are average Z-statistics. Daily option premiums over a period from January 2005 to May 2014 are used for the calculation. * denotes significantly different from zero at the 1% level.

	2005-2007	**2008-2011**	**2012-2014**	**Overall**
No. of days	754	1008	600	2362
$\widehat{GVIX}_a - \widehat{VIX}_a$	0.22 (41.38)*	0.83 (29.59)*	0.28 (54.91)*	0.49 (35.24)*
$\widehat{VIX}_a - \sqrt{\widehat{V}_a}$	-0.26 (-54.77)*	-0.86 (-28.95)*	-0.29 (-54.53)*	-0.52 (-35.68)*
$\frac{1}{3}\widehat{W}_a$	-7.09 (-21.96)*	-82.76 (-15.72)*	-10.42 (-35.70)*	-40.24 (-16.41)*
$\frac{1}{12}\widehat{X}_a$	0.61 (14.97)*	18.38 (10.92)*	0.90 (26.20)*	8.27 (10.84)*

Table 3
Correlation matrix of VIX bias and distribution moments

The VIX-bias is measured by $(\widehat{GVIX}_a - \widehat{VIX}_a)$. $\widehat{V}_a$, $\widehat{W}_a$, and $\widehat{X}_a$ measures daily (annualized) ex-ante second, third and fourth-moment, respectively.

Panel A: (2005-2007)	$\widehat{V}_a$	$\frac{1}{3}\widehat{W}_a$	$\frac{1}{12}\widehat{X}_a$
$\widehat{GVIX}_a - \widehat{VIX}_a$	0.91	-0.97	0.89
$\widehat{V}_a$		-0.93	0.80
$\frac{1}{3}\widehat{W}_a$			-0.95
Panel B: (2008-2011)			
$\widehat{GVIX}_a - \widehat{VIX}_a$	0.95	-0.97	0.92
$\widehat{V}_a$		-0.95	0.88
$\frac{1}{3}\widehat{W}_a$			-0.97
Panel C: (2012-2014)			
$\widehat{GVIX}_a - \widehat{VIX}_a$	0.82	-0.97	0.96
$\widehat{V}_a$		-0.91	0.84
$\frac{1}{3}\widehat{W}_a$			-0.98
Panel D: Overall			
$\widehat{GVIX}_a - \widehat{VIX}_a$	0.96	-0.97	0.90
$\widehat{V}_a$		-0.95	0.86
$\frac{1}{3}\widehat{W}_a$			-0.97
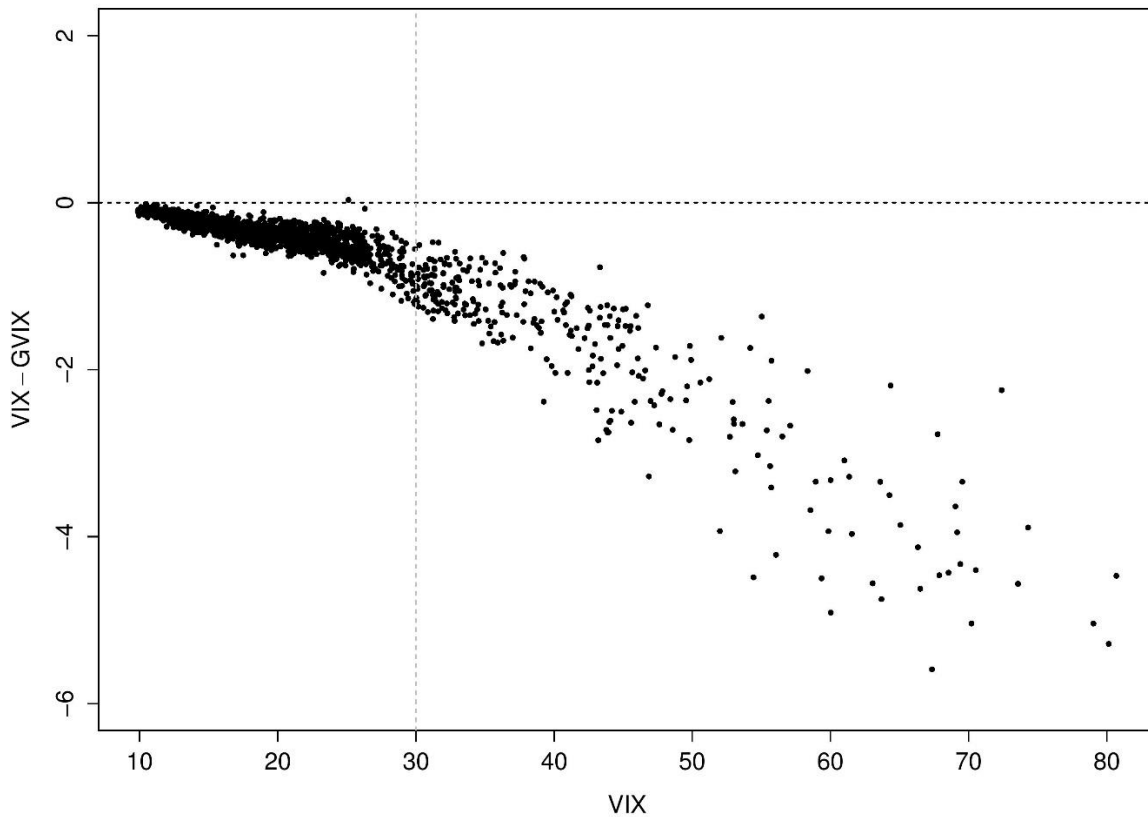
Plot 2. VIX downward bias. This Plot plots the VIX downward bias, measured by $\left(\widehat{\text{VIX}}_a - \widehat{\text{GVIX}}_a\right)$, corresponding to different levels of VIX values. For a numerical illustration, coincided with the Plot, we present, in the following table, the minimum, medium and maximum value of VIX-bias, measured by $\left(\widehat{\text{GVIX}}_a - \widehat{\text{VIX}}_a\right)$, according to three levels of VIX: $\widehat{\text{VIX}}_a < 30$, $30 \le \widehat{\text{VIX}}_a < 50$, and $\widehat{\text{VIX}}_a > 50$, respectively. It has been widely viewed that VIX values greater than 30 are generally associated with a large amount of volatility as a result of investor fear, while values below 20 generally correspond to less stressful times in the markets.

VIX-bias (GVIX – VIX)								
$\widehat{\text{VIX}}_a < 30$			$30 \le \widehat{\text{VIX}}_a < 50$			$\widehat{\text{VIX}}_a > 50$		
Minimum	Median	Maximum	Minimum	Median	Maximum	Minimum	Median	Maximum
-0.03	0.28	1.21	0.47	1.26	3.28	1.36	3.38	5.59

To see the change of sensitivity of VIX-bias to the level of VIX, we let $I_1 = 1$, if $\widehat{\text{VIX}}_a < 30$, and 0 otherwise as well as $I_2 = 1$, if $\widehat{\text{VIX}}_a \ge 30$, and 0 otherwise. To this end, we estimate an ordinary least squares regression: We presents the results from a OLS regression, $\left(\widehat{\text{VIX}}_a - \widehat{\text{GVIX}}_a\right) = \alpha + \beta_1(\widehat{\text{VIX}}_a \times I_1) + \beta_2(\widehat{\text{VIX}}_a \times I_2) + \epsilon$, as follows:

α	$t(\alpha)$	VIX < 30		VIX $\ge$ 30		
		β_1	$t(\beta_1)$	β_2	$t(\beta_2)$	R^2
0.38	21.9	-0.039	-40.4	-0.053	-104.2	0.86

22

Table 4
Unit-root tests for GVIX, VIX and GV-Spread

This table reports union-root tests for the time series of GVIX, VIX and GV-Spread, respectively. The augmented Dickey-Fuller (ADF) tests a unit root present in an autoregressive progress. Phillips–Perron (PP) test makes a non-parametric correction for any serial correlation and heteroskedasticity in the regressive errors by modifying the ADF test statistics. Zivot-Andrews (ZA) test examines unit-root of time series with endogenous structural break. Specifically, ZA test is a sequential test which utilizes the full sample and uses a different dummy variable for each possible break date. The break date is selected where the t-statistic from the ADF test of unit root is at a minimum (most negative). This table presents the least ZA test statistic of each time series of GVIX, VIX, and GV-Spread, respectively. * denotes significant rejection of union-root at the 5% level.

	Dickey-Fuller	Phillips-Perron	Zivot-Andrews
A: GVIX			
Test-Statistics	-1.79	-3.42*	-4.80
B: VIX			
Test-Statistics	-1.75	-3.42*	-4.80
C: GV-Spread			
Test-Statistics	-3.93*	-5.07*	-6.15*
5% *Critical Value*	-1.95	-2.86	-4.80

GVIX

VIX
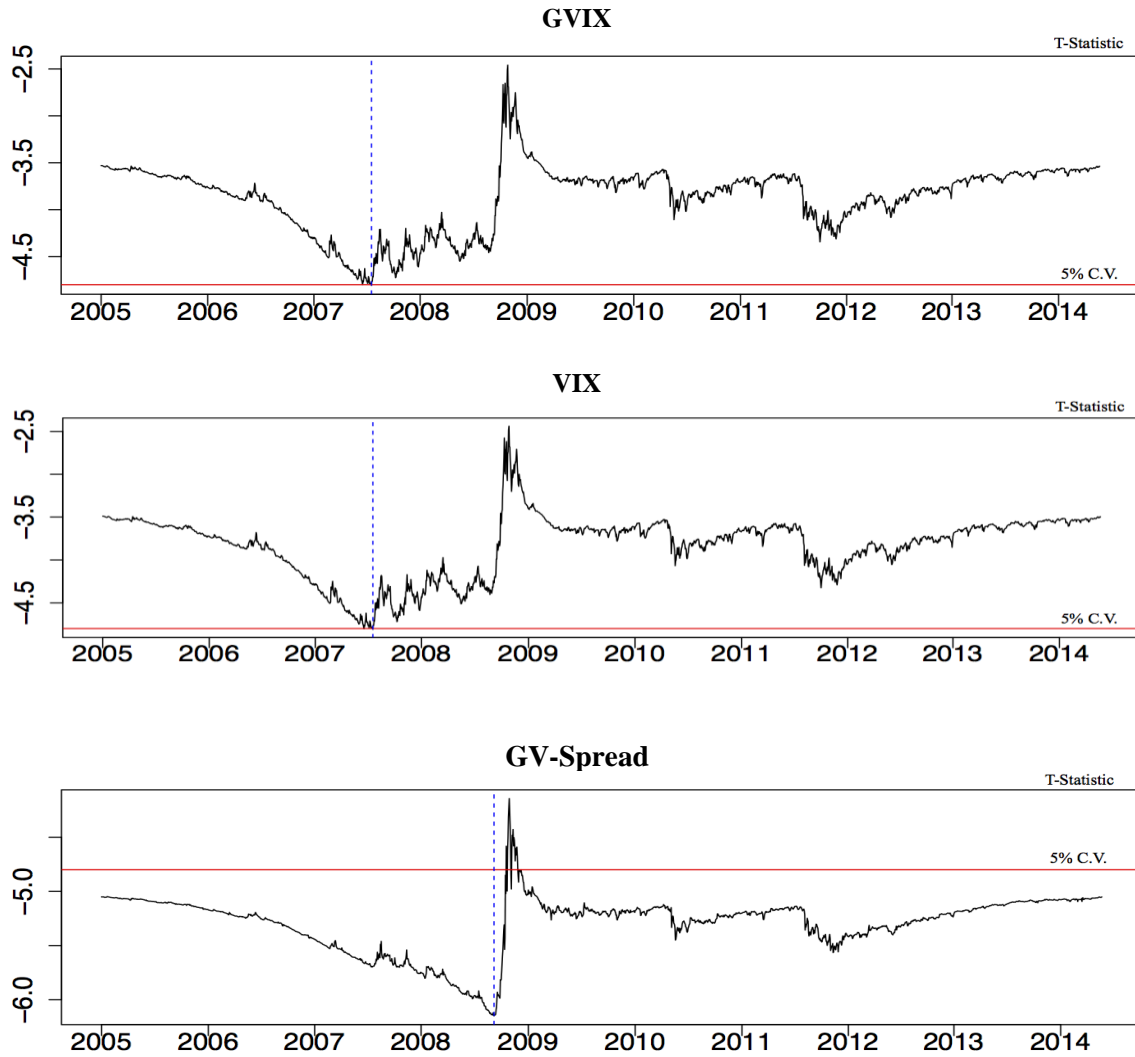
GV-Spread

Plot 3. Trend Stationarity of GV-Spread with a break. The Plot depicts the *time plots* of Zivot-Andrews (ZA) test statistic (*daily t-statistics*) over the entire sample period from January 2005 to May 2014. The null hypothesis of ZA test is a unit-root process without any exogenous structural breaks, and the alternative hypothesis is a trend-stationary process with possible structural change occurring at an unknown point in time. The 5% critical value is -4.80. Therefore, any *daily t-statistic* that has a value below (above) -4.80 is statistically significant (insignificant). The break date of the trend is indicated by the minimum test statistic and marked by a vertical "dot-line". The breakpoint date for the GV-Spread is September 8th, 2008.

Chapter 2   Decomposing the VIX: Implications for the Predictability of Stock Returns

Chapter Abstract

The VIX index is not only a volatility index but also a polynomial combination of all possible higher moments in market return distribution under the risk-neutral measure. This chapter formulates the VIX as a linear decomposition of four fundamentally different elements: the realized variance (RV), the variance risk premium (VRP), the realized tail (RT), and the tail risk premium (TRP), respectively. The VRP compensates the anticipated (normal) market volatility, and the TRP prices the potentially (unusual) large and asymmetric market movements. The chapter uses an innovative and nonparametric tail risk measure and finds that approximately one-third of the VIX's formation is attributed to the TRP. In addition to VRP, RT and TRP are crucial components for predicting future returns on equity portfolios.

2.1   Introduction

The VIX index enjoys tremendous popularity as a risk-neutral, forward-looking measure of the market's return volatility, and is a key driver of the equity variance risk premium (VRP) in Bollerslev, Tauchen, and Zhou (2009). Specifically, the VRP is calculated as the difference between the physical measure of the realized variance (RV) and the square of the VIX, and it serves as an important indicator of aggregate risk aversion of market participants.[9] Recent empirical evidence suggests that the VRP is a superior predictor of future aggregate market returns compared to the traditional predictor variables such as the dividend-price ratio and other valuation ratios, particularly for shorter time horizons.[10] Interestingly, Bollerslev, Tauchen, and Zhou (2009) find that neither the square of the VIX nor the RV is a good predictor of stock market returns, but that their difference (the VRP) is. So, a puzzle emerges. If neither the square of the VIX nor the RV can predict stock returns, then why does their difference provide such strong predictive power? This chapter attempts to unravel the puzzle.

Specifically, the VIX index is not just a measure of volatility (that is, a pure measure of the second moment of a return distribution); it is also a polynomial combination of all possible higher moments in the market return distribution under the risk-neutral measure. Therefore, to explain the puzzle, one needs to find a way to carve out the impact of these higher moments. To this end, this chapter provides a novel methodology for decomposing the VIX index and documents that it is indeed the higher moments, that is the tail-risk components of the VIX, that are driving the returns.

The VIX index was originally designed to measure the quadratic variation ($\mathbb{QV}$) of a jump-free process.[11] Nevertheless, Du and Kapadia (2012) and Chow, Jiang, and Li (2014) observe that the VIX index rapidly deviates from the true volatility measure when a larger proportion of stock return variability is determined by substantial jumps of returns. Also, the deviation of VIX from $\mathbb{QV}$ estimation is proportional to the jump intensity. In fact, it has often been overlooked that Bakshi-Kapadia-Madan's (2003) measure of variance (V^{BKM}) is insensitive to tail variation and can serve as an unbiased ex-ante estimate of $\mathbb{QV}$.[12] A question then arises: if the VIX is not simply a $\mathbb{QV}$ (volatility risk) measure, then what does it truly measure?

Empirical findings of Todorov and Tauchen (2011) suggest that the volatility risk either coincides or is highly correlated with the price jump risk, while Bollerslev and Todorov (2011) show that the risk premium for unusual tail events cannot be explained exclusively by the level of volatility and argue that the jump-tail risk is still present even if the investment opportunity set is

---

[9] See Campbell and Cochrane (1999); Bekaert and Engstrom (2010); Bollerslev, Gibson and Zhou (2011); Bekaert, Hoerova; and Lo Duca (2013); and Bekaert and Hoerova (2014).

[10] These studies include but are not limited to Bollerslev, Tauchen and Zhou (2009); Drechsler and Yaron (2010); Han and Zhou (2012); Du and Kapadia (2012); Andreou and Ghysels (2013); Bondarenko (2014); Eraker and Wang (2015); Almeida, Vicente, and Guillen (2013); Bekaert and Hoerova (2014); Bali and Zhou (2016); Camponovo, Scaillet, and Trojani (2014); Kelly and Jiang (2014); Li and Zinna (2018); Vilkov and Xiao (2013) and Bollerslev, Marrone, Xu, and Zhou (2014).

[11] See Carr and Madan (2001); Demeter, Derman, Kamal and Zou (1999a, 1999b); and Britten-Jones and Neuberger (2000).

[12] Du and Kapadia (2012) and Chow, Jiang and Li (2014) explicitly demonstrate that the Bakshi, Kapadia and Madan (2003)'s measure of the variance of the holding period return is the most appropriate for measuring quadratic variation.

approximately constant. Cremers, Halling, and Weinbaum (2015) also show that aggregate jump and volatility risk collectively explain variation in expected returns, and aggregate stock market jump risk is priced in the cross section. Thus, volatility and price jump-tail risk premia share compensations for similar risks and therefore should be modeled jointly. Recently, Bollerslev, Todorov, and Xu (2015) reveal that most of the predictability for the aggregate market portfolio previously attributed to the VRP stems from not just the volatility but the tail risk component; and that the compensation for tail risk drives out most of the predictability stemming from the part of the VRP associated with "normal" sized price fluctuations.[13] Intuitively, the compensation demanded by investors for bearing tail risk contributes to the expectation as well as the predictability of future market returns.

The main goals of this chapter are twofold. First, by explicitly recognizing the underlying stochastic process of the VIX index that follows the polynomial (not quadratic) variation, we formulate the (squared) VIX as a linear decomposition of four fundamentally different elements: the RV, the VRP, the realized tail (RT), and the tail risk premium (TRP). Through the process of VIX decomposition, we are able to differentiate the TRP from the VRP, both of which are embedded in the VIX index. Second, relying on our decomposition of the VIX index, we seek to clarify where the inherent market return predictability of the conventional variance risk premium, or VRP^c (i.e., VIX^2 - RV), is coming from and how it plays out over different return horizons and for various portfolios with different risk exposures.

Our empirical results confirm that the return predictability for the aggregate market portfolio afforded by the VRP^c is attributed to the return predictability of the decomposed components: the unbiased VRP, the RT, and the TRP. Importantly, the tail variation and its risk premia do not just offer some additional predictability for the market portfolio over and above that of the VRP but also provide the main impetus for the total predictability. This is consistent with recent findings of Bollerslev, Todorov, and Xu (2015) that most of the predictability for market return previously ascribed to the VRP originates from the tail risk component.

This chapter differentiates itself from Bollerslev, Todorov, and Xu (2015) in important ways. First, we formally define TRP through VIX decomposition and through quantifying a converged tail risk measure. The polynomial combination of all possible moments with orders higher than two boils down to one simple analytical form, the difference between the squared realized VIX and the realized variance. Second, the methodology used to estimate market TRP in this chapter is nonparametric in nature, which ensures a more accurate estimation process. Third, we find even greater increases in the predictive performance of RT and TRP from decomposed market portfolios: Size, Value, and Momentum as well as Industrial Sectors. In summary, the significant empirical evidence of the market returns predictability of the VRP previously documented in the literature is dominated by the predictability of the TRP in this chapter and, more importantly, to a larger extent.

---

[13] In addition, several papers have related jump-tail risk to asset risk premia. For example, Naik and Lee (1990); Longstaff and Piazzesi (2004); Liu, Pan, and Wang (2004); Bollerslev and Todorov (2011, 2014); Kelly and Jinag (2014); and Andersen, Fusari and Todorov (2015) model jump-tail risk premia in equity returns, while Gabaix (2012) and Wachter (2013), extending initial work of Rietz (1988) and Barro (2006), relate equity risk premia to time-varying consumption disaster risk.

The rest of the chapter is organized as follows: Section 2.2 begins with a simple derivation of the VIX formulation, wherein the realized VIX and polynomial variation are formally defined. A simple approach for determining the market TRP as well as our decomposition of the VIX index are also presented, and we show the sample estimating procedures of the statistics in Section 2.3. Section 2.3 also presents the statistical estimations for both unconditional and conditional risk premiums of return variation. Section 2.4 describes the data and illustrates our empirical analysis of the VIX decomposition. Section 2.5 reports our empirical findings of equity return predictability of the four VIX decomposed components, Section 2.6 performs several robustness checks. Section 2.7 contains brief concluding remarks.

2.2   VIX Decomposition

The Chicago Board Options Exchange's (CBOE) VIX index is the most widely used option-based (forward-looking) measure of stock return variability. Nevertheless, it is well known that the index contains compensation for risk in addition to that for time-varying volatilities. Those include risk premium of jump intensities as well as that of jump-tail events. As such, this does lend acceptance to the common use of the term "investor's fear gauge" as an epithet for the VIX volatility index, although admittedly an imperfect proxy. This section presents an unambiguous approach to distinguish risk between volatility and the tail variation embedded in the VIX index. We begin with a simple formulating process of the VIX index.

2.2.1   A Simple VIX Formulation

Without any specification of the return generating process, Chow, Jiang, and Li (2014) show that the formulation of VIX can be derived mathematically and straightforwardly as follows: Let $R_{t+1} \left(= \frac{S_{t+1}-S_t}{S_t}\right)$ be the forward arithmetic return and $r_{t+1} \left(= \ln\left(\frac{S_{t+1}}{S_t}\right)\right)$ denote the logarithmic forward return over a period from t to $t+1$. Employing the Taylor series expansion and the expansion with the remainder, the difference between the arithmetic and logarithmic returns can be expressed as follows:

$$R_{t+1} - r_{t+1} = \left[\int_{S_t}^{\infty} \frac{1}{K^2}(S_{t+1}-K)^+ dK + \int_0^{S_t} \frac{1}{K^2}(K-S_{t+1})^+ dK\right] = \sum_{n=2}^{\infty} \frac{1}{n!} r_{t+1}^n. \qquad (1)$$

Now, let $\mathbb{Q}$ denote the risk-neutral distribution associated with the time dynamic of forward returns. Under the no-arbitrage framework, the time-series conditional expected return-difference can be measured by current option prices, which is equivalent to the basic formulation of the (squared) VIX:[14]

---

[14] Under a purely continuous process of the quadratic variation, equation (2) serves as a basis for the derivation of the VIX. See Carr and Madan (2001); Demeterfi, Derman, Kamal, and Zou (1999a, 1999b); and Britten-Jones and Neuberger (2000), and others.

$$E_t^{\mathbb{Q}}(R_{t+1} - r_{t+1}) = e^{r_f}\left\{\int_{S_t}^{\infty} \frac{1}{K^2} C_{t,t+1}(K)dK + \int_0^{S_t} \frac{1}{K^2} P_{t,t+1}(K)dK\right\}$$

$$= \frac{1}{2}\text{VIX}_t^2 = \frac{1}{2}\left[E_t^{\mathbb{Q}}(r_{t+1}^2) + \sum_{n=3}^{\infty} \frac{2}{n!} E_t^{\mathbb{Q}}(r_{t+1}^n)\right] \qquad (2)$$

where $E_t^{\mathbb{Q}}(\cdot)$ is the risk-neutral conditional expectation operator at time t, r_f is the annualized risk-free rate corresponding to expiration date $t+1$, and $C_{t,t+1}(K)$ and $P_{t,t+1}(K)$ are the current (at time t) premiums of call and put option contracts with a strike K and expiration $t+1$, respectively. That is, the arbitrage-free argument implies that the VIX index can be extracted from the market price of a portfolio composed of all possible out-of-the-money (OTM) call/put options of the underlying index with weight inversely proportional to the square value of the strike price. Equivalently, equation (2) shows that instead of employing a long list of OTM options, the VIX also can be simply replicated by a portfolio of only two assets: a long position of a forward contract with a settlement price, S_{t+1} and a short position of a log contract with a settlement price, $\ln(S_{t+1})$, where the log contract has been proposed by Neuberger (1994) for hedging volatility.[15]

2.2.2 The Polynomial Variation and the Realized Tail

The most notable result from equation (2) is that the VIX index, calculated from the fair market price of either an options portfolio or that of long-short forward contracts, provides not only a forward-looking estimate of the market volatility but information about the future return distribution in its entirety. The distributional information in addition to the volatility (the second moment) is characterized by a polynomial combination of a series of all higher distributional moments (e.g. skewness, kurtosis, etc.). This aggregate of high moments implanted in the VIX formulation perhaps explains why the VIX index is often referred to as the investor's fear gauge. To examine and analyze the VIX index as a market fear indicator, decomposing the index regarding different risk characteristics is necessary. For convenience, we define $RVIX_{t+1}$ as the future realized outcomes of the VIX such that

$$RVIX_{t+1}^2 = 2(R_{t+1} - r_{t+1}). \qquad (3)$$

Then, the (squared) VIX is a conditionally risk-neutral estimate of twice the future arithmetic and logarithmic return differences (as called $RVIX_{t+1}$):

$$VIX_t^2 = E_t^{\mathbb{Q}}(RVIX_{t+1}^2). \qquad (4)$$

Next, following the classical approach and without losing generality, we assume that asset returns follow Merton's (1976) diffusion-jump process:

$$R_{t+1} = \int_t^{t+1}(\alpha_t - \lambda\mu_J)dt + \int_t^{t+1}\sigma_t dW_t + \int_t^{t+1}\int_{\mathbb{R}^0}(e^x - 1)\,\mu[dx, dt]. \qquad (5)$$

---

[15] Precisely, the replicated portfolio consists $\frac{1}{S_t}$ long position for every short position.

$$r_{t+1} = \int_t^{t+1} \left(\alpha_t - \frac{1}{2}\sigma_t^2 - \lambda\mu_J \right) dt + \int_t^{t+1} \sigma_t dW_t + \int_t^{t+1} \int_{\mathbb{R}^0} x\,\mu[dx, dt], \tag{6}$$

where α_t is the instantaneous expected return of the asset, σ_t is the volatility, W_t is standard Brownian motion, $\mathbb{R}^0$ is the real line excluding zero, and $\mu[dx, dt]$ is the Poisson random measure for the compound Poisson process with the compensator equal to $\lambda \frac{1}{\sqrt{2\pi\sigma_J^2}} e^{-\frac{1}{2}(x-\alpha)^2}$, with λ as the jump intensity. Now, by taking the square of (6) and based on the Brownian properties, the future quadratic return, r_{t+1}^2, can be expressed by a sum of two decomposed components: the integrated value of a continuously instant variance ($\mathbb{CV}$) and that of a discontinuously (or jump) quadratic variability ($\mathbb{JQV}$). This decomposed process of return variability is the $\mathbb{QV}$, and r_{t+1}^2 is the future realized outcome of the quadratic variation (denoted RV_{t+1}).[16] We summarize this as follows:

$$\begin{aligned} \mathbb{QV}_{[t,t+1]} = r_{t+1}^2 &= \int_t^{t+1} \sigma_t^2\,dt + \int_t^{t+1} \int_{\mathbb{R}^0} x^2\mu(dx, dt) \\ &= \mathbb{CV}_{[t,t+1]} + \mathbb{JQV}_{[t,t+1]}. \end{aligned} \tag{7}$$

Carr and Wu (2008) have shown that the theoretical determination of the VIX is inconsistent with the $\mathbb{QV}$ in that

$$VIX_t^2 = E_t^{\mathbb{Q}}\big(\mathbb{QV}_{[t,t+1]}\big) + \frac{1}{2}E_t^{\mathbb{Q}}\left(\int_t^{t+1} \int_{\mathbb{R}^0} (e^x - 1 - x^2)\,\mu[dx, dt] \right). \tag{8}$$

A question then arises: What should be the fundamental process of determining the VIX value? To answer this question, we consider (3), (5) and (6) and define a generalized stochastic process of return variations, polynomial in form, as follows:

Definition 1. *The infinite-order polynomial variation* ($\mathbb{PV}$) *of returns, based on the return generating process of* (5) *and* (6), *from time t to $t+1$ is defined as*

$$\begin{aligned} \mathbb{PV}_{[t,t+1]} = 2(R_{t+1} - r_{t+1}) &= \int_t^{t+1} \sigma_t^2 dt + \sum_{n=2}^{\infty} \frac{2}{n!} \int_t^{t+1} \int_{\mathbb{R}^0} x^n\,\mu(dx, dt) \\ &= \mathbb{CV}_{[t,t+1]} + \mathbb{JPV}_{[t,t+1]}. \end{aligned} \tag{9}$$

where $\mathbb{JPV}$ denotes a weighted sum of all of the predictable jumps of the $\mathbb{PV}$. The linear combination of all orders of the return variability in (9) characterizes the entire probability distribution of R_{t+1}, and thus the $\mathbb{QV}$ is just a special case of the $\mathbb{PV}$, if $n = 2$. It is also important to note that since the continuous component of the polynomial variation converges to that of the $\mathbb{QV}$ under the Brownian motion, $\mathbb{PV}$ equals $\mathbb{QV}$ with the absence of jump.

Theorem 1. *Based on Definition 1 as well as equations* (3) *and* (4), *the theoretical value of the VIX index at time t is the* (*square-rooted*) *risk-neutral estimate of the polynomial variation from*

---

[16] See Andersen *et al.* (2001); Cont and Tankov (2003), and others.

time t to t + 1:

$$VIX_t = \sqrt{E_t^{\mathbb{Q}}\left(\mathbb{PV}_{[t,t+1]}\right)} \tag{10}$$

In short, the VIX index is a risk-neutral forward-looking measure of the polynomial variation of log-returns: not that of the $\mathbb{QV}$. Consequently, the realized variance (RV_{t+1}) is <u>not</u> generally the future realized outcome of the VIX. This highlights the potential bias of the conventional calculation of the VRP by simply taking the difference between the squared VIX and the RV.

Structurally, although polynomial and quadratic variations are similar in form, $\mathbb{PV}$ provides additional information beyond the jump process of return variability. That is, statistically, the difference between $\mathbb{PV}$ and $\mathbb{QV}$ simultaneously captures the asymmetry, tail thickness, and other characteristics of the return distribution. We refer to this difference as the tail variation (hereafter $\mathbb{TV}$) or whose physical measure is the realized tail (RT) of returns:

Corollary 1. *From* (10) *in Theorem 1 and* (7)*, the difference between the polynomial and the quadratic variations of returns characterizes the jump tail variation (denoted $\mathbb{TV}$), which can be measured by the realized tail (denoted* RT)*. The realized tail is a polynomial combination of all possible higher orders (higher than the 2nd order) of log-returns that are calculated by the spread between the squared realized VIX and the realized variance:*

$$
\begin{aligned}
RT_{t+1} \equiv \mathbb{TV}_{[t,t+1]} \quad &= \mathbb{PV}_{[t,t+1]} - \mathbb{QV}_{[t,t+1]} \\
&= [2(R_{t+1} - r_{t+1}) - r_{t+1}^2] = \sum_{n=3}^{\infty} \frac{2}{n!} r_{t+1}^n \, .
\end{aligned} \tag{11}
$$

Note that, based on (6), the higher order of the jump process, x^n for $n > 2$, is equivalent to the same order of the log-returns, r^n for $n > 2$. Therefore, the expected RT is a polynomial sum of all higher order moments of an asset's log-return distribution. Corollary 1 highlights the important relationship between the $\mathbb{QV}$ and the VIX: Under the risk-neutral framework as well as from (4) and (11),

$$E_t^{\mathbb{Q}}\left(\mathbb{QV}_{[t,t+1]}\right) = VIX_t^2 - E_t^{\mathbb{Q}}\left(\mathbb{TV}_{[t,t+1]}\right). \tag{12}$$

Consistent with Proposition 1 of Carr and Wu (2008), we show that the (risk-neutral) conditional $\mathbb{QV}$ is just a tail-free VIX^2. Further, the option based conditional tail variation can then be measured by the spread between the squared VIX and the BKM's unbiased variance measure (V^{BKM}):

$$E_t^{\mathbb{Q}}(RT_{t+1}) = VIX_t^2 - V_t^{BKM} = \sum_{n=3}^{\infty} \frac{2}{n!} E_t^{\mathbb{Q}}(r_{t+1}^n) \tag{13}$$

where

$$V_t^{BKM} \equiv E_t^{\mathbb{Q}}(\mathbb{QV}_{[t,t+1]}) = e^{r_f}\left[\int_{S_t}^{\infty} \frac{2\left[1 - ln\left(\frac{K}{S_t}\right)\right]}{K^2} C_{t,t+1}(K)dK + \int_0^{S_t} \frac{2\left[1 + ln\left(\frac{S_t}{K}\right)\right]}{K^2} P_{t,t+1}(K)dK\right] \quad (14)$$

The spread between the squared VIX and the BKM variance equals the negative value of Du and Kapadia (2012) jump and tail index, the JTIX. Here the V_t^{BKM} serves as an appropriate (risk-neutral) forward-looking measure of the quadratic variation, while the JTIX serves as an appropriate (risk-neutral) forward-looking measure of the TV. As discussed in Du and Kapadia (2012), the JTIX is a short position in a risk reversal and the hedge that a dealer in short variance swaps would buy to protect against the risk of discontinuities.

2.2.3 VIX Decomposition

Following Bollerslev, Tauchen and Zhou (2009)'s basic notion, we define formally three different risk premiums: the VIX risk premium (VIXRP), the unbiased VRP, and the TRP as follows: First,

$$VIXRP_{[t,t+1]} = E_t^{\mathbb{Q}}(\mathbb{PV}_{[t,t+1]}) - E_t^{\mathbb{P}}(\mathbb{PV}_{[t,t+1]}) = VIX_t^2 - E_t^{\mathbb{P}}(RVIX_{t+1}^2), \quad (15)$$

where $E_t^{\mathbb{P}}(RVIX_{t+1})$ is the physical measure of the polynomial variation in the actual probability space $\mathbb{P}$, and VIX_t^2, as shown in (4), is the risk-neutral estimation of $\mathbb{PV}$. Since $\mathbb{PV}$ identifies the overall variation of returns, VIXRP contains both the risk premium of return volatility and that of potentially abnormal variability. Second,

$$VRP_{[t,t+1]} = E_t^{\mathbb{Q}}(\mathbb{QV}_{[t,t+1]}) - E_t^{\mathbb{P}}(\mathbb{QV}_{[t,t+1]}) = V_t^{BKM} - E_t^{\mathbb{P}}(RV_{t+1}). \quad (16)$$

VRP serves as a risk premium proxy for ordinary price fluctuation with normal jumps. Third,

$$TRP_{[t,t+1]} = (VIX_t^2 - V_t^{BKM}) - E_t^{\mathbb{P}}(RT_{t+1}), \quad (17)$$

where

$$E_t^{\mathbb{P}}(RT_{t+1}) = E_t^{\mathbb{P}}(RVIX_{t+1}^2) - E_t^{\mathbb{P}}(RV_{t+1}). \quad (18)$$

TRP is the difference between VIXRP and VRP, which characterizes the compensation for the prospectively unusual jumps of the market return distribution. Finally, the VIX index can then be decomposed into four fundamentally different constituents such that:

$$VIX_t^2 = \left[E_t^{\mathbb{P}}(RV_{t+1}) + VRP_{[t,t+1]}\right] + \left[E_t^{\mathbb{P}}(RT_{t+1}) + TRP_{[t,t+1]}\right]. \quad (19)$$

Intuitively, the first two components of the (squared) VIX index reflect the conditional (physical) expectation of future volatility and the risk compensation of the future variability from normal economic uncertainty. The third and fourth elements characterize the conditional (physical)

TV of returns and the corresponding TRP for compensating the potentially abnormal market variation. The RT and TRP could be negative if market returns are negatively skewed. This implies that the VIX index could understate the true return volatility due to negative RT and TRP, although VIX tends to be highly correlated with return volatility. Importantly, the conventional Bollerslev, Tauchen and Zhou (2009) VRP^c, is also biased toward the true variance risk premium from (19),

$$VRP^c_{[t,t+1]} = VIX^2_t - E^{\mathbb{P}}_t(RV_{t+1}) = VRP_{[t,t+1]} + \left[E^{\mathbb{P}}_t(RT_{t+1}) + TRP_{[t,t+1]}\right]. \qquad (20)$$

It is clear that the widely used VRP^c is influenced by not only the volatility risk premium but the RT and its associated risk premium. Consequently, the impact of tail risk on future market price fluctuation could be the source of the predictability of Bollerslev, Tauchen and Zhou (2009) VRP^c to US aggregate equity returns. This chapter addresses this issue by empirically examining the return predictability of our four decomposed VIX measures.

Traditionally, the past realized variation is often used as the $\mathbb{P}$ estimate of the conditional variation of stock market returns, which is, in fact, an unconditional sample estimate of the historical return variability. Consequently, to ensure the accuracy of risk estimation, developing robust statistical methods for measuring conditional (physical) return variation is necessary. We present our estimation procedures of conditional RVIX, RV, and RT based on Bekaert and Hoerova (2014) forecasting models in the next section.

2.3 Unconditional and Conditional Estimates

For quantifying the actual return variations, standard approaches employ high-frequency price observations, and the time interval $[t-1, t]$ is split into n equally spaced increments. (e.g. 78, 5-minute trading intervals in a day). Let p_t denote the logarithmic price of the asset. The jth intraday return r_j on day t is defined as $r_j = p_{t-1+\frac{j}{n}} - p_{t-1+\frac{j-1}{n}(\Delta)}$. According to Andersen and Bollerslev (1998), the unconditional (ex-post) estimate of the realized variance can be defined:

$$\widehat{RV}_t = \sum_{j=1}^{n} r_j^2 \xrightarrow{p} \mathbb{QV}_{[t-1,t]}, \qquad \text{for } n \to \infty \qquad (21)$$

where $\xrightarrow{p}$ standard for convergence in probability. Analogous to RV estimation, Jiang and Oomen (2008) show that the sum of twice the difference between arithmetic and logarithmic returns convergence in probability are limited to quadratic variation plus jumps in exponential form. Mathematically, that is, $\underset{n \to \infty}{plim} \sum_{j=1}^{n} 2(R_j - r_j) = \mathbb{QV}_{[t-1,t]} + 2\int_{t-1}^{t}[\exp(J_u) - J_u^2 - J_u - 1]dq_u = \mathbb{PV}_{[t-1,t]}$, with J being the jump process. Therefore, the sample estimate of our realized VIX can be calculated as:

$$\widehat{RVIX}^2_t = \sum_{j=1}^{n} 2(R_j - r_j) \xrightarrow{p} \mathbb{PV}_{[t-1,t]}, \qquad \text{for } n \to \infty, \qquad (22)$$

and the asymptotically unbiased, unconditional measure of the RT can thus be computed by $\widehat{RT}_t = \sum_{j=1}^{M}[2(R_j - r_j) - r_j^2] \xrightarrow{p} \mathbb{TV}_{[t-1,t]} = [\mathbb{PV}_{[t-1,t]} - \mathbb{QV}_{[t-1,t]}]$, for $n \to \infty$. Further, the estimation of VIX (denoted $\widehat{VIX}_t$) based on finite option prices can be obtained from CBOE. We also apply the same procedure as the CBOE's VIX formulation to the unbiased variance measure of V_t^{BKM} (denoted $\hat{V}_t^{BKM}$).[17] Then, the calculation of our risk premiums can be summarized as follows:[18]

Unconditional VIX Risk Premium: $\qquad\qquad\qquad \widehat{VIXRP}_t = \widehat{VIX}_t^2 - \widehat{RVIX}_t^2,$

Unconditional Unbiased Variance Risk Premium: $\qquad \widehat{VRP}_t = \hat{V}_t^{BKM} - \widehat{RV}_t,$

Unconditional Tail Risk Premium: $\qquad\qquad\qquad \widehat{TRP}_t = \left(\widehat{VIX}_t^2 - \hat{V}_t^{BKM}\right) - \widehat{RT}_t.$

Economically, the return VRP, as shown in (15), (16), and (17), is the difference between the conditional variation using a risk-neutral probability measure and that uses the actual physical probability measure. Both of the option-based estimates of $\hat{V}_t^{BKM}$ and $\widehat{VIX}_t$ are risk-neutral conditional measures. Conventionally, the physical measures employed are backward-looking (past) sample estimations, where the options based $\mathbb{Q}$ measures are forward-looking. This counterintuitive approach used for calculating VRP could naturally produce biased results. Recently, Bekaert and Hoerova (2014) evaluate a plethora of state-of-the-art volatility forecasting models based on the decomposition of the squared VIX index to produce an accurate measure of the conditional variance. We adopt one of Bekaert and Hoerova (2014) winning models (model 11) as our forecasting model for estimating conditional return variation. Bekaert and Hoerova (2014) model 11 features continuous and jump variations at three frequencies: 1-day, 5-day, and 22-day, respectively, in that the presence of realized variability at all three frequencies is important in delivering lower error statistics. We present the application of Bekaert and Hoerova (2014) model 11 to our variables as follows.

We begin with daily measures of RV, RVIX, and RT, calculated from 5-minute intraday returns as well as an overnight close-to-open return (79 increments in total per day). They are $\widehat{RV}_t^{(1)} = \frac{79}{\kappa}\sum_{i=1}^{\kappa} r_i^2$, $\widehat{RVIX}_t^{2(1)} = \frac{79}{\kappa}\sum_{i=1}^{\kappa} 2(R_i - r_i)$, and $\widehat{RT}_t^{(1)} = \frac{79}{\kappa}\sum_{i=1}^{\kappa} 2(R_i - r_i) - r_i^2$, respectively, where κ is the actual trading increment. Next, the h-day estimate of the continuous as well as the discontinuous components of the quadratic and polynomial variations in (7) and (10) are calculated: $\mathbb{CV}_t^{(h)} = \left(\frac{22}{h}\sum_{j=1}^{h} \widehat{RV}_{t-j+1}^{(1)}\right) - \mathbb{JQV}_t^{(h)}$, $\mathbb{JQV}_t^{(h)} = \frac{22}{h}\sum_{j=1}^{h} max\left[\left(\widehat{RV}_{t-j+1}^{(1)} - TBPV_{t-j+1}^{(1)}\right), 0\right]$, and $\mathbb{JPV}_t^{(h)} = \frac{22}{h}\sum_{j=1}^{h}\left(\widehat{RVIX}_{t-j+1}^{2(1)} - \mathbb{CV}_{t-j+1}^{(1)}\right)$, where $TBPV_t^{(1)}$ stands for the daily threshold bipower variation defined in Corsi *et al.* (2010). Note that we scale up all measures to the monthly (22-day) basis. Then, three rollover series of continuous and discontinuous sample estimates, daily ($h = 1$), weekly ($h = 5$), and monthly ($h = 22$), accordingly, are used as independent variables for the following forecasting models:

$$2\left[R_t^{(22)} - r_t^{(22)}\right] = \alpha + \beta^m \mathbb{CV}_{t-22}^{(22)} + \beta^w \mathbb{CV}_{t-22}^{(5)} + \beta^d \mathbb{CV}_{t-22}^{(1)} \qquad (23)$$

---

[17] See the VIX white paper, URL: http://www.cboe.com/micro/vix/vixwhite.pdf.

[18] All variables are annualized whenever appropriate.

$$+ \gamma^m \mathbb{JPV}_{t-22}^{(22)} + \gamma^w \mathbb{JPV}_{t-22}^{(5)} + \gamma^d \mathbb{JPV}_{t-22}^{(1)} + \varepsilon_t$$

$$
\begin{aligned}
\left[r_t^{(22)} \right]^2 = {}& a + b^m \mathbb{CV}_{t-22}^{(22)} + b^w \mathbb{CV}_{t-22}^{(5)} + b^d \mathbb{CV}_{t-22}^{(1)} \\
& + c^m \mathbb{JQV}_{t-22}^{(22)} + c^m \mathbb{JQV}_{t-22}^{(5)} + c^d \mathbb{JQV}_{t-22}^{(1)} + e_t
\end{aligned}
\tag{24}
$$

and

$$
\begin{aligned}
2\left[R_t^{(22)} - r_t^{(22)} \right] - \left[r_t^{(22)} \right]^2 = {}& \mathcal{A} + \mathcal{C}^m \left[\mathbb{JPV}_{t-22}^{(22)} - \mathbb{JQV}_{t-22}^{(22)} \right] \\
& + \mathcal{C}^w \left[\mathbb{JPV}_{t-22}^{(5)} - \mathbb{JQV}_{t-22}^{(5)} \right] \\
& + \mathcal{C}^d \left[\mathbb{JPV}_{t-22}^{(1)} - \mathbb{JQV}_{t-22}^{(1)} \right] + \epsilon_t
\end{aligned}
\tag{25}
$$

where $R_t^{(22)}$ and $r_t^{(22)}$ are the monthly rollover arithmetic and logarithmic returns over the time interval $[t - 22, t]$, respectively. Consequently, the conditional measures of return variations as well as their risk premiums can be computed using the estimated coefficients from regressions of (23), (24), and (25), accordingly. We summarize the calculation as follows: Let $\overline{RVIX}_t^2 = \hat{E}_t^{\mathbb{P}}(RVIX_{t+22}^2)$, $\overline{RV}_t = \hat{E}_t^{\mathbb{P}}(RV_{t+22})$, and $\overline{RT}_t = \hat{E}_t^{\mathbb{P}}(RT_{t+22})$ be the empirical conditional estimates of next month's return variations.

Conditional VIX Risk Premium:
$$\overline{VIXRP}_t = \frac{1}{12} \widehat{VIX}_t^2 - \overline{RVIX}_t^2, \text{ where}$$
$$
\begin{aligned}
\overline{RVIX}_t^2 = {}& \hat{\alpha} + \hat{\beta}^m \mathbb{CV}_t^{(22)} + \hat{\beta}^w \mathbb{CV}_t^{(5)} + \hat{\beta}^d \mathbb{CV}_t^{(1)} \\
& + \hat{\gamma}^m \mathbb{JPV}_t^{(22)} + \hat{\gamma}^w \mathbb{JPV}_t^{(5)} + \hat{\gamma}^d \mathbb{JPV}_t^{(1)},
\end{aligned}
$$

Conditional Unbiased Variance Risk Premium:
$$\overline{VRP}_t = \frac{1}{12} \hat{V}_t^{BKM} - \overline{RV}_t, \text{ where}$$
$$
\begin{aligned}
\overline{RV}_t = {}& \hat{a} + \hat{b}^m \mathbb{CV}_t^{(22)} + \hat{b}^w \mathbb{CV}_t^{(5)} + \hat{b}^d \mathbb{CV}_t^{(1)} \\
& + \hat{c}^m \mathbb{JQV}_t^{(22)} + \hat{c}^w \mathbb{JQV}_t^{(5)} + \hat{c}^d \mathbb{JQV}_t^{(1)},
\end{aligned}
$$

Conditional Tail Risk Premium:
$$\overline{TRP}_t = \frac{1}{12} \left(\widehat{VIX}_t^2 - \hat{V}_t^{BKM} \right) - \overline{RT}_t, \text{ where}$$
$$
\begin{aligned}
\overline{RT}_t = {}& \hat{\mathcal{A}} + \hat{\mathcal{C}}^m \left[\mathbb{JPV}_t^{(22)} - \mathbb{JQV}_t^{(22)} \right] + \hat{\mathcal{C}}^w \left[\mathbb{JPV}_t^{(5)} - \mathbb{JQV}_t^{(5)} \right] \\
& + \hat{\mathcal{C}}^d \left[\mathbb{JPV}_t^{(1)} - \mathbb{JQV}_t^{(1)} \right].
\end{aligned}
$$

2.4 Empirical Analysis of VIX Decomposition

This section describes data and empirical analysis of our VIX decomposition. Particularly, the focus is on examining the source of the intrinsic market return predictability on different return horizons as well as different decomposed aggregate market portfolios with various types of risk exposures.

2.4.1 Data Description

We employ the aggregate S&P 500 composite index as a proxy for the aggregate market portfolio. Our high-frequency data for the S&P 500 index span the period of January 2, 1990 to October 10, 2014. The prices are recorded at 5-minute intervals, with the first price for the day at 9:30 A.M. and the last price at 4:00 P.M.[19] Along with the close-to-open overnight return, this leaves us with a total of 79 intraday return observations for each of the 5,979 trading days in the sample. The daily VIX index is obtained directly from the website of the CBOE. We also obtain the daily CBOE SKEW index for comparison purposes. The CBOE SKEW index typically ranges from 100 to 150. A SKEW value of 100 means that the perceived distribution of S&P 500 log-returns is normal and the probability of outlier returns is therefore negligible. As SKEW rises above 100, the left tail of the S&P 500 distribution acquires more weight, and the probabilities of outlier returns become more significant.

For calculating V^{BKM}, we use closing bid and ask quotes for all S&P 500 options traded on the CBOE.[20] Further, for analyzing the predictive performance of VRP and TRP on various size, book-to-market, and momentum sorted portfolios, we downloaded return data from Kenneth R. French's data library.[21] Finally, the data of the control variables in our analytical models are from Compustat and the Federal Reserve Bank dataset and the Federal Reserve Bank of St. Louis website.

2.4.2 Sample Estimates of the VIX Decomposed Components

Basic summary statistics for the daily, weekly and monthly measures of return variations and risk premiums are provided in Table 1. In addition to ex-post (unconditional) sample estimates, we calculate the daily conditional measures of return variations using the resulting coefficients from the forecasting regressions of (23), (24) and (25) over the full sample as follows:

$$
\overline{RVIX_t^2} = \begin{array}{c} 12.747 - 0.164\ \mathbb{CV}_t^{(22)} + 0.456\ \mathbb{CV}_t^{(5)} - 0.071\ \mathbb{CV}_t^{(1)} \\ (2.184)\quad (0.132)\qquad\qquad (0.169)\qquad\qquad (0.089) \\[4pt] +\ 0.544\ \mathbb{JPV}_t^{(22)} + 0.370\ \mathbb{JPV}_t^{(5)} - 0.028\ \mathbb{JPV}_t^{(1)}, \\ (0.226)\qquad\qquad (0.316)\qquad\quad (0.041) \end{array} \tag{26}
$$

$$
\overline{RV_t} = \begin{array}{c} 12.752 - 0.171\ \mathbb{CV}_t^{(22)} + 0.487\ \mathbb{CV}_t^{(5)} - 0.086\ \mathbb{CV}_t^{(1)} \\ (2.212)\quad (0.143)\qquad\qquad (0.180)\qquad\qquad (0.091) \\[4pt] +\ 0.530\ \mathbb{JQV}_t^{(22)} + 0.392\ \mathbb{JQV}_t^{(5)} + 0.029\ \mathbb{JQV}_t^{(1)}, \\ (0.229)\qquad\qquad (0.334)\qquad\quad (0.042) \end{array} \tag{27}
$$

$$
\overline{RT_t} = \begin{array}{c} -0.234 - 2.013\left[\mathbb{JPV}_t^{(22)} - \mathbb{JQV}_t^{(22)}\right] - 0.608\left[\mathbb{JPV}_t^{(5)} - \mathbb{JQV}_t^{(5)}\right] \\ (0.195)\quad (1.300)\qquad\qquad\qquad\qquad\qquad (0.890) \end{array} \tag{28}
$$

---

[19] The source of our high-frequency data is from Genesis Financial Technologies.
[20] We obtained options data from Ivolatility.com
[21] Website: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

$$+ \; 0.129 \atop (0.123) \left[\mathbb{JPV}_t^{(1)} - \mathbb{JQV}_t^{(1)} \right];$$

[Insert Table 1 here]

The heteroscedasticity-robust standard errors are reported in parenthesis. Numerically, due to the similarity of scale between $RVIX^2$ and RV, the magnitude of the RT measures is quite small in that the means of daily, weekly and monthly RT are only -0.172, -0.312, and -0.206 percentage points, respectively.[22] Nevertheless, the significant *t*-statistics for all RT estimates indicate that $RVIX^2$ is statistically different from RV, and thus the higher order jump (tail) process of market returns cannot be ignored. Implicitly, it shows that market return variability results from two parts: volatility as well as TV. Therefore, the risk compensation of market variation can be decomposed by the risk premium of the VRP and that of the TRP, accordingly. Empirically, both the VRP and the TRP are statistical non-zero. To illustrate, Plot 1 plots the daily time series of the VRP and the TRP based on conditional measures. We also plot the CBOE VIX index and SKEW index, the BKM volatility, the TJIX, and the unconditional measures of RV and TV for comparison purposes.

[Insert Plot 1 here]

Consistent with empirical evidence in previous studies, the spread between the unbiased implied (risk-neutral, V^{BKM}) and RV is generally positive. We show that the spread between the realized and implied TV, on the other hand, is mostly negative and seems to be highly and negatively correlated with the VRP. The dynamics in the TRP capture compensation for unusually large and asymmetric risks in the market return distribution. From Plot 1, the VIX index and the BKM variance are very close to each other, which results that the average magnitude of the JTIX is relatively small. The TRP is driven more by large volatility values (as shown by the BKM V and Unconditional Realized V in Plot 1) than SKEW values, although some spikes (the spikes around 1990-1991 and 1998-1999) seem correlated to the high SKEW level. Back in Table 1, conditional and unconditional risk premium measures are alike on average. However, conditional (unconditional) VRP tends to be positively (negatively) skewed. This highlights the potential difference between ex-ante and ex-post approaches in market return predictability analysis.

2.4.3 Decomposable Goodness-of-Fit Test

To examine fundamental attributions of the four individual components (RV, VRP, RT, and TRP) to the variation of VIX, we employ Klein and Chow (2013, KC hereafter) decomposed R-square approach. Based on an optimal simultaneous orthogonal data transformation, KC methodology allows us to identify the underlying uncorrelated components of RV, VRP, RT, and TRP, respectively. Since squared VIX is a linear combination of the four decomposed factors as shown in (19), without losing generality and to avoid the problem of multicollinearity, a multiple factor regression model with orthogonally transformed variables can be set up as follows:

---

[22] Bondarenko (2014) also shows the numerical similarity between $\sum_{j=1}^{n} 2(R_j - r_j)$ and $\sum_{j=1}^{n} r_j^2$.

$$VIX_t^2 = \alpha_t + \beta_{RV}^{\perp} RV_t^{\perp} + \beta_{VRP}^{\perp} VRP_t^{\perp} + \beta_{RT}^{\perp} RT_t^{\perp} + \beta_{VRT}^{\perp} TRP_t^{\perp} + e_t, \qquad (29)$$

where $\perp$ denotes variables or coefficients after orthogonal data transformation. Specifically, the essential components of the four factors retain their original variances before orthogonalization, but their cross-sectional covariances are zero. In addition, the multi-factor regression (29) maintains the same coefficient of determination (R-square, i.e. the ratio of systematic variation to the overall variability of the VIX) as that using the fundamental, non-orthogonalized factors. Since R-square represents a goodness-of-fit of the VIX from data of the four components, disentangling the R-square, based on factors' volatility and their corresponding betas, is thus able to determine the individual contribution of to the VIX's variation from different components. Statistically, this decomposition of the R-square can be expressed as:

$$
\begin{aligned}
R_{\text{VIX}}^2 &= \left(\beta_{RV}^{\perp} \frac{\sigma_{RV}}{\sigma_{VIX}}\right)^2 + \left(\beta_{VRP}^{\perp} \frac{\sigma_{VRP}}{\sigma_{VIX}}\right)^2 + \left(\beta_{RT}^{\perp} \frac{\sigma_{RT}}{\sigma_{VIX}}\right)^2 + \left(\beta_{TRP}^{\perp} \frac{\sigma_{TRP}}{\sigma_{VIX}}\right)^2 \\
&= DR_{\text{RV}}^2 \quad + \quad DR_{\text{VRP}}^2 \quad + \quad DR_{\text{RT}}^2 \quad + \quad DR_{\text{TRP}}^2,
\end{aligned}
\qquad (30)
$$

where DR^2 denotes the decomposed R-square. Further, note that from (19), since the squared VIX is a sum of the four factors, R_{VIX}^2 in (30) is one.

[Insert Table 2 here]

[Insert Table 3 here]

Table 2 and 3 report the empirical results of (29) and (30) with the unconditionally and conditionally daily estimations of RV, VRP, RT, and TRP, respectively. For preventing bias results due to measurements at different scales, all variables are standardized for the analyses. Over the sample period of January 1993 to September 2014, the (unconditional) realized volatility characterizes more than two-thirds (66.83 percent) of the VIX daily variation. Notably, the decomposed R-square of the TRP is 26.64 percent, which is almost four times larger than that of the VRP. The impact of both unconditional and conditional RT on VIX's variability appears to be small. This demonstrates that the investors' required compensation of potential significant market movements (tail risk) is a major factor in determining the variation of the VIX.

Nevertheless, as shown in Table 3, the influence of conditional RV to the VIX is much smaller than that of unconditional RV, where the decomposed R-square drops to 29.8 percent. Since the conditional RV is an ex-ante measure calculated from the forecasting model (27), the results of both unconditional and conditional analyses (Tables 2 and Table 3) strongly indicate that although the formulation of the VIX is a forward-looking (options) based measure, the major determinant of the VIX is actually the physically (ex-post) realized volatility.

To further analyze the impact of the decomposed components of the VIX variability under different market conditions, we divide the overall sample into sub-samples based on three distinct levels of the VIX: (1) nervous market condition for VIX $\geq$ 23.32 (75 percentile), (2) normal market condition for $14.17 \leq$ VIX < 23.32, and (3) calm market condition for VIX < 14.17 (25 percentile), correspondingly. It appears that the TRP has the largest influence on the VIX determination during Nervous Market Condition. This suggests that the VIX is not only a volatility index but the market's fear gauge regarding the higher moments of the market return distribution.

2.5   Stock Return Predictability

Mounting empirical evidence suggests that equity market future returns could be predicted by the long-term VRP, defined as the difference between the risk-neutral and the actual expectations (i.e. VIX^2 – RV), especially over a 3- to 6-month time horizon. Bollerslev, Todorov, and Xu (2015) argues that the VRP can be naturally decomposed into two fundamentally different sources of market variance risk: normal size price fluctuations and jump tail risk. Specifically, by differentiating the left and right (risk-neutral) jump components from the $\mathbb{QV}$ based on a threshold of log-jump size, the part of the VRP associated with compensation for left jump (tail) risk may be seen as a proxy for market fears. Bollerslev, Todorov, and Xu (2015) show that the left jump (or tail variation) serves as a predictor variable for market future returns. Instead of discriminating the quadratic jump variation between left and right, we measure tail risk based on the spread between the $\mathbb{PV}$ and $\mathbb{QV}$ (i.e. $\mathbb{TV} = \mathbb{PV} - \mathbb{QV}$).

2.5.1   S&P 500 Index Return Predictability

Following the analytical procedures of Bollerslev, Tauchen and Zhou (2009) and Bekaert and Hoerova (2014), we investigate the relationship between aggregate stock market (the S&P 500 Index) monthly excess returns and a set of lagged predictor variables with a focus on the RT and the TRP. The main predictive variables include the four decomposed VIX risk factors: RV, VRP, RT, and TRP, respectively. In addition, to ensure the robustness of our analysis, we also include a set of control variables employed by Bekaert and Hoerova (2014) that consists of the real 3-month rate (the 3-month T-bill minus CPI inflation, denoted 3MTB), the logarithm of the dividend yield (denoted Log(DY)), the credit spread (the difference between Moody's BAA and AAA bond yield indices, denoted CS) and the term spread (the difference between the ten-year and the 3-month Treasury yields, denoted TS). Table 4 reports two correlation matrices of predictor variables with respect to the unconditional and conditional measures. The RT has relatively low cross-sectional correlations with other variables. It ranges from -0.18 (with VRP) to 0.24 (with TRP) for the unconditional RT, and from -0.28 (with TRP) to 0.50 (with RV) for the conditional RT.

[Insert Table 4 here]

Our main analytical results of stock market predictability appear in Table 5. We employ the standard approaches of Bollerslev, Tauchen and Zhou (2009) and Bekaert and Hoerova (2014)

by regressing excess stock returns (the annualized monthly S&P500 return in excess of the annualized 3-month T-bill rate) against the risk factors described above. All variables, except RT, are expressed in annualized percentages; the RT is expressed in basis points. The analysis is also based on three different horizons, monthly, quarterly and annual (denoted by 1, 3 and 12, respectively), averaging returns over a quarter/year. To correct for serial correlation, the Newey-West t-statistics with a relatively large number of lags is adopted.[23] For each Panel of Table 5, we report the results from simple regressions with respect to each risk variable and their risk premium individually as well as with multiple regressions that consider jointly individual risk factors, its premium, and control variables. Panel A reveals monthly return predictability.

There are fairly different outcomes between unconditional and conditional measurements. Based on a conventional ex-post approach of simple historical (unconditional) estimation, individual t-statistics for all risk factors (except the RT), extending from -2.311 to 3.204, are significant at the 5% level. At the monthly prediction horizon, TRP is significant for both unconditional and conditional measures; this result is in line with the findings in Table 2 and 3 that TRP is of larger significance among the four decomposed VIX components.

[Insert Table 5 here]

Importantly, as shown in Panel A of Table 5, almost an opposite result appears when we employ the Bekaert and Hoerova (2014) conditional approach. From simple return predictability regressions, the RT and TRP are the only significant predictors for future monthly market returns. A similar result holds from the multiple variable regression, except that VRP is significant, where Newey-West t-statistics of conditional VRP, RT, and TRP regressor coefficients are 3.172, 2.651, and 2.513, respectively. By extending the prediction period from a month to a quarter, Panel B of Table 5 shows that from the regression with multiple control variables, both conditional and unconditional RT still retain their predictive power of stock market returns. However, Panel C of Table 5 reports that both conditional RT and TRP fail to predict stock market returns. Therefore, the tail risk factor and its risk premium have predictive power for stock return over a relatively short period of time. On the other hand, the predictability of VRP increases as the time horizon increases from a month to a quarter. In summary, the empirical evidence from Table 5 concludes that from multiple regressions including control variables, RV does not predict S&P 500 index returns for almost all time horizons (except the unconditional monthly prediction). Nevertheless, the time series conditional tail risk factor and its premium proxy, on the contrary, statistically predict the next month's (and quarter's) stock market returns.

Next, consider that the two decomposed components of the VIX risk premium derived from the polynomial variation (i.e., VRP and TRP) are separate potential predictors of stock market returns. To compare the predictability of VRP with that of TRP, we plot the corresponding Newey-West t-statistics and adjusted regression R^2s for all of the 1- through 12-month return regressions in Plot 2. The t-statistics from the simple regressions based on unconditional (conditional) VRP are all significant (insignificant), and the R^2s increase with the return horizons. However, the R^2s of the unconditional VRP regression decreases after they reach the maximum value of 10% at the

---

[23] Econometrics literature such as Newey and West (1987) and Smith (2005) documents that the Newey-West t-statistics are Heteroscedasticity and Autocorrelation Consistent (HAC).

four-month horizon. Consistent with the results in Table 5, the *t*-statistics from the simple regressions based on either unconditional or conditional TRP are significant in the short time horizon (shorter than two-month), and the R^2s decrease with the return horizon. In addition, the adjusted R^2s from the multiple regressions based on both unconditional (conditional) predictor variables are higher but close to those from the simple regressions based on unconditional VRP (conditional TRP) only. In summary, the risk premium of the market return variation contains two components: compensation for economic uncertainty, measured by the VRP, and that for the unusually large and asymmetric market movements, measured by the TRP. To further examine the sources of the predictability, we follow Bollerslev, Todorov, and Xu (2015) by analyzing a series of predictability regressions for various style portfolios.

[Insert Plot 2 here]

2.5.2 Return Predictability of Style Portfolios

Portfolios with different styles represent different risk characteristics and exposures. Therefore, their reaction to a change in aggregate risk and risk-aversion could vary. Table 6 reports the results from multiple regressions based on lagged RV, VRP, TRP, RT, and control variables similar to those in Table 5. The dependent variables are based on monthly excess returns of different style portfolios. The style portfolios are classified by three different risk factors of Fama-French-Carhart: Size, Value/Growth, and Momentum, accordingly. The six equally weighted portfolios, obtained from the data library of Kenneth R. French, comprise the top and bottom quintiles for each of the three different stock sorts according to their market capitalization, book-to-market (B/M) value, and most recent annual return. The predictability analysis is again based on three different horizons: monthly, quarterly and annual.

The most notable result shown in Table 6 is that neither conditional nor unconditional RV predicts style portfolios for all time horizons. Now, we begin with the analysis relating to the size-sorted portfolios. From the monthly and quarterly results, both unconditional and conditional measures of VRP and those of TRP are significant predictors for the small-stock portfolio. The influence of the conditional RT to the small-stock portfolio is insignificant till the predictive time horizon increases to one quarter (one-year), where the *t*-statistics of conditional (unconditional) RT reaches 2.967 (2.536). The predictability of a big-stock portfolio mainly comes from the VRP, although conditional TRP and RT show some influence on monthly and quarterly predictability. Further, the zero-cost long-short portfolio of small minus big (SMB) is a proxy portfolio that removes the market risk but retains only the size effect. From Panels A and B of Table 6, in contrast to Bollerslev, Todorov, and Xu (2015), we find that the TRP contributes to the predictability of the SMB portfolio, where the VRP shows no impact on SMB prediction at all.

For the B/M sorted value and growth portfolios, both the conditional and unconditional VRPs (TRPs) seem to be significant predictors for the monthly and quarterly (annual) returns on the zero-cost High-Minus-Low (HML) portfolios. The *t*-statistics of conditional TRP and RT predictors for the next month returns on the growth (low B/M) portfolios are significant at the 5% percent level. However, this tail risk influence on the value portfolios declines as the predictive time horizon increases. Both the VRP and TRP appear to have an impact on the monthly and

quarterly return prediction for the value portfolios. Our VIX decomposed measures seem to have relatively low predictability for the returns on the momentum (WML) portfolios. Particularly, none of the *t*-statistics of the quarterly predictive regression coefficients is significant. However, the unconditional VRP and RT, as well as the conditional TRP, retains some predictive power on the monthly return prediction of the WML portfolios. Both the winner and loser portfolios have some influence from the VRP, TRP, and RT.

[Insert Table 6 here]

Plot 3 shows the predictability patterns (*t*-statistics and R^2) of VRP (solid lines) and TRP (dashed lines) over time for size, value/growth, and momentum portfolios. The general patterns are similar between unconditional and conditional measures. The impact of TRP (VRP) on SMB appears to be relatively short-term (long-term). For the HML portfolios, the predictive power of TRP seems to be much larger than that of VRP, where the R^2s of TRP for the HML portfolio appear to be maximized at the intermediate four-month horizon. Finally, the pattern of increasing (decreasing) predictability from TRP (VRP) on the WML portfolio indicates that the short-term (long-term) predictability of momentum portfolios is attributable to VRP (TRP). In summary, the results of Table 6 and Plot 3 describe that variance and tail risk have various impacts on portfolios with different fundamental risk exposures. In addition to style portfolios, we further investigate the effects of our decomposed VIX premiums on disintegrative equity market portfolio based on various mutually exclusive industrial sectors.

[Insert Plot 3 here]

2.5.3 Return Predictability of Industrial Portfolios

Table 7 reports results from multiple predictability regressions that include the four conditional measures of the VIX decomposed components (RV, VRP, TRP, and RT) as well as all control variables. Once again, the RV has no influence on return predictability for all sector portfolios. The conditional RT of the S&P index return distribution (RT), on the other hand, significantly attributes monthly return predictability to industrial sectors of non-durables, chemicals, equipment, telecommunication, utilities, and wholesale. By extending the predictive time horizon from a month to a quarter, RT has significant impact on 11 of the 12 sectors. Although both VRP and TRP have predictive power for monthly and quarterly returns on some industrial stocks, it is less significant than the predictability of RT. This suggests that the realized jump-tail could be a significant risk factor in determining future returns on disintegrative market portfolios or even on individual assets. The insignificance of *t*-statistics of all our predictor variables in Panel C of Table 7 suggests that the influence of VRP and TRP to less diversified market portfolios (e.g., industrial equity funds) occurs only in the relative short run. Interestingly, from our empirical outcomes shown in all Panels of Table 7, returns on energy stocks appear to be independent of both equity market volatility and jump-tail risk.

[Insert Table 7 here]

2.6 Robustness Checks

In this section, we discuss the robustness of our main findings to the methods used in calculating the variables of interest. We discuss if the TRP provides additional information to VRP in 2.6.1. We then calculate a new "unbiased" conditional measure in 2.6.2 and test the predictive regressions based on this measure in 2.6.3.

2.6.1 Regression Using Residuals from TRP and VRP

In Table 4, the correlation between the conditional VRP and TRP is -0.62, which may raise concern if the TRP provides additional information content to the VRP. To further test the information content between the VRP and the TRP, we use the residuals from the regression of the TRP on the VRP (and regression of the VRP on the TRP) to repeat the results in the Conditional Panel of Table 5.

We run the two-stage regression as follows. First, regress the conditional VRP on the conditional TRP,

$$VRP = a^1 + b^1 \times TRP + \varepsilon^1 \tag{31}$$

We find that the coefficient b^1 is statistically significant. We use the residuals, ε^1, to repeat the regressions in the Conditional Panel of Table 5. We report the results in Panel A of Table 8 which indicate that the part of the VRP that is not linearly correlated with the TRP significantly predicts future index returns.

We then try a reverse version of the above two-stage regression. We first run a regression of the TRP on the VRP,

$$TRP = a^2 + b^2 \times VRP + \varepsilon^2 \tag{32}$$

The coefficient b^2 is statistically significant. We then use the residuals, ε^2, to repeat the regressions in the Conditional Panel of Table 5 and report the results in Panel A of Table 8. The results indicate that the part of the TRP that is not linearly correlated with the VRP significantly predicts future index returns.

Overall, by combining the two sets of results, we observe that both the VRP and the TRP possess a unique component that is not linearly correlated with the other. This unique component significantly predicts future index returns.

2.6.2 An "Unbiased" Conditional Measure

The conditional measures from Bekaert and Hoerova that are implemented in this chapter are estimated over the entire period. To verify that the predictive results from the conditional measures are not driven by this estimation, we calculate a new conditional measure using subsample estimation and is not biased estimation, and it is not biased from future data.

We split our sample into two equal subsamples in time series. The first subsample is from January 31, 1990, to November 14, 2002. The second half of the sample is from November 15, 2002, to September 10, 2014. We estimate the conditional regression parameters using the first subsample of data and then apply them to the second subsample to calculate the "unbiased" conditional variables for November 15, 2002, to September 10, 2014. To compare this conditional variable with the unconditional one, we repeat the tests in Table 2 and Table 3 using the conditional measure and a new unconditional measure for the second half of the sample. We report the results in Table 9 and Table 10, respectively.

The coefficient estimates and the decomposed R^2 using the new "unbiased" conditional measure for the second subsample in Table 9 are consistent to what has been reported in Table 2, which uses the conditional measure estimated over the entire sample. The RV is still the dominating component and the systematic attribution of the TRP accounts for more than that of the VRP. These results are also comparable to the results in Table 10, which uses the same subsample but the unconditional measure. Note that the decomposed R^2s for the TRP in Table 9 and Table 10 are much higher than those reported in Table 2 and Table 3. (The decomposed R^2 for the entire sample in Table 2 is 26.64% and that for the second sample in Table 9 is 41.86%). This observation is mainly driven by the differences in sample periods, in which the second subsample is a much more volatile period.

2.6.3 Predictive Power using the "Unbiased" Conditional Measure

To test the actual predictive power of the proposed decomposition, we first apply the conditional measure calculated in 6.2 to investigate the out-of-sample performance of the "unbiased" conditional measure. We compare the in-sample R^2 and the out-of-sample R^2 for the second half of the sample and report the results in Table 11. The in-sample R^2 is obtained using the conditional measure which is calculated from the parameters estimated for the same (second) subsample, while the out-of-sample R^2 is obtained using the "unbiased" conditional measure which is calculated from the parameters estimated for the first subsample. From Table 11, we observe that the TRP has the highest R^2 out-of-sample and outperforms other decomposed components of the VIX index.

We then repeat the predictive regressions in the Conditional Panel of Tables 5 through 7. Due to the space constraint, we report only regression results for monthly returns in Table 12-14, respectively. From Table 12, it is evident that the TRP calculated using the new conditional measure for the second subsample has significant predictive power for the monthly index returns. Overall, the new conditional measures yield similar results to the conditional variables estimated using the entire sample in terms of index return prediction, style portfolio return prediction and industry portfolio return prediction.

2.7 Conclusion

Based on our notion of the PV, the VIX index is composed of four fundamentally different elements: the realized variance (RV), the underlying (unbiased) variance risk premium (VRP), the

realized tail (RT), and the tail risk premium (TRP). RV measures the current (normal) volatility of returns; VRP quantifies the risk premium of anticipated (normal) market volatility; RT captures the present (abnormal) jumps of market returns; and TRP compensates the potentially (unusual) large and asymmetric market price movements, respectively. In short, the VIX index consists of investors' required compensations to two separately expected market risks: the volatility risk (normal price fluctuations from economic uncertainty) and jump-tail risk (abnormally large and asymmetric price movement). Empirically, although the daily variation of the VIX index is largely attributed to the contemporarily realized volatility, premiums of both the volatility and tail risk play a major role in formatting the VIX.

Our VIX decomposition also highlights the bias of the conventional measure of variance risk premium (VRP^c; the squared VIX minus RV) toward the actual premium of its underlying variance risk (VRP) in that VRP^c is actually the sum of VRP, RT, and TRP. We investigate if the high predictive power of the popular VRP^c previously reported in the literature can be actually from the predictability of the conditional RT and that of the TRP by investigating empirically the joint predictive ability of the decomposed VIX components for future returns on the S&P 500, style, and sector portfolios. To ensure the accuracy of risk estimation, we employ both the Bollerslev, Tauchen and Zhou (2009) unconditional and the Bekaert and Hoerova (2014) conditional approaches for calculating RV, VRP, RT, and TRP, respectively.

Statistically, our analysis, consistent with previous researchers' findings, also shows that the RV has no predictive power of future market returns. However, the RT, on the other hand, has a significant influence on market return prediction, particularly, for relatively short time horizons. In addition, both the unbiased premiums of variance risk and tail risk play an important role in predicting future returns on the market, and style as well as different sector portfolios. Specifically, the predictability of the zero-cost small-minus-big (size) portfolios appears to be driven by the TRP. The VRP has a significant impact on the return prediction of the HML B/M (growth/value) portfolios. Nevertheless, the influence of the four VIX decomposed components on return prediction of the winners-minus-losers (momentum) portfolios is quite weak. Finally, although none of our VIX decomposed measures has long-term predictive power for forecasting (annual) returns on industrial portfolios, the conditional RT and TRP, particularly, appear to be strong return predictors for monthly and quarterly returns on almost all sector portfolios. Interestingly, the insignificance of all of our predictors for predicting returns on the energy portfolio demonstrates the unique pricing behavior of energy stocks from other sectors.

Perceptibly, despite the fact that the physical measure of the RT is numerically unnoticeable, our empirical evidence reveals that its impact on future returns is statistically significant and should not be ignored. Notably, the increase in statistical significance from the market indexes to less diversified industrial portfolios indicates that the influence of tail risk on individual stocks could be nontrivial. Therefore, mapping the cross-sectional dynamics of time-varying tail variations in individual asset prices so that the asset pricing model can generate sufficient compensations for investors' fear of potential disasters becomes a consequential line of further research.

References

Almeida, C., Vicente, J. and Guillen, O., 2013. Nonparametric tail risk and stock returns: predictability and risk premia. Unpublished working paper. Getulio Vargas Foundation, Rio de Janeiro.

Andreou, E. and Ghysels, E., 2013. What drives the VIX and the Volatility Risk Premium. unpublished University of Cyprus and University of North Carolina working paper (September).

Andersen, T.G. and Bollerslev, T., 1998. Deutsche mark–dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependencies. the Journal of Finance, 53(1), pp.219-265.

Andersen, T.G., Fusari, N. and Todorov, V., 2015. Parametric inference and dynamic state recovery from option panels. Econometrica, 83(3), pp.1081-1145.

Andersen, T.G., Bollerslev, T., Diebold, F.X. and Labys, P., 2001. The distribution of realized exchange rate volatility. Journal of the American statistical association, 96(453), pp.42-55.

Bakshi, G., Kapadia, N. and Madan, D., 2003. Stock return characteristics, skew laws, and the differential pricing of individual equity options. The Review of Financial Studies, 16(1), pp.101-143.

Bali, T.G. and Zhou, H., 2016. Risk, uncertainty, and expected returns. Journal of Financial and Quantitative Analysis, 51(3), pp.707-735.

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

Bekaert, G. and Engstrom, E., 2010. Inflation and the stock market: Understanding the "Fed Model". Journal of Monetary Economics, 57(3), pp.278-294.

Bekaert, G. and Hoerova, M., 2014. The VIX, the variance premium and stock market volatility. Journal of Econometrics, 183(2), pp.181-192.

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

Bollerslev, T., Marrone, J., Xu, L. and Zhou, H., 2014. Stock return predictability and variance risk premia: statistical inference and international evidence. Journal of Financial and Quantitative Analysis, 49(3), pp.633-661.

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

Bollerslev, T., Gibson, M. and Zhou, H., 2011. Dynamic estimation of volatility risk premia and investor risk aversion from option-implied and realized volatilities. Journal of Econometrics, 160(1), pp.235-245.

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

Bollerslev, T. and Todorov, V., 2014. Time-varying jump tails. Journal of Econometrics, 183(2), pp.168-180.

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

Bondarenko, O., 2014. Variance trading and market price of variance risk. Journal of Econometrics, 180(1), pp.81-97.

Britten‐Jones, M. and Neuberger, A., 2000. Option prices, implied price processes, and stochastic volatility. The Journal of Finance, 55(2), pp.839-866.

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), pp.205-251.

Camponovo, L., Scaillet, O. and Trojani, F., 2014. Predictability hidden by anomalous observations. Swiss Finance Institute Research Paper, (13-05).

Carr, P. and Madan, D., 1999. Option valuation using the fast Fourier transform. Journal of computational finance, 2(4), pp.61-73.

Carr, P., Lee, R. and Wu, L., 2012. Variance swaps on time-changed Lévy processes. Finance and Stochastics, 16(2), pp.335-355.

Carr, P. and Wu, L., 2008. Variance risk premiums. The Review of Financial Studies, 22(3), pp.1311-1341.

Chow, V., Jiang, W. and Li, V., 2018. Does VIX Truly Measure Return Volatility?. Available at SSRN 2489345.

Corsi, F., Pirino, D. and Reno, R., 2010. Threshold bipower variation and the impact of jumps on volatility forecasting. Journal of Econometrics, 159(2), pp.276-288.

Cremers, M., Halling, M. and Weinbaum, D., 2015. Aggregate jump and volatility risk in the cross‑section of stock returns. The Journal of Finance, 70(2), pp.577-614.

Demeterfi, K., Derman, E., Kamal, M. and Zou, J., 1999. More than you ever wanted to know about volatility swaps. Goldman Sachs quantitative strategies research notes, 41, pp.1-56.

Demeterfi, K., Derman, E., Kamal, M. and Zou, J., 1999. A guide to volatility and variance swaps. The Journal of Derivatives, 6(4), pp.9-32.

Du, J. and Kapadia, N., 2012. Tail and volatility indices from option prices. Working paper, University of Massachusetts, Amhurst.

Drechsler, I. and Yaron, A., 2010. What's vol got to do with it. The Review of Financial Studies, 24(1), pp.1-45.

Eraker, B. and Wang, J., 2015. A non-linear dynamic model of the variance risk premium. Journal of Econometrics, 187(2), pp.547-556.

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

Han, B. and Zhou, Y., 2012. Variance risk premium and cross-section of stock returns. unpublished paper, University of Texas at Austin.

Jiang, G.J. and Oomen, R.C., 2008. Testing for jumps when asset prices are observed with noise–a "swap variance" approach. Journal of Econometrics, 144(2), pp.352-370.

Klein, R. and Chow, V., 2013. Orthogonalized Equity Risk Premia and Systematic Risk Decomposition. Quarterly Review of Economics and Finance 53(2), 175-187.

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

Li, J. and Zinna, G., 2018. The variance risk premium: Components, term structures, and stock return predictability. Journal of Business & Economic Statistics, 36(3), pp.411-425.

Liu, J., Pan, J. and Wang, T., 2004. An equilibrium model of rare-event premia and its implication for option smirks. The Review of Financial Studies, 18(1), pp.131-164.

Longstaff, F.A. and Piazzesi, M., 2004. Corporate earnings and the equity premium. Journal of Financial Economics, 74(3), pp.401-421.

Merton, R.C., 1976. Option pricing when underlying stock returns are discontinuous. Journal of financial economics, 3(1-2), pp.125-144.

Naik, V. and Lee, M., 1990. General equilibrium pricing of options on the market portfolio with discontinuous returns. The Review of Financial Studies, 3(4), pp.493-521.

Neuberger, A., 1994. The log contract. Journal of portfolio management, 20(2), p.74.

Newey, W.K. and West, K.D., 1986. A simple, positive semi-definite, heteroskedasticity and autocorrelationconsistent covariance matrix.

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

Smith, R.J., 2005. Automatic positive semidefinite HAC covariance matrix and GMM estimation. Econometric Theory, 21(1), pp.158-170.

Tankov, P., 2003. Financial modelling with jump processes. Chapman and Hall/CRC.

Todorov, V. and Tauchen, G., 2011. Volatility jumps. Journal of Business & Economic Statistics, 29(3), pp.356-371.

Vilkov, G. and Xiao, Y., 2013. Option-implied information and predictability of extreme returns. Working paper (2013), Frankfurt School of Finance & Management.

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

Tables and Plots

Table 1. Summary Statistics
This table reports descriptive statistics for our realized volatility (RV), realized tail (RT) as well as both the conditional and unconditional (annualized) risk premiums of the variance (VRP) and these of the jump-tail (TRP) with respect to no overlapping 1-day, 5-day (weekly) and 22-day (monthly) time horizon, respectively. The sample of 5-minute returns of S&P 500 index extends from January 31, 1990 to September 10, 2014. The conditional measures are based on the forecasting models shown in equations (26), (27) and (28), accordingly. All our measures are on the daily overlapping basis with 5979 observations in total. RT is in annualized basis point, and all other variables are in annualized percentage. In addition, all numbers are scaled up by a factor of 100.

	RV	RT	Unconditional		Conditional	
			VRP*	TRP	VRP*	TRP
Panel A. Daily Measure						
Mean	3.803	-0.172	1.013	-0.179	2.051	-0.156
Std. Dev.	8.716	6.523	6.241	0.740	3.971	0.744
Skewness	11.912	-49.538	-15.959	-5.038	5.014	-5.051
Max	282.739	50.690	25.947	3.606	59.072	3.289
Min	0.099	-429.90	-226.595	-9.521	-5.869	-9.620
t-value	33.611	-2.035	12.505	-18.673	39.791	-16.160
Panel B. Weekly Measure						
Mean	3.943	-0.312	0.901	-0.183	2.054	-0.161
Std. Dev.	7.479	4.770	4.752	0.701	3.784	0.708
Skewness	6.910	-15.493	-8.144	-4.913	4.480	-4.846
Max	109.415	26.492	14.991	1.328	43.839	1.325
Min	0.234	-100.389	-69.897	-7.625	-3.886	-7.545
t-value	18.723	-2.320	6.736	-9.264	19.274	-8.100
Panel C. Monthly Measure						
Mean	3.879	-0.206	0.970	-0.177	2.044	-0.154
Std. Dev.	6.154	1.548	2.840	0.637	3.606	0.641
Skewness	6.616	-7.694	-5.208	-4.600	4.076	-4.515
Max	74.069	4.630	9.610	0.648	31.758	0.490
Min	0.389	-19.569	-30.766	-5.611	-2.166	-5.555
t-value	10.790	-2.276	5.848	-4.752	9.705	-4.122
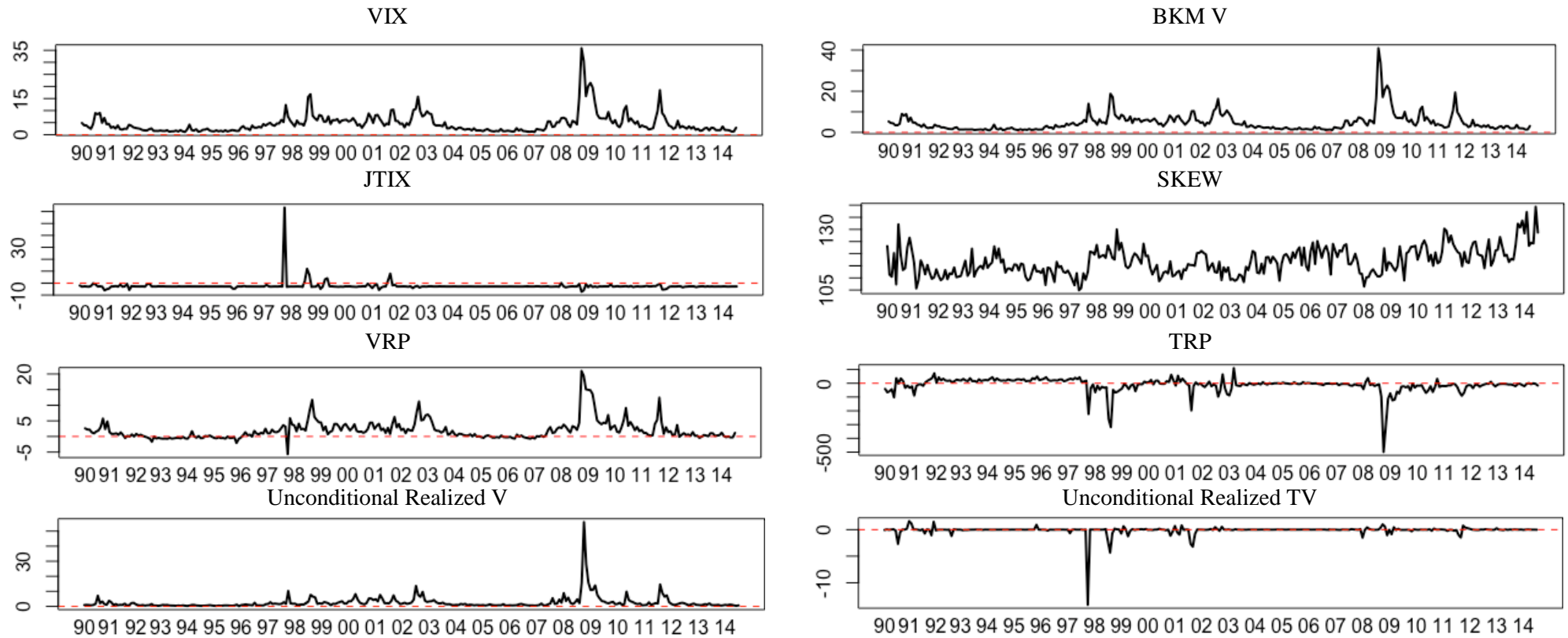
Plot 1. Variance and Tail Risk Premiums
The monthly conditional estimates of VRP and TRP are based on 5-minute sample returns of the S&P 500 index and extends from January 31, 1990 to September 10, 2014 for a total of 5979 trading days. VRP = V^{BKM} – RV, TRP = (VIX^2– RVIX) – VRP, and the conditional measures are based on the forecasting models shown in equations (26), (27), and (28), accordingly. Both VRP and TRP are reported in annualized percentage and scaled up by a factor of 100. The VIX index and the SKEW index are obtained from the CBOE website. The unconditional measures are calculated using equations (21) and (22).

Table 2. Goodness-of-Fit of the VIX Decomposition (Unconditional Estimates)

This table reports the *decomposed R^2* of the orthogonalized VIX components, $RV^{\perp}$, $VRP^{\perp}$, $RT^{\perp}$, and $TRP^{\perp}$, respectively. In addition to the overall sample analysis, we also examine the decomposed goodness of fit from three different sub-samples. The sub-samples are classified by three different levels of the VIX. These include (1) *Nervous Market Condition*: VIX ≥ 23.32 (75 percentile), (2) *Normal Market Condition*: 14.17 ≤ VIX < 23.32, and (3) *Calm Market Condition*: VIX < 14.17 (25 percentile), correspondingly. We employ Klein and Chow (2013) to orthogonalize the VIX's decomposed variables and further calculate their *decomposed R^2*. Also, to avoid bias results due to measurements at different scales, all variables are standardized for the analyses. The overall sample period of daily data ranges from January 1993 to September 2014.

A. Overall Sample	$RV^{\perp}$	$VRP^{\perp}$	$RT^{\perp}$	$TRP^{\perp}$
Coefficient	0.82	0.26	-0.01	-0.53
Decomposed R^2(%)	66.83	6.52	0.01	26.64

B. Subsample for VIX ≥ 23.32 (*Nervous Market Condition*)				
Coefficient	0.68	0.16	-0.06	-0.49
Decomposed R^2(%)	63.70	3.84	0.12	32.35

C. Subsample for 14.17 ≤ VIX < 23.32 (*Normal Market Condition*)				
Coefficient	0.34	0.16	-0.00	-0.05
Decomposed R^2(%)	69.93	28.37	0.04	1.67

D. Subsample for VIX < 14.17 (*Calm Market Condition*)				
Coefficient	0.55	0.29	-0.02	0.02
Decomposed R^2(%)	65.81	33.32	0.01	0.87

Table 3. Goodness-of-Fit of the VIX Decomposition (Conditional Estimates)
This table reports the *decomposed R^2* of the four orthogonalized VIX components, $RV^\perp$, $VRP^\perp$, $RT^\perp$, and $TRP^\perp$, respectively. In addition to the overall sample analysis, we also examine the decomposed goodness of fit from three different sub-samples. The sub-samples are classified by three different levels of the VIX. These include (1) *Nervous Market Condition*: VIX $\geq$ 23.32 (75 percentile), (2) *Normal Market Condition*: 14.17 $\leq$ VIX < 23.32, and (3) *Calm Market Condition*: VIX < 14.17 (25 percentile), correspondingly. We employ Klein and Chow (2013) to orthogonalize the VIX's decomposed variables and further calculate their *decomposed R^2*. In addition, to avoid bias results due to measurements at different scales, all variables are standardized for the analyses. The conditional measures are based on the forecasting models shown in equations (26), (27) and (28), accordingly. The overall sample period of daily data ranges from January 1993 to September 2014.

A. Overall Sample	$RV^\perp$	$VRP^\perp$	$RT^\perp$	$TRP^\perp$
Coefficient	0.55	0.76	0.03	-0.34
Decomposed R^2 (%)	29.80	58.37	0.08	11.74

B. Subsample for VIX $\geq$ 23.32 (*Nervous Market Condition*)				
Coefficient	0.46	0.71	-0.01	-0.34
Decomposed R^2 (%)	28.82	55.24	0.01	15.93

C. Subsample for 14.17 $\leq$ VIX < 23.32 (*Normal Market Condition*)				
Coefficient	0.34	0.68	-0.00	-0.02
Decomposed R^2 (%)	22.86	76.94	0.00	0.20

D. Subsample for VIX < 14.17 (*Calm Market Condition*)				
Coefficient	0.29	0.62	-0.01	0.04
Decomposed R^2 (%)	24.70	72.94	0.00	2.36

Table 4. Correlation Matrices

This table depicts pairwise correlations for monthly non-overlapping measures of variation (i.e., VIX, RV, V^{BKM}, and RT), those of risk premiums (i.e., VRP and, TRP) as well as those of our control variables including 3MTB (3-month T-bill minus CPI inflation), Log(DY), the log-dividend yield, CS (the spread between Moody's BAA and AAA bond yield), and TS (the spread between 10-year and 3-month Treasury yields), respectively. The sample period extends from January 31, 1990 to September 10, 2014.

| | | *Unconditional* | | | | | | | |
		RV	VRP	RT	TRP	3MTB	log(DY)	CS	TS
Unconditional	VIX	0.86	0.35	-0.18	-0.69	-0.17	0.12	0.66	0.09
	RV	1.00	-0.17	-0.10	-0.71	-0.15	0.16	0.59	0.08
	VRP		1.00	-0.18	-0.13	-0.06	-0.04	0.20	0.02
	RT			1.00	0.24	-0.06	0.04	0.09	0.08
	TRP				1.00	0.18	-0.23	-0.49	-0.04
	3MTB					1.00	-0.66	-0.43	-0.72
	log(DY)						1.00	0.31	0.08
	CS							1.00	0.31

| | | *Conditional* | | | | | | | |
		RV	VRP	RT	TRP	3MTB	log(DY)	CS	TS
Conditional	VIX	0.80	0.93	0.15	-0.69	-0.17	0.12	0.66	0.09
	RV	1.00	0.53	0.50	-0.71	-0.07	0.08	0.43	0.03
	VRP		1.00	-0.09	-0.62	-0.21	0.14	0.68	0.11
	RT			1.00	-0.28	0.08	-0.04	-0.10	-0.08
	TRP				1.00	0.17	-0.22	-0.46	-0.03
	3MTB					1.00	-0.66	-0.43	-0.72
	log(DY)						1.00	0.31	0.08
	CS							1.00	0.31

Table 5. S&P 500 Return Predictability Regressions

This table reports the estimated regression coefficients and adjusted R^2's from return predictability regressions for monthly, quarterly, and annual excess returns on the S&P 500 market portfolio, respectively. RV is the realized variance; VRP is the unbiased variance risk premium; RT is the realized tail; and TRP is the tail risk premium. The term 3MTB is the 3-month T-bill minus CPI inflation; Log(DY) is the log-dividend yield; CS is the spread between Moody's BAA and AAA bond yield; and TS is the term spread between 10-year and 3-month Treasury yields. The sample extends from January 31, 1990 to September 10, 2014. Newey-West t-statistics are reported in parentheses. *Adj. R^2* is the adjusted coefficient of determination. RT is in annualized basis point, and all other variables are measured by annualized percentage.

Panel A. Monthly Return Prediction

Unconditional

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
VIX	-0.281					0.523						
	(-0.221)					(0.408)						
RV		-1.788				1.668						
		(-2.311)				(1.260)						
VRP			4.598			6.597						
			(3.204)			(4.756)						
RT				0.110		0.543						
				(0.036)		(0.132)						
TRP					16.164	26.064						
					(2.448)	(1.961)						

Conditional

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
RV							-2.127					-2.207
							(-0.835)					(-0.660)
VRP								-0.052				5.168
								(-0.033)				(3.172)
RT									0.484			2.376
									(2.056)			(2.651)
TRP											15.155	29.260
											(2.470)	(2.513)
3MTB										-1.784	1.260	0.769
										(-0.358)	(0.328)	(0.195)
Log(DY)										-5.023	15.831	12.016
										(-0.224)	(0.843)	(0.656)
CS										-13.477	-12.808	-12.593
										(-0.911)	(-0.923)	(-0.806)
TS										-0.970	2.674	2.126
										(-0.140)	(0.503)	(0.388)
Constant	5.618	8.850	-5.819	4.341	6.600	10.243	4.430	5.489	6.113	36.084	-52.486	-19.072
	(1.126)	(3.390)	(-1.543)	(1.342)	(2.409)	(1.599)	(1.469)	(1.692)	(2.174)	(0.392)	(-0.680)	(-0.253)
Adj. R^2(%)	-0.290	2.022	4.034	-0.343	2.573	0.405	-0.343	0.183	2.341	-1.046	6.766	4.602

Panel B. Quarterly Return Prediction

Unconditional

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
VIX	0.086									0.720		
	(0.106)									(0.732)		
RV		-1.149								0.480		
		(-2.588)								(0.630)		
VRP			4.096							4.765		
			(6.855)							(6.228)		
RT				-2.288						-0.910		
				(-1.940)						(-0.923)		
TRP					5.501					8.678		
					(0.898)					(1.707)		

Conditional

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
RV						-0.937						-2.673
						(-0.457)						(-1.668)
VRP							0.401					3.553
							(0.288)					(3.313)
RT								0.562				1.779
								(1.631)				(3.699)
TRP									4.851			10.824
									(0.795)			(2.161)
3MTB										-3.582	-1.043	-1.612
										(-0.929)	(-0.346)	(-0.480)
Log(DY)										-12.740	1.878	-2.160
										(-0.637)	(0.109)	(-0.123)
CS										-11.194	-9.818	-10.985
										(-0.814)	(-0.799)	(-0.806)
TS										-3.787	-0.461	-1.139
										(-0.720)	(-0.107)	(-0.241)
Constant	3.878	7.198	-4.776	4.011	5.049	6.889	3.488	5.621	4.847	64.995	1.279	33.498
	(1.114)	(4.059)	(-1.516)	(1.489)	(1.759)	(1.330)	(1.220)	(1.841)	(1.671)	(0.824)	(0.020)	(0.500)
Adj. R^2(%)	-0.332	2.356	9.265	0.192	0.589	0.056	-0.177	0.253	0.416	0.084	10.200	4.820

Panel C. Annual Return Prediction

Unconditional

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
VIX	0.460									0.353		
	(1.327)									(0.820)		
RV		0.057										-0.112
		(0.193)										(-0.177)
VRP			1.452									1.144
			(3.099)									(2.079)
RT				-0.125								0.764
				(-0.144)								(0.735)
TRP					-2.897							-3.594
					(-1.369)							(-0.755)

Conditional

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
RV						0.416						-0.767
						(0.904)						(-0.994)
VRP							0.722					0.629
							(1.524)					(0.798)
RT								0.074				0.225
								(0.585)				(0.827)
TRP									-2.813			-2.519
									(-1.380)			(-0.591)
3MTB										-3.007	-2.281	-2.581
										(-0.985)	(-0.788)	(-0.813)
Log(DY)										-20.012	-16.750	-18.599
										(-0.872)	(-0.748)	(-0.793)
CS										2.325	1.466	1.718
										(0.385)	(0.267)	(0.297)
TS										-3.173	-2.039	-2.467
										(-0.682)	(-0.467)	(-0.521)
Constant	2.144	4.153	1.052	4.287	3.892	3.129	2.844	4.479	3.967	72.049	58.536	68.997
	(0.668)	(1.421)	(0.319)	(2.531)	(1.545)	(0.945)	(0.937)	(1.460)	(1.690)	(0.917)	(0.762)	(0.861)
Adj. R^2(%)	0.942	-0.335	3.614	-0.352	0.500	-0.097	1.445	0.3224	0.490	2.976	4.612	2.685

Unconditional Measures
t-Statistics

Unconditional Measures
R^2

Conditional Measures
t-Statistics

Conditional Measures
R^2
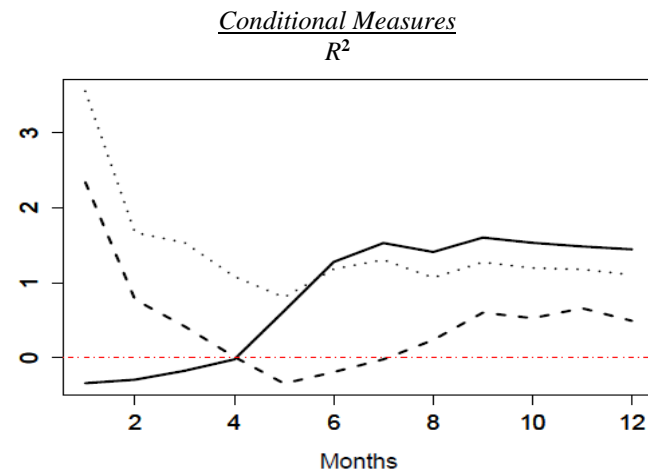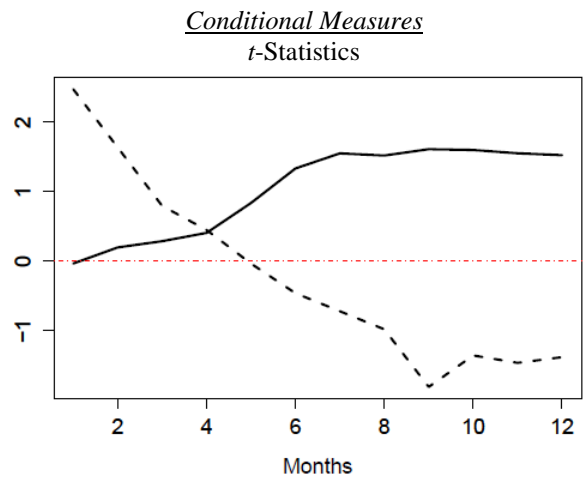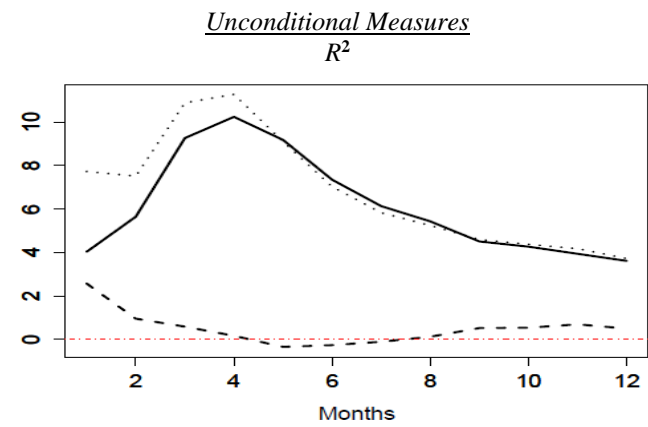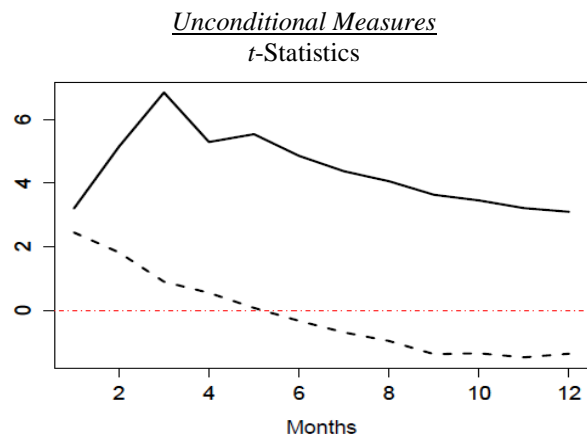
Plot 2. S&P 500 Return Predictability Regressions
The left panels show the Newey-West *t*-statistics from the simple return predictability regressions for the S&P 500 portfolio based on the unbiased variance risk premiums, VRP (solid line), and the tail risk premium, TRP (dashed line), respectively. The right panels depict the corresponding R^2s along with the R^2s from multiple regressions including both VRP and TRP (dotted line). The results shown on the top (bottom) panels are based on unconditional (conditional) measures of risk premiums.

Table 6. Style Portfolio Return Predictability Regressions

This table reports the predictability regression results from excess returns on Size (20% smallest and biggest firms), Book-to-Market (20% highest and lowest B/M ratios), and Momentum (20% top and bottom performance), along with the corresponding zero-cost portfolios. All other variables are described in Table 3.

Panel A. Monthly Return Prediction

| | Constant | Unconditional | | | | | 3MTB | Log(DY) | CS | TS | Adj. R^2(%) |
		RV	VRP	TRP	RT						
Small	3.804	0.110	0.568	3.133	0.504		-0.326	-1.101	0.265	-0.059	6.238
	(0.429)	(0.805)	(3.236)	(2.207)	(1.049)		(-0.699)	(-0.491)	(0.201)	(-0.098)	
Big	-4.792	0.152	0.599	1.957	0.003		0.147	1.464	-1.187	0.273	6.797
	(-0.776)	(1.331)	(4.930)	(1.737)	(0.011)		(0.463)	(0.985)	(-1.029)	(0.604)	
SMB	8.596	-0.042	-0.032	1.176	0.501		-0.474	-2.565	1.453	-0.332	3.707
	(1.479)	(-0.390)	(-0.215)	(1.774)	(2.017)		(-1.369)	(-1.810)	(1.776)	(-0.669)	
High	0.098	0.085	0.447	2.300	0.195		-0.124	0.259	-0.586	0.028	4.213
	(0.014)	(0.763)	(3.029)	(1.986)	(0.476)		(-0.313)	(0.148)	(-0.391)	(0.057)	
Low	-8.822	0.140	0.701	1.821	0.020		0.376	2.289	-0.761	0.527	7.276
	(-1.290)	(0.995)	(4.935)	(1.479)	(0.074)		(0.990)	(1.378)	(-0.742)	(0.937)	
HML	8.920	-0.054	-0.255	0.479	0.175		-0.499	-2.030	0.174	-0.499	2.693
	(2.071)	(-0.658)	(-2.164)	(0.769)	(0.742)		(-1.894)	(-2.025)	(0.213)	(-1.094)	
Winners	-2.819	0.021	0.557	1.124	0.270		0.132	1.048	-1.600	0.472	4.932
	(-0.452)	(0.202)	(4.338)	(1.100)	(0.920)		(0.438)	(0.698)	(-1.684)	(0.909)	
Losers	-2.436	0.318	1.080	4.450	0.141		-0.265	0.114	0.779	-0.069	8.492
	(-0.191)	(1.577)	(4.622)	(2.419)	(0.317)		(-0.382)	(0.036)	(0.309)	(-0.079)	
WML	-0.384	-0.297	-0.523	-3.326	0.129		0.398	0.934	-2.378	0.540	5.392
	(-0.040)	(-1.836)	(-2.566)	(-2.507)	(0.460)		(0.773)	(0.385)	(-1.261)	(0.777)	

| | | Conditional | | | | | | | | | |
		RV	VRP	TRP	RT						
Small	6.534	-0.205	0.406	3.561	0.144		-0.379	-1.465	0.428	-0.123	4.461
	(0.755)	(-0.662)	(2.436)	(2.920)	(1.587)		(-0.831)	(-0.660)	(0.286)	(-0.213)	
Big	-1.564	-0.161	0.455	2.251	0.203		0.087	1.053	-1.186	0.200	3.906
	(-0.257)	(-0.555)	(3.513)	(2.330)	(2.909)		(0.264)	(0.730)	(-0.916)	(0.431)	
SMB	8.098	-0.044	-0.049	1.309	-0.059		-0.465	-2.518	1.614	-0.324	3.142
	(1.413)	(-0.198)	(-0.338)	(1.978)	(-0.883)		(-1.298)	(-1.751)	(1.836)	(-0.622)	
High	2.736	-0.101	0.293	2.628	0.127		-0.196	-0.152	-0.493	-0.066	2.502
	(0.381)	(-0.249)	(1.850)	(2.781)	(1.105)		(-0.474)	(-0.089)	(-0.282)	(-0.127)	
Low	-5.075	-0.331	0.560	2.130	0.254		0.331	1.893	-0.828	0.481	4.073
	(-0.779)	(-0.895)	(3.529)	(1.754)	(3.001)		(0.913)	(1.244)	(-0.677)	(0.912)	
HML	7.810	0.231	-0.268	0.498	-0.128		-0.526	-2.046	0.335	-0.547	2.716
	(1.879)	(1.200)	(-2.964)	(0.830)	(-1.897)		(-1.915)	(-2.046)	(0.500)	(-1.202)	
Winners	1.053	-0.267	0.344	1.566	0.145		0.031	0.457	-1.558	0.342	1.266
	(0.153)	(-0.885)	(2.437)	(1.617)	(2.004)		(0.080)	(0.288)	(-1.266)	(0.587)	
Losers	2.275	-0.386	0.913	4.873	0.379		-0.298	-0.308	0.739	-0.090	6.781
	(0.197)	(-0.585)	(2.949)	(2.941)	(1.964)		(-0.452)	(-0.115)	(0.217)	(-0.101)	

		Unconditional								
WML	-1.222	0.118	-0.569	-3.307	-0.235	0.329	0.765	-2.297	0.431	5.860
	(-0.126)	(0.374)	(-2.967)	(-2.513)	(-3.070)	(0.664)	(0.293)	(-1.230)	(0.652)	

Panel B. Quarterly Return Prediction

			Unconditional							
	Constant	RV	VRP	TRP	RT	3MTB	Log(DY)	CS	TS	Adj. R^2(%)
Small	7.424	0.111	0.458	2.142	0.041	-0.501	-2.061	0.527	-0.396	8.769
	(0.881)	(1.342)	(4.690)	(3.478)	(0.229)	(-1.063)	(-0.967)	(0.408)	(-0.654)	
Big	-0.507	0.041	0.412	0.471	-0.101	-0.026	0.350	-0.855	0.029	10.240
	(-0.096)	(0.611)	(5.910)	(1.084)	(-1.164)	(-0.104)	(0.252)	(-0.868)	(0.078)	
SMB	7.931	0.071	0.046	1.671	0.141	-0.475	-2.410	1.383	-0.425	9.358
	(1.569)	(0.901)	(0.509)	(3.049)	(1.168)	(-1.577)	(-1.946)	(2.117)	(-0.945)	
High	3.051	0.034	0.320	1.617	-0.200	-0.267	-0.610	0.122	-0.239	7.480
	(0.507)	(0.468)	(3.833)	(3.296)	(-1.546)	(-0.830)	(-0.388)	(0.086)	(-0.579)	
Low	-3.182	0.057	0.466	0.148	-0.054	0.142	0.900	-0.716	0.217	9.634
	(-0.564)	(0.776)	(5.800)	(0.303)	(-0.527)	(0.499)	(0.626)	(-0.757)	(0.509)	
HML	6.233	-0.023	-0.146	1.469	-0.146	-0.408	-1.511	0.839	-0.456	11.250
	(1.710)	(-0.326)	(-1.932)	(2.866)	(-1.505)	(-1.912)	(-1.721)	(1.286)	(-1.136)	
Winners	1.331	0.004	0.365	0.337	0.019	-0.049	-0.048	-1.325	0.236	6.655
	(0.246)	(0.059)	(4.030)	(0.646)	(0.186)	(-0.170)	(-0.035)	(-1.482)	(0.533)	
Losers	6.077	0.118	0.713	1.757	-0.178	-0.613	-2.145	1.750	-0.689	13.910
	(0.545)	(1.041)	(4.094)	(2.082)	(-0.677)	(-1.127)	(-0.733)	(0.823)	(-0.981)	
WML	-4.746	-0.114	-0.348	-1.420	0.197	0.564	2.097	-3.075	0.924	16.450
	(-0.591)	(-1.078)	(-1.746)	(-1.664)	(0.884)	(1.370)	(1.040)	(-2.206)	(1.622)	

			Conditional							
		RV	VRP	TRP	RT					
Small	9.976	-0.129	0.352	2.345	0.136	-0.551	-2.400	0.459	-0.457	6.659
	(1.193)	(-0.734)	(3.403)	(4.404)	(2.967)	(-1.204)	(-1.090)	(0.338)	(-0.784)	
Big	2.268	-0.242	0.313	0.643	0.155	-0.071	0.013	-0.976	-0.024	4.543
	(0.414)	(-1.759)	(3.304)	(1.398)	(3.454)	(-0.257)	(0.009)	(-0.890)	(-0.060)	
SMB	7.709	0.113	0.039	1.703	-0.018	-0.479	-2.413	1.435	-0.433	9.289
	(2.200)	(0.798)	(0.458)	(3.835)	(-0.479)	(-2.280)	(-2.757)	(2.710)	(-1.346)	
High	5.899	-0.051	0.180	1.809	0.123	-0.361	-1.106	0.071	-0.365	4.082
	(1.022)	(-0.262)	(1.800)	(3.084)	(2.604)	(-1.119)	(-0.746)	(0.049)	(-0.851)	
Low	-0.382	-0.300	0.378	0.320	0.167	0.112	0.612	-0.860	0.188	4.895
	(-0.066)	(-1.867)	(3.477)	(0.642)	(3.261)	(0.357)	(0.417)	(-0.817)	(0.409)	
HML	6.281	0.249	-0.198	1.489	-0.043	-0.472	-1.718	0.931	-0.553	13.020
	(1.844)	(2.369)	(-2.354)	(2.945)	(-1.477)	(-2.251)	(-2.046)	(1.588)	(-1.436)	
Winners	3.881	-0.271	0.266	0.533	0.132	-0.090	-0.359	-1.401	0.188	3.140
	(0.666)	(-1.925)	(2.565)	(1.196)	(3.226)	(-0.289)	(-0.247)	(-1.321)	(0.396)	
Losers	11.109	-0.200	0.498	2.078	0.208	-0.742	-2.907	1.533	-0.856	8.772
	(1.558)	(-0.578)	(3.083)	(2.121)	(1.983)	(-1.915)	(-1.656)	(0.988)	(-1.652)	
WML	-7.228	-0.072	-0.231	-1.545	-0.076	0.652	2.549	-2.934	1.043	14.560
	(-0.852)	(-0.354)	(-1.314)	(-1.629)	(-0.954)	(1.517)	(1.101)	(-1.868)	(1.738)	

Panel C. Annual Return Prediction

	Constant	RV	VRP	TRP	RT	3MTB	Log(DY)	CS	TS	Adj. R^2(%)
		Unconditional								
Small	12.109	0.071	0.169	0.270	0.229	-0.521	-3.518	0.908	-0.544	21.380
	(1.698)	(1.519)	(2.548)	(0.819)	(2.536)	(-1.828)	(-1.697)	(1.962)	(-1.402)	
Big	4.090	-0.013	0.102	-0.393	0.050	-0.132	-1.127	0.096	-0.097	4.825
	(0.670)	(-0.236)	(2.202)	(-0.976)	(0.638)	(-0.567)	(-0.637)	(0.214)	(-0.266)	
SMB	8.019	0.084	0.068	0.663	0.178	-0.389	-2.390	0.812	-0.447	19.060
	(2.279)	(1.689)	(1.035)	(2.352)	(2.613)	(-2.348)	(-2.655)	(1.608)	(-1.488)	
High	5.118	0.009	0.025	0.201	0.030	-0.166	-1.545	0.942	-0.154	5.108
	(0.750)	(0.191)	(0.421)	(0.594)	(0.291)	(-0.558)	(-0.784)	(1.730)	(-0.378)	
Low	1.883	-0.009	0.149	-0.582	0.084	-0.014	-0.602	0.157	0.012	8.814
	(0.326)	(-0.154)	(3.234)	(-1.438)	(1.012)	(-0.061)	(-0.365)	(0.376)	(0.030)	
HML	3.236	0.018	-0.124	0.783	-0.054	-0.152	-0.943	0.785	-0.166	10.510
	(0.869)	(0.390)	(-2.266)	(2.594)	(-0.840)	(-0.722)	(-1.051)	(2.143)	(-0.378)	
Winners	4.437	-0.035	0.144	-0.503	0.065	-0.087	-1.224	-0.261	0.154	10.170
	(0.657)	(-0.625)	(3.014)	(-1.192)	(0.866)	(-0.339)	(-0.633)	(-0.583)	(0.365)	
Losers	11.586	0.050	0.112	-0.382	0.127	-0.578	-3.814	2.101	-0.597	26.440
	(1.200)	(0.622)	(1.347)	(-0.682)	(1.150)	(-1.570)	(-1.338)	(3.553)	(-1.188)	
WML	-7.149	-0.085	0.032	-0.121	-0.061	0.491	2.590	-2.363	0.751	37.160
	(-1.231)	(-1.510)	(0.443)	(-0.383)	(-0.922)	(1.982)	(1.574)	(-4.887)	(2.353)	
		Conditional								
		RV	VRP	TRP	RT					
Small	13.353	0.125	0.079	0.442	-0.041	-0.591	-3.870	0.998	-0.651	19.820
	(1.769)	(1.873)	(1.042)	(1.621)	(-1.816)	(-1.932)	(-1.770)	(2.085)	(-1.537)	
Big	5.058	-0.074	0.055	-0.300	0.023	-0.159	-1.296	0.113	-0.136	2.579
	(0.794)	(-1.088)	(0.821)	(-0.823)	(1.036)	(-0.621)	(-0.699)	(0.233)	(-0.344)	
SMB	8.295	0.199	0.023	0.741	-0.064	-0.432	-2.574	0.885	-0.515	20.370
	(2.360)	(2.419)	(0.317)	(3.019)	(-2.689)	(-2.580)	(-2.854)	(1.899)	(-1.760)	
High	5.590	0.060	-0.012	0.259	-0.011	-0.199	-1.698	0.994	-0.204	5.213
	(0.787)	(1.097)	(-0.171)	(0.840)	(-0.482)	(-0.632)	(-0.826)	(1.721)	(-0.471)	
Low	3.209	-0.091	0.085	-0.451	0.027	-0.052	-0.835	0.179	-0.042	4.902
	(0.529)	(-1.209)	(1.226)	(-1.219)	(1.210)	(-0.203)	(-0.481)	(0.381)	(-0.098)	
HML	2.381	0.151	-0.096	0.710	-0.039	-0.147	-0.863	0.814	-0.162	8.095
	(0.634)	(2.258)	(-1.773)	(2.451)	(-2.401)	(-0.676)	(-0.927)	(2.104)	(-0.360)	
Winners	6.120	-0.103	0.060	-0.346	0.031	-0.143	-1.545	-0.229	0.074	5.520
	(0.846)	(-1.310)	(0.849)	(-0.956)	(1.278)	(-0.493)	(-0.748)	(-0.458)	(0.158)	
Losers	12.340	0.074	0.057	-0.273	-0.021	-0.619	-4.021	2.173	-0.659	26.040
	(1.231)	(0.763)	(0.589)	(-0.524)	(-0.734)	(-1.573)	(-1.351)	(3.343)	(-1.218)	
WML	-6.220	-0.177	0.002	-0.073	0.052	0.476	2.475	-2.402	0.733	36.110
	(-1.053)	(-1.662)	(0.035)	(-0.260)	(1.950)	(1.835)	(1.473)	(-4.643)	(2.149)	

Unconditional Measures ... *Conditional Measures*

SMB *t*-Statistics   SMB R^2 SMB *t*-Statistics   SMB R^2

HML *t*-Statistics   HML R^2 HML *t*-Statistics   HML R^2

WML *t*-Statistics   WML R^2 WML *t*-Statistics   WML R^2

Plot 3. Sorted Zero-Cost Style Portfolio Return Predictability Regressions
This Plot depicts the Newey-West *t*-statistics and the corresponding R^2s from simple return predictability regressions for the sorted zero-cost style portfolios based on the unbiased variance risk premiums, VRP (solid lines), and the tail risk premium, TRP (dashed lines), respectively. The dotted lines are the *t*-statistics and adjusted R^2s from multiple regressions including both VRP and TRP. SMB, HML, and WML stands for Small Minus Big, High Minus Low, and Winners Minus Losers, accordingly.

61

Table 7. Industry Portfolio Return Predictability Regressions

This table reports the predictability regression results from excess returns on twelve industry portfolios. All other variables are described in Table 3. We employ the data directly from the Fama-French data library.

Panel A. Monthly Return Prediction

| | Constant | Conditional | | | | 3MTB | Log(DY) | CS | TS | Adj. R^2(%) |
		RV	VRP	TRP	RT					
Non-Durables	-2.939	-0.230	0.229	1.706	0.213	0.202	1.099	0.408	0.271	2.619
	(-0.576)	(-1.063)	(1.584)	(1.942)	(3.556)	(0.696)	(0.893)	(0.403)	(0.745)	
Durables	4.102	-0.560	0.751	3.806	0.269	-0.252	-0.655	0.033	0.182	6.054
	(0.395)	(-1.122)	(3.068)	(2.614)	(1.931)	(-0.438)	(-0.262)	(0.015)	(0.251)	
Manufacturing	-0.374	-0.265	0.525	3.126	0.244	0.002	0.704	-0.432	0.154	4.296
	(-0.047)	(-0.584)	(2.466)	(2.460)	(1.925)	(0.004)	(0.374)	(-0.233)	(0.270)	
Energy	4.095	-0.136	0.181	0.700	0.009	-0.286	-0.190	-1.306	-0.344	-1.131
	(0.605)	(-0.461)	(1.258)	(0.822)	(0.137)	(-0.754)	(-0.114)	(-1.533)	(-0.651)	
Chemicals	-2.197	-0.220	0.325	2.197	0.202	0.124	0.986	-0.251	0.417	3.268
	(-0.362)	(-0.601)	(1.791)	(2.466)	(1.986)	(0.346)	(0.697)	(-0.163)	(0.833)	
Equipment	-4.373	-0.364	0.968	3.899	0.350	0.281	2.006	-1.902	0.543	5.055
	(-0.416)	(-0.957)	(4.549)	(2.321)	(4.053)	(0.533)	(0.764)	(-1.246)	(0.615)	
Telecommunications	3.529	0.046	0.404	1.766	0.196	-0.329	-0.080	-1.789	-0.234	1.621
	(0.501)	(0.153)	(2.059)	(1.451)	(2.384)	(-0.895)	(-0.044)	(-1.281)	(-0.405)	
Utilities	-0.912	-0.090	0.026	-0.037	0.108	0.034	0.915	-0.925	0.068	-0.764
	(-0.168)	(-0.482)	(0.223)	(-0.046)	(2.544)	(0.127)	(0.657)	(-0.921)	(0.192)	
Wholesale	-2.329	-0.187	0.509	2.662	0.233	0.108	1.073	-0.439	0.272	4.310
	(-0.325)	(-0.566)	(3.236)	(2.647)	(2.785)	(0.241)	(0.641)	(-0.356)	(0.488)	
Healthcare	-4.089	-0.182	0.264	0.552	0.111	0.325	1.667	-0.831	0.176	-0.599
	(-0.617)	(-0.555)	(1.973)	(0.577)	(1.296)	(0.842)	(1.037)	(-0.742)	(0.316)	
Finance	3.581	-0.096	0.429	3.661	0.183	-0.112	-0.588	-0.463	-0.045	4.746
	(0.433)	(-0.178)	(2.075)	(3.145)	(1.173)	(-0.217)	(-0.309)	(-0.201)	(-0.071)	
Other	-0.372	-0.159	0.507	2.976	0.167	-0.101	0.798	-1.358	0.154	4.674
	(-0.053)	(-0.337)	(2.635)	(2.166)	(1.419)	(-0.259)	(0.491)	(-0.785)	(0.291)	

Panel B. Quarterly Return Prediction

	Constant	Conditional				3MTB	Log(DY)	CS	TS	Adj. R^2(%)
		RV	VRP	TRP	RT					
Non-Durables	-2.202	-0.202	0.204	0.752	0.133	0.184	0.889	-0.021	0.269	2.398
	(-0.502)	(-1.524)	(2.711)	(1.916)	(3.636)	(0.724)	(0.827)	(-0.025)	(0.845)	
Durables	10.763	-0.312	0.438	1.904	0.214	-0.637	-2.576	0.642	-0.442	7.689
	(1.221)	(-1.180)	(2.906)	(2.410)	(3.205)	(-1.341)	(-1.144)	(0.364)	(-0.692)	
Manufacturing	5.166	-0.155	0.312	1.435	0.152	-0.297	-0.798	-0.392	-0.268	3.469
	(0.765)	(-0.846)	(3.263)	(2.785)	(2.907)	(-0.848)	(-0.465)	(-0.280)	(-0.584)	
Energy	3.168	-0.116	0.147	0.489	-0.006	-0.223	-0.037	-1.080	-0.308	0.883
	(0.634)	(-0.643)	(1.574)	(0.998)	(-0.117)	(-0.860)	(-0.029)	(-1.225)	(-0.758)	
Chemicals	3.169	-0.176	0.135	0.906	0.119	-0.151	-0.537	0.226	-0.027	2.202
	(0.664)	(-0.855)	(1.309)	(1.886)	(2.140)	(-0.560)	(-0.465)	(0.229)	(-0.073)	
Equipment	3.566	-0.300	0.597	0.866	0.210	-0.097	-0.180	-1.670	-0.004	5.009
	(0.456)	(-1.297)	(3.607)	(1.035)	(2.691)	(-0.235)	(-0.089)	(-1.152)	(-0.006)	
Telecommunications	2.124	-0.140	0.358	0.612	0.178	-0.207	0.289	-1.690	-0.028	5.359
	(0.322)	(-0.836)	(2.344)	(1.006)	(3.905)	(-0.638)	(0.164)	(-1.361)	(-0.053)	
Utilities	-4.198	-0.029	0.132	0.516	0.069	0.192	1.726	-1.296	0.347	1.978
	(-0.896)	(-0.259)	(1.553)	(1.343)	(2.113)	(0.841)	(1.379)	(-1.349)	(1.269)	
Wholesale	3.803	-0.256	0.370	0.670	0.163	-0.171	-0.485	-0.611	-0.094	5.679
	(0.787)	(-1.499)	(3.272)	(1.261)	(3.107)	(-0.538)	(-0.418)	(-0.730)	(-0.231)	
Healthcare	-2.107	-0.183	0.210	-0.171	0.099	0.241	1.136	-0.913	0.096	2.957
	(-0.392)	(-1.439)	(2.306)	(-0.394)	(2.742)	(0.767)	(0.815)	(-0.988)	(0.202)	
Finance	9.651	-0.241	0.290	2.013	0.188	-0.370	-2.253	0.231	-0.457	7.465
	(1.605)	(-0.921)	(2.279)	(2.982)	(2.505)	(-1.025)	(-1.533)	(0.169)	(-0.964)	
Other	5.336	-0.254	0.329	1.324	0.172	-0.367	-0.746	-0.688	-0.265	5.683
	(0.912)	(-1.203)	(3.159)	(2.523)	(3.057)	(-1.116)	(-0.516)	(-0.571)	(-0.610)	

Panel C. Annual Return Prediction

	Constant	Conditional				3MTB	Log(DY)	CS	TS	Adj. R^2(%)
		RV	VRP	TRP	RT					
Non-Durables	0.170	-0.010	0.043	0.186	0.004	0.072	-0.082	0.447	0.088	1.750
	(0.032)	(-0.213)	(0.834)	(0.916)	(0.237)	(0.305)	(-0.055)	(1.213)	(0.233)	
Durables	6.142	0.015	0.064	0.174	0.030	-0.244	-2.259	1.863	-0.067	17.170
	(0.617)	(0.213)	(0.597)	(0.404)	(1.060)	(-0.593)	(-0.763)	(2.435)	(-0.111)	
Manufacturing	7.961	-0.029	0.034	-0.155	0.000	-0.345	-2.121	0.671	-0.380	6.759
	(1.052)	(-0.579)	(0.508)	(-0.460)	(0.017)	(-1.210)	(-0.944)	(1.215)	(-0.917)	
Energy	5.663	-0.074	0.012	-0.359	-0.017	-0.300	-1.018	-0.484	-0.288	2.595
	(0.969)	(-1.366)	(0.200)	(-1.254)	(-0.863)	(-1.272)	(-0.604)	(-0.962)	(-0.883)	
Chemicals	4.607	-0.029	-0.018	-0.057	0.001	-0.154	-1.353	0.872	-0.143	6.068
	(0.905)	(-0.633)	(-0.326)	(-0.221)	(0.044)	(-0.751)	(-0.916)	(2.250)	(-0.455)	
Equipment	8.948	-0.124	0.129	-0.803	0.037	-0.285	-2.279	-0.144	-0.254	6.417
	(1.146)	(-0.923)	(1.292)	(-1.278)	(1.202)	(-0.925)	(-1.070)	(-0.199)	(-0.393)	
Telecommunications	3.302	-0.142	0.059	-0.609	0.052	-0.166	-0.700	-0.044	0.085	4.184
	(0.422)	(-1.151)	(0.501)	(-1.087)	(1.601)	(-0.473)	(-0.323)	(-0.059)	(0.153)	
Utilities	-2.808	-0.051	0.054	0.040	0.015	0.171	0.956	-0.317	0.358	-0.894
	(-0.466)	(-0.996)	(0.807)	(0.146)	(0.633)	(0.687)	(0.561)	(-0.777)	(1.124)	
Wholesale	2.535	-0.017	0.123	0.015	0.024	-0.064	-0.632	0.084	-0.003	8.430
	(0.478)	(-0.311)	(1.989)	(0.061)	(1.237)	(-0.246)	(-0.420)	(0.180)	(-0.008)	
Healthcare	2.354	-0.009	0.060	-0.002	0.008	-0.025	-0.452	0.124	-0.235	1.360
	(0.414)	(-0.134)	(0.905)	(-0.006)	(0.322)	(-0.083)	(-0.283)	(0.264)	(-0.524)	
Finance	9.994	0.057	0.041	0.371	-0.009	-0.252	-3.104	0.770	-0.292	10.720
	(1.059)	(0.823)	(0.502)	(0.961)	(-0.350)	(-0.622)	(-1.122)	(1.379)	(-0.540)	
Other	5.799	-0.048	0.028	-0.108	0.003	-0.282	-1.681	0.649	-0.152	7.538
	(0.746)	(-0.854)	(0.414)	(-0.319)	(0.112)	(-0.919)	(-0.741)	(1.368)	(-0.352)	

Table 8. Regressions Using Residuals

This table reports the predictability regression results using the two-stage regression that is explained in section 2.6.1. All other variables are described in Table 3. We employ the data directly from the Fama-French data library.

Panel A. Regression using residuals from VRP on TRP			
Conditional	Monthly Return	Quarterly Return	Annual Return
RV	-2.207	-2.673	-0.768
	(-0.660)	(-1.674)	(-0.994)
ε^1	5.168	3.553	0.629
	(3.172)	(3.297)	(0.798)
RT	2.376	1.779	0.225
	(2.651)	(3.683)	(0.827)
TRP	11.044	-1.696	-4.732
	(0.996)	(-0.312)	(-1.460)
3MTB	0.769	-1.612	-2.581
	(0.195)	(-0.476)	(-0.813)
Log(DY)	12.016	-2.160	-18.599
	(0.656)	(-0.120)	(-0.793)
CS	-12.593	-10.985	1.718
	(-0.806)	(-0.816)	(0.297)
TS	2.126	-1.139	-2.467
	(0.388)	(-0.237)	(-0.521)
Constant	-11.111	39.008	70.004
	(-0.148)	(0.567)	0.877
Adj. R^2(%)	4.602	4.820	2.685

Panel B. Regression using residuals from TRP on VRP			
Conditional	Monthly Return	Quarterly Return	Annual Return
RV	-2.207	-2.673	-0.768
	(-0.660)	(-1.668)	(-0.994)
ε^2	29.260	10.824	-2.519
	(2.513)	(2.161)	(-0.591)
RT	2.376	1.779	0.225
	(2.651)	(3.699)	(0.827)
VRP	2.009	2.383	0.904
	(1.192)	(2.213)	(1.502)
3MTB	0.769	-1.612	-2.581
	(0.195)	(-0.480)	(-0.813)
Log(DY)	12.016	-2.160	-18.599
	(0.656)	(-0.123)	(-0.793)
CS	-12.593	-10.985	1.718
	(-0.806)	(-0.806)	(0.297)
TS	2.126	-1.139	-2.467
	(0.388)	(-0.241)	(-0.521)
Constant	-16.340	34.523	68.743
	(-0.218)	(0.515)	(0.859)
Adj. R^2(%)	4.602	4.820	2.685

Table 9. Goodness-of-Fit of the VIX Decomposition for the second subsample (Unconditional Estimates)

This table reports the *decomposed R^2* of the orthogonalized VIX components, $RV^\perp$, $VRP^\perp$, $RT^\perp$, and $TRP^\perp$, respectively, for the second subsample from November 15, 2002, to September 10, 2014. We also examine three different VIX levels for the second subsample, including (1) *Nervous Market Condition*: VIX $\geq$ 23.07 (75 percentile), (2) *Normal Market Condition*: 13.74 $\leq$ VIX < 23.07, and (3) *Calm Market Condition*: VIX < 13.74 (25 percentile), correspondingly. We employ Klein and Chow (2013) to orthogonalize the VIX's decomposed variables and further calculate their *decomposed R^2*. Also, to avoid bias results due to measurements at different scales, all variables are standardized for the analyses.

A. Second Subsample	$RV^\perp$	$VRP^\perp$	$RT^\perp$	$TRP^\perp$
Coefficient	0.76	0.21	-0.12	-0.75
Decomposed R^2(%)	54.60	3.39	0.15	41.86

B. Second Subsample for VIX $\geq$ 23.07 (*Nervous Market Condition*)				
Coefficient	0.56	0.08	-0.05	-0.61
Decomposed R^2(%)	45.72	1.20	0.41	52.68

C. Second subsample for 13.74 $\leq$ VIX < 23.07 (*Normal Market Condition*)				
Coefficient	0.77	0.33	0.01	-0.36
Decomposed R^2(%)	67.00	26.69	0.02	6.30

D. Second subsample for VIX < 13.74 (*Calm Market Condition*)				
Coefficient	0.55	0.25	-0.00	-0.12
Decomposed R^2(%)	64.13	34.00	0.00	1.87

Table 10. Goodness-of-Fit of the VIX Decomposition for the second subsample ("Unbiased" Conditional Estimates)

This table reports the *decomposed R^2* of the four orthogonalized VIX components, $RV^\perp$, $VRP^\perp$, $RT^\perp$, and $TRP^\perp$, respectively, for the second subsample from November 15, 2002, to September 10, 2014. We also examine three different VIX levels for the second subsample, including (1) *Nervous Market Condition*: VIX $\geq$ 23.07 (75 percentile), (2) *Normal Market Condition*: 13.74 $\leq$ VIX < 23.07, and (3) *Calm Market Condition*: VIX < 13.74 (25 percentile), correspondingly. We employ Klein and Chow (2013) to orthogonalize the VIX's decomposed variables and further calculate their *decomposed R^2*. The conditional measures for the second subsample are calculated using the parameters estimated from the first subsample based on the forecasting models shown in equations (26), (27) and (28). In addition, to avoid bias results due to measurements at different scales, all variables are standardized for the analyses.

A. Second Subsample	$RV^\perp$	$VRP^\perp$	$RT^\perp$	$TRP^\perp$
Coefficient	0.63	0.61	-0.02	-0.47
Decomposed R^2(%)	39.76	37.66	0.04	22.54

B. Second Subsample for VIX $\geq$ 23.07 (*Nervous Market Condition*)				
Coefficient	0.56	0.53	-0.04	-0.46
Decomposed R^2(%)	43.11	27.56	0.24	29.08

C. Second Subsample for 13.74 $\leq$ VIX < 23.07 (*Normal Market Condition*)				
Coefficient	0.73	0.62	-0.02	-0.17
Decomposed R^2(%)	29.04	69.45	0.10	1.41

D. Second Subsample for VIX < 13.74 (*Calm Market Condition*)				
Coefficient	0.59	0.54	-0.02	-0.04
Decomposed R^2(%)	27.06	72.45	0.28	0.20

Table 11. Out-of-Sample Prediction (December 2002 to September 2014)

This table reports the out-of-sample performance of the four variables decomposed from the VIX. The in-sample R^2 is obtained using the conditional measure which is calculated from the parameters estimated for the same (second) subsample, while the out-of-sample R^2 is obtained using the "unbiased" conditional measure, which is calculated from the parameters estimated for the first subsample.

Conditional	In-Sample R^2	Prediction Begins From December 2002	
		In-Sample R^2	Out-of-Sample R^2
RV	0.405	2.924	1.927
VRP	-0.343	-0.431	-0.517
RT	-0.183	-0.589	-0.597
TRP	2.341	2.832	2.773

Table 12. Monthly Return Prediction using the "Unbiased" Conditional Variable for the Second Subsample

This table reports the regression coefficients and adjusted R^2 from return predictability regressions for monthly excess returns on the S&P 500 market portfolio using the "unbiased" conditional measures for the second subsample. We use the first subsample to estimate the conditional regression parameters and then apply the estimated parameters to calculate the "unbiased" conditional RV, VRP, RT, and TRP. The second subsample extends from December 2002 to September 2014. Newey-West t-statistics are reported in parentheses. Adj. R^2 is the adjusted coefficient of determination. RT is in annualized basis point, and all other variables are measured by annualized percentage.

	December 2002 to September 2014			
RV	-2.685			
	(-1.709)			
VRP		-0.742		
		(-0.287)		
RT			-5.703	
			(-0.452)	
TRP				16.795
				(2.761)
Constant	13.514	6.493	4.534	9.263
	(3.536)	(1.663)	(0.895)	(2.264)
Adj. R^2 (%)	1.927	-0.517	-0.594	2.773

Table 13. Monthly Return Prediction using the "Unbiased" Conditional Variable for the Second Subsample

This table reports the monthly predictability regression results from excess returns on *Size* (20% smallest and biggest firms), *Book-to-Market* (20% highest and lowest B/M ratios), and *Momentum* (20% top and bottom performance), along with the corresponding zero-cost portfolios using the "unbiased" conditional variable for the second subsample, which is described in 6.2. All other variables are described in Table 3.

	Constant	RV	VRP	TRP	RT	3MTB	Log(DY)	CS	TS	Adj. R^2(%)
	December 2002 to September 2014									
Small	10.780	0.022	0.161	3.234	-0.754	-0.244	-3.342	1.039	0.201	2.411
	(0.843)	(0.041)	(0.391)	(1.496)	(-0.436)	(-0.319)	(-1.030)	(0.562)	(0.197)	
Big	3.330	0.093	0.386	2.262	-0.450	-0.204	-0.416	-1.134	-0.135	0.566
	(0.404)	(0.226)	(1.535)	(1.280)	(-0.436)	(-0.425)	(-0.209)	(-0.852)	(-0.211)	
SMB	7.450	-0.071	-0.225	0.672	-0.304	-0.040	-2.926	2.173	0.336	6.017
	(1.191)	(-0.328)	(-1.109)	(0.904)	(-0.405)	(-0.105)	(-1.799)	(2.510)	(0.584)	
High	7.027	0.387	0.021	4.181	-0.836	-0.397	-1.847	0.366	-0.173	0.155
	(0.605)	(0.717)	(0.052)	(1.706)	(-0.526)	(-0.517)	(-0.667)	(0.208)	(-0.173)	
Low	-0.297	-0.013	0.500	2.587	-0.605	-0.054	0.437	-0.710	0.066	1.721
	(-0.036)	(-0.031)	(2.186)	(1.295)	(-0.514)	(-0.118)	(0.219)	(-0.649)	(0.104)	
HML	7.324	0.400	-0.480	1.594	-0.231	-0.343	-2.284	1.076	-0.239	4.006
	(1.230)	(1.865)	(-2.308)	(2.423)	(-0.194)	(-0.706)	(-1.759)	(0.982)	(-0.361)	
Winners	6.235	0.087	0.285	2.507	0.159	-0.164	-1.150	-1.629	0.101	0.450
	(0.596)	(0.186)	(0.911)	(1.296)	(0.131)	(-0.227)	(-0.490)	(-1.009)	(0.099)	
Losers	7.555	0.277	0.836	6.469	-3.554	-0.254	-2.617	1.139	-0.248	5.226
	(0.419)	(0.339)	(1.813)	(1.685)	(-1.240)	(-0.274)	(-0.575)	(0.412)	(-0.201)	
WML	-1.321	-0.189	-0.551	-3.962	3.713	0.091	1.467	-2.767	0.349	7.750
	(-0.105)	(-0.290)	(-1.564)	(-1.207)	(1.375)	(0.123)	(0.472)	(-1.516)	(0.327)	

Table 14. Monthly Return Prediction using the "Unbiased" Conditional Variable for the Second Subsample
This table reports the monthly predictability regression results from excess returns on twelve industry portfolios using the "unbiased" conditional variable for the second subsample, which is described in 6.2. All other variables are described in Table 3. We employ the data directly from the Fama-French data library.

	Constant	RV	VRP	TRP	RT	3MTB	Log(DY)	CS	TS	Adj. R²(%)
				December 2002 to September 2014						
Non-Durables	2.594	-0.057	0.198	1.519	0.380	-0.147	-0.336	-0.304	0.025	0.107
	(0.451)	(-0.164)	(0.849)	(0.936)	(0.484)	(-0.406)	(-0.227)	(-0.279)	(0.048)	
Durables	1.718	-0.024	0.951	4.794	-2.587	0.038	-0.554	-0.517	0.331	2.351
	(0.099)	(-0.033)	(1.789)	(1.204)	(-0.856)	(0.034)	(-0.134)	(-0.228)	(0.228)	
Manufacturing	4.697	-0.024	0.506	2.391	-1.140	0.138	-1.057	-1.163	0.245	-1.531
	(0.385)	(-0.045)	(1.144)	(0.944)	(-0.616)	(0.198)	(-0.359)	(-0.627)	(0.263)	
Energy	12.857	-0.096	0.187	0.337	0.862	-0.180	-2.792	-1.628	-0.303	-2.992
	(1.202)	(-0.182)	(0.463)	(0.128)	(0.556)	(-0.239)	(-1.183)	(-0.854)	(-0.287)	
Chemicals	0.378	-0.051	0.527	2.369	-0.495	0.188	0.305	-1.496	0.350	2.215
	(0.044)	(-0.137)	(2.157)	(1.341)	(-0.379)	(0.329)	(0.160)	(-1.061)	(0.456)	
Equipment	-2.595	0.090	0.538	3.913	-0.643	0.188	0.649	-0.057	0.423	1.329
	(-0.261)	(0.185)	(1.859)	(1.804)	(-0.437)	(0.332)	(0.254)	(-0.043)	(0.515)	
Telecommunications	1.835	0.328	0.225	2.754	-0.471	-0.244	-0.084	-0.909	-0.186	-3.345
	(0.195)	(0.591)	(0.906)	(0.965)	(-0.377)	(-0.387)	(-0.038)	(-0.709)	(-0.224)	
Utilities	9.655	0.079	0.086	0.710	0.646	-0.153	-2.156	-1.460	-0.062	-1.245
	(1.552)	(0.197)	(0.351)	(0.278)	(0.895)	(-0.447)	(-1.377)	(-1.066)	(-0.121)	
Wholesale	4.209	0.361	0.403	3.990	-0.300	-0.419	-0.686	-0.900	-0.332	2.183
	(0.538)	(0.876)	(1.899)	(2.289)	(-0.262)	(-0.834)	(-0.353)	(-1.043)	(-0.516)	
Healthcare	0.230	-0.195	0.070	0.067	1.040	-0.406	0.645	-0.041	-0.217	-0.853
	(0.032)	(-0.504)	(0.314)	(0.041)	(1.249)	(-0.903)	(0.364)	(-0.035)	(-0.342)	
Finance	11.345	0.225	0.471	4.235	-2.340	-0.802	-2.415	-1.187	-0.878	3.363
	(0.891)	(0.349)	(1.284)	(1.711)	(-1.106)	(-0.957)	(-0.802)	(-0.538)	(-0.843)	
Other	7.245	0.222	0.603	3.815	-1.143	-0.248	-1.385	-2.254	-0.141	2.792
	(0.656)	(0.371)	(1.972)	(1.322)	(-0.742)	(-0.361)	(-0.536)	(-1.352)	(-0.159)	

Chapter 3   The Predictive Power of Tail Risk Premia on Individual Stock Returns

Chapter Abstract

This chapter introduces a novel, option-free methodology to calculate the tail risk premium for individual stocks and examines the characteristics of this premium in the cross-section of stock returns. The existence of a premium for bearing negative tail risk is significantly associated with negative returns up to one month in the future. In contrast, the premium for bearing positive tail risk has no significant predictive power. Further, the larger the magnitude of the premium for bearing negative tail risk, the greater and longer lasting its impact on expected future returns.

This chapter is coauthored with Victor Chow (West Virginia University) and Ben Sopranzetti (Rutgers Business School – Newark and New Brunswick).

3.1   Introduction

Compensation for extreme tail event risk is formally referred to in academic finance literature as a "tail risk premium." Bollerslev, Todorov and Xu (2015) shows that a majority of the predictability in the variance risk premium is attributed to this premium for bearing jump tail risk, and that, specifically, it is negative tail risk and not positive tail risk that seems to be priced. The impact of tail events on returns is well-documented at the aggregate market level, but not so much is known about its impact at the individual stock level. One reason why is that out of the money call and put options are required to determine the tail risk in the risk-neutral probability space. Although out of the money options are prevalent for an index such as the S&P500, they either do not exist or are illiquid for most stocks. For this reason, to date no paper has directly examined both the impact of tail risk and its premium on the cross-section of individual stock returns. Given that the return distribution for individual stocks will likely exhibit a greater proclivity for extreme events than the return distribution for a diversified market portfolio (where extreme negative events in some securities might be tempered by extreme positive events in others), one would expect that tail risk should play a more prominent role in the returns for individual stocks than it would for a market portfolio. Consequently, a careful study of tail risk premia for individual stocks may yield new and heretofore unseen insights into their predictive power for future returns.

Kelly and Jiang (2014) are the first to examine tail risk in the cross section of individual stock returns. This important paper employs an aggregate measure of time-varying tail risk that relies on panel estimation from the cross-section of stock returns. It then measures a stock's sensitivity to this measure of tail risk by sorting portfolios into quintiles based on tail beta-exposure, and documents that the lowest tail beta quintile is associated with the lowest future returns, while the highest tail beta quintile is associated with the highest future returns. Although Kelly and Jiang (2014) provide strong evidence that tail risk is priced in individual stocks, their paper does not directly calculate the tail risk premium nor does it examine any asymmetry in the way positive and negative tail risk premia affect future returns.

The current chapter differentiates itself from Kelly and Jiang (2014) in three critical ways. First, rather than using an aggregate measure of tail risk and indirectly examining the sensitivity of a stock's return to this aggregate measure, the current chapter directly calculates the tail risk premium for individual stocks and examines how this premium varies across the cross-section of stock returns. Specifically, the chapter introduces a novel, nonparametric approach to directly determine the tail risk premium. The approach avoids the need for the use of liquid out of the money stock options (which don't exist for most stocks). The second contribution is that this new approach allows stocks to be sorted by their exposure to tail risk, so that the impact of positive and negative tail risk premia on future returns can be examined separately. Stocks with exposure to negative tail risk require a tail risk premium that is positive (investors demand a higher return today than otherwise expected for bearing negative tail risk), while those with positive tail risk require a tail risk premium that is negative (investors are willing to accept a lower return today when there is a chance for extreme positive events). Third, this chapter documents that almost all of the extreme jumps are concentrated in the first and tenth deciles; consequently, an analysis of deciles, and even percentiles, rather than the quintiles examined in prior studies, is necessary if researchers are to better understand how extreme jump tail risk is priced.

The results in this chapter provide evidence on the differential pricing of information related to negative and positive tail risk. Bollerslev, Todorov and Xu (2015) and Kelly and Jiang (2014) find evidence of pricing for negative tail risk, but neither fully examines the extent to which positive tail risk is priced. The current chapter documents that the existence of a premium for bearing negative tail risk today is associated with significantly lower future monthly returns, but that the existence of a premium for positive tail risk does not have statistically significant predictive power in the cross-section of individual stock returns. In addition, the current chapter presents evidence that it is not only the sign of the tail risk premium that matters in predicting future returns but also its magnitude. The larger and more positive the current tail risk premium (that is, the greater the concerns about a big negative jump), the more negative and persistent the association with future returns will be.

The chapter's empirical methodology controls for several explanations previously offered in the literature for the existence and pricing of tail risk, including momentum (Lehmann, 1990; Jegadeesh & Titman, 1993), lottery effects (Barberis & Huang, 2008; Bali, Cakici & Whitelaw, 2011), idiosyncratic volatility (Ang, Hoderick, Xing, & Zhang, 2006), illiquidity (Amihud, 2002), market beta (Scholes & Williams, 1977; Dimson, 1979), maximum and minimum monthly return (Bali, Cakici, & Whitelaw, 2011). The predictive power of the premium for bearing negative tail risk on future returns survives the inclusion of these control variables.

This chapter is organized as follows: Section 3.2 contains the literature review; Section 3.3 demonstrates individual stock level tail risk premium estimation and the data; Section 3.4 contains the tail risk premium cross-sectional pricing characteristics and cross-sectional return tests; Section 3.5 includes robustness checks; and Section 3.6 concludes the chapter.

3.2   Literature Review

In addition to the papers mentioned in the introduction, there are several recent papers that examine the pricing of downside risk that are related to the current chapter. Ang, Chen, and Xing (2006) find that stocks that covary strongly with the market during periods of market decline tend to have higher average returns than other stocks. Investors are downside risk averse and therefore require a premium to hold these assets. Bali, Cakici and Whitelaw (2014) introduce a hybrid tail covariance risk measure that measures stock return tail covariance risk. The measure is based on the basic form of lower partial moments. The paper documents a significant positive premium for bearing negative tail risk captured in the cross-section.

This chapter is also related to the literature on crash risk. Kelly and Jiang (2014) are among the first of the papers that examine extreme crash risk on stock returns. The paper finds that stocks with high loadings on market tail risk earn higher abnormal returns. Chabi-Yo, Ruenzi, and Weigert (2018) find that investors are crash-averse; that is, they receive positive compensation for holding crash-sensitive stocks through the measure of "lower tail dependence" from individual stock price distributions. The findings in these papers are consistent with the downside risk

literature (Ang, Chen, & Xing, 2006; Bali, Cakici, & Whitelaw, 2014) that investors are downside risk averse and require a positive premium for holding the crash risk sensitive stocks.

Bali, Cakici, and Whitelaw (2014) construct a firm-specific tail risk measure based on lower partial moments of stock returns and find that it negatively predicts future stock returns. Almeida, Ardison, Garcia, and Vicente (2017) adopt a risk-neutral excess expected shortfall approach to construct a nonparametric tail risk measure. The paper finds that the risk-neutral tail risk measure possesses negative predictive power for intermediate horizon stock returns. Lu and Murray (2017) construct a proxy for bear-market risk and finds it to be negatively priced; that is, stocks with a high sensitivity to bear-market risk are found to underperform their low-sensitivity counterparts.

This chapter is also related to the asset pricing literature on higher moments. Traditional finance theory assumes a normal distribution of asset returns, for which mean and variance together are sufficient to characterize the entire return distribution. The capital asset pricing model (Sharpe, 1964; Lintner, 1965; Mossin, 1966) predicts that market volatility is a determinant of the market equity premium. Contrary to this notion, Ang et al. (2006) examines whether aggregate volatility innovation is priced in the cross-section of stock returns, and concludes that high sensitivity stocks have subsequently lower average returns. Given this controversy, it is natural to ask whether other return distributional characteristics are also priced in the cross-section. Chang, Christofferson and Jacobs (2013) show that the cross-section of stock returns has substantial exposure to higher moments. Cremers, Halling and Weinbaum (2015) find that although both jumps and volatility are priced in the cross-section, jumps seem to have a larger impact on returns than does volatility. Bali, Cakici, and Whitelaw (2011) find that stocks with maximum returns have a significant negative return in the following month. These pricing findings are consistent with the erroneous probability weighting of investors as modeled in Barberis and Huang (2008) and the optimal belief framework of economic agents modeled in Brunnermeier, Gollier, and Parker (2007).

3.3   Calculation of the Tail Risk Premium in the Cross Section of Individual Stock Returns
3.3.1   Methodology

This section discusses the construction of the tail risk premium associated with jumps in returns for individual stocks. The methodology is an innovation on the well-established notion – Bollerslev, Todorov and Xu (2015), Carr and Wu (2008), and others – that the jump tail risk premium can be calculated as the difference between the expectation of the tail variation in the physical probability space ($\mathbb{P}$-space) and its counterpart in the risk-neutral probability space ($\mathbb{Q}$-space).

To this end, we define the infinite-order polynomial variation of log returns, which captures not only the second-order (quadratic) variation (see Carr & Wu, 2008), but also the higher-order (third-order and up) variations, which Jiang and Oomen (2008) have shown to be associated with jumps in stock returns. We denote the simple return $R_{t+1} = \frac{S_{t+1} - S_t}{S_t}$ and logarithmic return $r_{t+1} =$

$\ln\left(\frac{S_{t+1}}{S_t}\right)$ over a period from t to $t+1$. Formally, based on Merton's (1976) jump diffusion process, the realized infinite-order *polynomial variation* ($\mathbb{PV}$) for individual asset i at time $t+1$ can be expressed as follows:

$$
\begin{aligned}
\mathbb{PV}_{i,[t,t+1]} = 2\big(R_{i,t+1} - r_{i,t+1}\big) &\underbrace{= 2 \times \left(\frac{1}{2!} \times r_{i,t+1}^2 + \frac{1}{3!} \times r_{i,t+1}^3 + \cdots + \frac{1}{n!} \times r_{i,t+1}^n\right)}_{Polynomial\ form\ of\ \log return\ variations} \\
&= \int_t^{t+1} \sigma_{i,t}^2 dt + \sum_{n=2}^{\infty} \frac{2}{n!} \int_t^{t+1} \int_{\mathbb{R}^0} x_i^n \mu(dx_i, dt) \\
&= \mathbb{CV}_{i,[t,t+1]} + \mathbb{JPV}_{i,[t,t+1]}.
\end{aligned} \tag{1}
$$

where σ is the volatility. $\mu(dx_i, dt)$ is the Poisson random measure for the compound Poisson process with compensator equal to $\lambda \frac{1}{\sqrt{2\pi\sigma_j^2}} e^{-\frac{1}{2}(x-\alpha)^2}$, with λ as the jump intensity. $\mathbb{CV}$ is the integral of the continuously instantaneous variance (often referred to as the *integrated volatility*), and $\mathbb{JPV}$ represents the realized jump component of the infinite-order *polynomial variation.* Analogously, the second-order polynomial variation (the realized quadratic variance, denotes $\mathbb{QV}$) can be written by the following equation:

$$
\mathbb{QV}_{i,[t,t+1]} = r_{i,t+1}^2 = \int_t^{t+1} \sigma_{i,t}^2 \, dt + \int_t^{t+1} \int_{\mathbb{R}^0} x_i^2 \mu(dx_i, dt) = \mathbb{CV}_{i,[t,t+1]} + \mathbb{JQV}_{i,[t,t+1]}. \tag{2}
$$

By subtracting Equation 2 from Equation 1, we then have the realized tail-jump variation at time *t* such that

$$
\begin{aligned}
\mathbb{TV}_{i,[t,t+1]} &= 2\big(R_{i,t+1} - r_{i,t+1}\big) - r_{i,t+1}^2 = \mathbb{PV}_{i,[t,t+1]} - \mathbb{QV}_{i,[t,t+1]} \\
&= \sum_{n=3}^{\infty} \frac{2}{n!} \int_t^{t+1} \int_{\mathbb{R}^0} x_i^n \mu(dx_i, dt).
\end{aligned} \tag{3}
$$

Now that we have the unconditional realized tail-jump variation, we next present the conditional ex-ante estimation of the tail-jump variation and then will develop a proxy for the tail-risk premium.

Following Bollerslev, Tauchen and Zhou (2009), under the assumption that $\mathbb{TV}$ is a martingale[24], the $\mathbb{P}$-space expected tail-variation of returns at time t can be expressed as follows:

---

[24] Under the Merton (1976) jump diffusion model assumption, the compensated compound Poisson process $\mu(dx_i, dt)$, with compensator $\lambda \frac{1}{\sqrt{2\pi\sigma_j^2}} e^{-\frac{1}{2}(x-\alpha)^2}$, is a martingale process; consequently, $\mathbb{PV}_{i,[t,t+1]}$ is also a martingale. Todorov and Tauchen (2011) provide empirical evidence that the VIX index is a pure-jump process without a continuous component, which supports the notion that $\mathbb{PV}_{i,[t,t+1]}$ is a martingale process. [Du and Kapadia (2012) demonstrates that $\mathbb{PV}_{i,[t,t+1]}$ is the underlying process of the VIX index]. Furthermore, it is widely accepted in both

$$E_t^{\mathbb{P}}\big(\mathbb{TV}_{i,[t,t+1]}\big) = 2\big(R_{i,t} - r_{i,t}\big) - r_{i,t}^2. \tag{4a}$$

Then, the difference between expected $\mathbb{TV}$ in the $\mathbb{P}$-space and that in the risk-neutral $\mathbb{Q}$-space, $E_t^{\mathbb{P}}\big(\mathbb{TV}_{i,[t,t+1]}\big) - E_t^{\mathbb{Q}}\big(\mathbb{TV}_{i,[t,t+1]}\big)$, serves as a proxy for the tail-risk premium. The advantages of using $E_t^{\mathbb{P}}\big(\mathbb{TV}_{i,[t,t+1]}\big)$ as a tail risk measure in the physical space are threefold. First, it is nonparametric, i.e., it does not require the estimation of a cutoff value as in Kelly and Jiang (2014) or Bollerslev and Todorov (2011b). Second, it does not require the estimation of a jump compensator in order for the instantaneous arithmetic stock return to be a semi-martingale process as in Bollerslev and Todorov (2011a), Bollerslev and Todorov (2011b), or Bollerslev, Todorov and Xu (2015). Third, $E_t^{\mathbb{P}}\big(\mathbb{TV}_{i,[t,t+1]}\big)$ only relies on stock price information and can be easily calculated using databases such as the WRDS CRSP database. Thus, relative to the measures used in the afformentioned papers, our measure not only lessens estimation error but also shortens the calculation time, and because it only relies on prices, it is broadly applicable to other asset classes. Once the $\mathbb{P}$-space tail risk measure is calculated, then one can examine the innovations in this measure

$$\Delta E_t^{\mathbb{P}}\big(\mathbb{TV}_{i,[t,t+1]}\big) = \Delta[2\big(R_{i,t} - r_{i,t}\big) - r_{i,t}^2] \tag{4b}$$

In contrast to the $\mathbb{P}$-space tail variation measure, which is easy to calculate, the corresponding calculation of Equation 4a in the $\mathbb{Q}$-space for individual stocks is much more problematical, since the necessary data is not readily available. If, instead of examining individual stocks, one were interested in calculating the $\mathbb{Q}$-space tail variation of the market as a whole, then the methodology would be relatively easy to implement. For example, Carr and Wu (2008) shows that the CBOE VIX index is a measure of moment combinations, and therefore a polynomial variation in the risk-neutral probability space. Specifically, that paper argues that the VIX index measures the risk-neutral expectation of the polynomial variation process for the S&P 500 market index,[25]

$$E_t^{\mathbb{Q}}\big(\mathbb{PV}_{M,[t,t+1]}\big) = VIX_t^2 \tag{5}$$

Once Equation 5 has been calculated, Du and Kapadia (2012) and Chow, Jiang, and Li (2014) show that the $\mathbb{Q}$-space measure of tail variation for the market can be calculated as the difference between the square of the VIX and the centralized Bakshi, Kapadia, and Madan (2003) volatility measure, V_{BKM}^C such that [26]

---

theory and practice that $\mathbb{QV}_{i,[t,t+1]}$ is a martingale process; consequently, the claim that $\mathbb{TV}_{i,[t,t+1]}$ is a martingale process has both theoretical and practical support.

[25] Note that, Equation 5 clearly indicates that the literature-prevalent variance risk premium estimation methodology, which takes the difference between the VIX index and physical probability space quadratic variation is biased. Specifically, the VIX, because it includes higher order moments, undervalues (overvalues) volatility when the market return is expected to be negatively (positively) skewed.

[26] $V_{BKM}^C = V_{BKM} - \mu_{BKM}^2$, where $\mu_{BKM} = \ln\left(\frac{K_0}{S_0}\right) + \left(\frac{F_0}{K_0} - 1\right) - e^{rT}[\int_{K_0}^{\infty} \frac{1}{K^2} C_T(K)dK + \int_0^{K_0} \frac{1}{K^2} P_T(K)dK]$ and $V_{BKM} = \ln^2\left(\frac{K_0}{S_0}\right) + 2\ln\left(\frac{K_0}{S_0}\right)\left(\frac{F_0}{K_0} - 1\right) + 2e^{rT}[\int_{K_0}^{\infty} \frac{[1-\ln(\frac{K}{S_0})]}{K^2} C_T(K)dK + \int_0^{K_0} \frac{[1+\ln(\frac{K}{S_0})]}{K^2} P_T(K)dK]\}$ as in Bakshi, Kapadia, and Madan (2003).

$$E_t^{\mathbb{Q}}\left(\mathbb{TV}_{M,[t,t+1]}\right) = VIX_t^2 - V_{BKM,t}^C \tag{6}$$

It is important to note that the calculation of both the VIX and V_{BKM}^C rely on highly liquid out of the money put and call options[27], which fortunately are prevalent on the S&P 500 index. Recent papers by Gao, Gao, and Song (2018) and Gao, Lu and Song (2018) estimate tail risk based on the $\mathbb{Q}$-space tail variation measure (in Equation 6) for the market index and portfolio of assets where there exists liquid option trading. Unfortunately, these options often either do not even exist for individual stocks or, if they do exist, are not frequently traded, and thus an analogue of the aforementioned methodology to calculate the $\mathbb{Q}$-space tail variation for individual stocks is impossible to implement. Consequently, an alternative is required.

To this end, we propose a methodology for the estimation of the $\mathbb{Q}$-space tail variation that is based on the groundbreaking work on tracking portfolios presented in Breeden, Gibbons, and Litzenberger (1989) and Lamont (2001). According to Lamont (2001), "A tracking portfolio for any variable *y* can be obtained as the fitted value of a regression of *y* on a set of base asset returns. The portfolio weights for the economic tracking portfolio for *y* are identical to the coefficients of an OLS regression." Ang et al. (2006) apply the tracking portfolio technology and use returns to capture innovations in the VIX index. We employ a modified version of Ang et al. (2006); specifically, we use first-order difference in the $\mathbb{Q}$-space tail variation measures, rather than the raw measures themselves, in order to capture innovations in $\mathbb{P}$-space tail variation measures. Accordingly, we estimate the following ordinary least squares regression for each stock in each month to obtain our portfolio weights, β_i[28]

$$\Delta[2\left(R_{i,t} - r_{i,t}\right) - r_{i,t}^2] = \alpha_i + \beta_i \cdot \Delta\left(VIX_{t-22}^2 - V_{BKM,t-22}^C\right) + \varepsilon_{i,t} \tag{7}$$

where $VIX_{t-22}^2 - V_{BKM,t-22}^C$ represents the tail variation in the $\mathbb{Q}$-space, as delineated in Equation 6, and $2\left(R_{i,t} - r_{i,t}\right) - r_{i,t}^2$ represents the tail variation in the $\mathbb{P}$-space, as shown in Equation 4a. Note that we follow the precedent set by Bekaert and Hoerova (2014) which estimates $\mathbb{P}$-spaced conditional realized variation utilizing a 22-day lag. Their approach is based on the notion that options-based $\mathbb{Q}$-spaced measures, such as the VIX, are forward-looking, and thus there is a time lag error of one month (22 trading days) that must be corrected.

Once the β_i coefficients have been obtained, then $\beta_i \cdot \Delta\left(VIX_{t-22}^2 - V_{BKM,t-22}^C\right)$ represents the $\mathbb{Q}$-space tracking portfolio that mimics innovations in tail variation that occur in the $\mathbb{P}$-space. Formally,

$$\Delta E_t^{\mathbb{Q}}\left(\mathbb{TV}_{i,[t,t+1]}\right) = \beta_i \cdot \Delta\left(VIX_{t-22}^2 - V_{BKM,t-22}^C\right) \tag{8}$$

---

[27] Demeterfi, Derman, Michael and Zou (1999) present a methodology for estimating the $\mathbb{Q}$-space measure of implied volatility for individual securities that is based on the variance swap concept, which requires highly liquid out of the money put and call options.

[28] Stocks must have at least 17 observations in any given month to be included in that month's regression.

Now that the daily innovations in the $\mathbb{P}$-space and $\mathbb{Q}$-space tail variation measures (Equations 4b and 8, respectively) have been obtained, the daily tail risk premium for any asset can be estimated by taking their difference

$$TRP^{daily}_{i,[t,t+1]} = \Delta E^{\mathbb{P}}_t\big(\mathbb{TV}_{i,[t,t+1]}\big) - \Delta E^{\mathbb{Q}}_t\big(\mathbb{TV}_{i,[t,t+1]}\big) \tag{9}$$

Since there are 22 trading days in a month, the corresponding monthly tail risk premium for each individual stock can be estimated as

$$TRP^{Monthly}_i = 22 \cdot TRP^{daily}_{i,[t,t+1]} \tag{10}$$

3.3.2 Data

We run the baseline regression model in Equation 7 for all common stocks on AMEX, NASDAQ, and NYSE, with more than 17 daily observations in any given month. Daily stock returns come from the WRDS CRSP database, over the sample period from January 1990 to September 2014. S&P index option data are obtained from IVolatility.com, which provides end-of-day and high frequency option data on major stock market indices across countries.

[Insert Table 1 Here]

Table 1 reports the summary statistics for portfolios sorted into deciles by the tail risk premium. Definitions for all the variables can be found in the Appendix. Panel A presents the decile portfolio firm-specific characteristics sorted by the tail risk premium. Firms with a higher tail risk premium tend to have lower lagged 1-month returns (short-term return reversal effect). Firms that fall into the extreme first and 10th decile also tend to be smaller firms that have higher market betas, higher idiosyncratic volatility, more illiquidity, higher maximum monthly returns, lower minimum monthly returns, lower trading volumes, and lower prices.

To examine the correlation structure among the explanatory variables, we report in percentage form Pearson correlation coefficients of the variables in Table 2.

[Insert Table 2 Here]

Idiosyncratic volatility (Ang et al., 2006) is negatively correlated with size (correlation coefficient of -49.70%), which is consistent with the findings in Fu (2009). Moreover, idiosyncratic volatility is also correlated with maximum and minimum monthly returns (Bali, Cakici, & Whitelaw, 2011), with correlation coefficients of 89.72% and 80.77%, respectively). Maximum and minimum monthly returns (Bali, Cakici, & Whitelaw, 2011) are correlated with size, but to a much lesser extent (-37.24% and 39.00%, respectively).

3.4 Predictive Power of the Tail Risk Premium on Future Returns
3.4.1 Portfolios Sorted by the Tail Risk Premium

To investigate the predictive power of the tail risk premium in the cross-section, we first calculate the tail risk premium for each of the stocks in our sample, and then sort the stocks into decile portfolios by the magnitude of their monthly tail risk premium. We next calculate the one month forward buy and hold returns for each decile portfolio. We term these returns as the $1/0/1$ (sort in one month, examine the one-month return for the following month) return. The results are reported in Panel A of Table 3.

[Insert Table 3 Here]

The lowest (Decile 1) tail risk premium portfolio earns the highest return of 2.02% in the following month, while the highest (Decile 10) portfolio earns the lowest return of 0.27%. The difference between the lowest and highest quintile portfolio is 1.75% monthly, and has a t-statistic of -11.26. After a Newey-West (1986) adjustment for heteroskedasticity and autocorrelation, the t-statistic is still strongly significant with a value of -7.30.

We next examine the length of time it takes for the market to correct this pricing error, by comparing the results for the $1/0/1$ (sort in one month, examine the one-month return for the following month) portfolio strategy discussed in the previous paragraph with $1/1/1$ and $1/2/1$ (sort in one month, examine the one-month return starting two months from now, and three months from now, respectively) portfolio strategies.

These results are reported in Panels B and C of Table 3. The existence of a tail risk premium at time *t* possesses virtually no impact on future returns moving from the second-next month into the future. This suggests that the adjustment period for market perception of tail risk seems to be somewhere between one month and two months, after which the market fully incorporates information about tail risk into the price.

3.4.2 Cross-Sectional Return Test for the Predictive Power of the Tail Risk Premium

The above evidence suggests that tail risk is priced at the individual stock level. Consistent with prior studies, we perform a more thorough firm-level cross-sectional return and examine whether the predictive power of the tail risk premium remains. Specifically, we estimate the following monthly regression:

$$R_{i,t+1} = \gamma_{0,t+1} + \gamma_{1,t+1} \times TRP_{i,t}^{Monthly} + \phi_{t+1}' \times Z_{i,t} + \varepsilon_{i,t+1} \tag{11}$$

where $R_{i,t+1}$ is the monthly stock return for stock i in month $t + 1$. $TRP_{i,t}^{Monthly}$ is the individual stock tail risk premium, delineated by Equation 11. $Z_{i,t}$ represents a vector of characteristics and controls for firm i at the end of month t, such as size, book-to-market ratio, and market beta. Controls are also provided for illiquidity following Amihud (2002), idiosyncratic volatility

following Ang et al. (2006), lagged 1-month return for short-term return reversal effect following Jegadeesh (1990) and Lehmann (1990), lagged 12-month return accounting for the momentum effect, and maximum and minimum monthly return following Bali, Cakici, and Whitelaw (2011).

Table 4 reports the time-series average of γ and ϕ coefficients for the cross-sectional regressions.

[Insert Table 4 Here]

Column 1 provides univariate results, and Column 2 adds firm-specific control variables.[29] The coefficient for the tail risk premium is negative and is statistically significant in both the univariate and multivariate regressions, with coefficients of -0.801 and -1.155 and Newey-West (1986) t-statistic equal to -5.60 and -5.50, respectively. Specifically, stocks with tail risk require a premium in the current month, and this premium is associated with lower returns the following month.

The results for the impact of overall tail risk on one-month future returns are interesting, but tail risk involves concerns about both extreme positive events and extreme negative events. Consequently, it may be of interest to examine, separately, the impact of positive and negative tail risk on future returns.

3.4.2.1 The Monthly Predictive Power of Positive versus Negative Tail Risk Premia

To investigate the extent to which positive and negative tail risk may be priced differentially in the cross-section of returns, we again perform firm-level cross-sectional monthly regressions, but this time we include dummy variables to identify those stocks in the top and bottom deciles when sorted by their tail risk premia. The regression is specified in Equation 12.

$$
\begin{aligned}
R_{i,t+n} = \gamma_{0,t+n} + \gamma_{1,t+n} \times TRP_{i,t}^{Monthly} \times I_{[Decile\ 1\ TRP_{i,t}^{Monthly}]} + \gamma_{2,t+n} \times TRP_{i,t}^{Monthly} \\
\times I_{[Decile\ 10\ TRP_{i,t}^{Monthly}]} + \phi_{t+n}^{'} \times Z_{i,t} + \varepsilon_{i,t+n}
\end{aligned}
\tag{12}
$$

where $I_{[Decile\ 1\ TRP_{i,t}^{Monthly}]}$ is a dummy variable that equals 1 if $TRP_{i,t}^{Monthly}$ is in Decile 1 and equals 0 otherwise, and $I_{[Decile\ 10\ TRP_{i,t}^{Monthly}]}$ is the corresponding dummy variable for Decile 10. $R_{i,t+n}$ is monthly stock return for stock i in month $t+n$, where $N = 1,2$. $TRP_{i,t}^{Monthly}$ is the individual stock tail risk premium calculated in Equation 10. $Z_{i,t}$ represents a vector of characteristics and controls for firm i at the end of month t such as size, B/M ratio, market beta, illiquidity, etc.

[Insert Table 5 Here]

---

[29] See Appendix for variable definitions.

Table 5 presents the results. In the t+1 regression, the coefficient on the interacted variable $TRP_{i,t}^{Monthly} \times I_{[Decile\ 10\ TRP_{i,t}^{Monthly}]}$ is statistically significant with γ coefficient of -1.971 and Newey-West (1986) t-statistic of -5.09. The negative coefficient implies that the greater the tail risk premium today (i.e., the greater the negative tail risk) the more negative the next month's return will be. The coefficient on the Decile 1 interacted variable is also negative with a γ coefficient of -0.830, but it is statistically significant at only the 10% level; thus, we refrain from making any claims about the premium for bearing positive tail risk's ability to impact future returns. In the t+2 regression, neither the Decile 1 nor the Decile 10 interacted dummies are significant different from zero.

3.4.2.2 The Daily Predictive Power of Positive versus Negative Tail Risk Premia

In order to more fully examine the relationship between negative tail risk premia and future returns, we replicate the study conducted in the previous section using daily returns. Specifically, we conduct a firm-level cross-sectional predictive regression as in Equation 13. We add the caveat that, at the daily level, there is likely to be some noise in our estimates; thus, any conclusions should be tempered somewhat.

$$R_{i,t+n}^{Daily} = \gamma_{0,t+n} + \gamma_{1,t+n} \times TRP_{i,t}^{Daily} \times I_{[Decile\ 10\ TRP_{i,t}^{Daily}]} + \phi_{t+n}' \times Z_{i,t} + \varepsilon_{i,t+n} \qquad (13)$$

[Insert Table 6 Here]

Panel 1 of Table 6 presents the results of the regression of the relationship in Equation 13, day into the future. The coefficient on the interacted variable for Decile 10 shows that the existence of a premium for bearing negative tail risk continues to have predictive power for about 10 days. Even on Day 10, the coefficient for the interacted variable for Decile 10 is -5.910 with NW t-statistic -2.68. The daily results corroborate the findings of the previous section and offer additional evidence on the way that concerns about extreme negative tail events impact future returns.

3.4.3 Do Larger Tail Risk Premia Have More Predictive Power?

The results in the previous section suggest that the larger the tail risk premium, the greater its impact on future returns. This suggests that using a finer grid to sort stocks, for example, sorting the stock by tail risk premia into percentiles rather than deciles, and then redoing the earlier analysis may yield interesting results. To this end, we sort and then split the stocks contained in Deciles 1 and 10 into deciles once again; that is, we effectively create 10 extreme high and low *percentile* portfolios, with percentiles 1-10 belonging to Decile Portfolio 1 and percentiles 91-100 belonging to Decile Portfolio 10. We then apply the 1/0/1 (sort in one month, examine the one-month return for the following month) portfolio strategy for percentile 1 and 100 portfolios, 2 and 99 portfolios, and 3 and 98 portfolios, and report the results in Table 7.

[Insert Table 7 Here]

The results are consistent with the notion that there is a monotonic relationship between the magnitude of the tail risk premium and the impact on future returns. The t+1 return difference is most negative when comparing the two most extreme (1 and 100) portfolios and decreases monotonically thereafter. The return difference (in percentage) for the 100-1 portfolio is -3.91, for the 99-2 portfolio is -2.50, and for the 98-3 portfolio is -2.44, respectively. All are significant at the 1% level.

As a robustness check, we combine percentiles 98-100 into one portfolio and percentiles 1-3 into another portfolio and then do the same for percentiles 97-99 and 2-4. The t+1 return difference for the (98-100)-(1-3) portfolio is -2.95 and for the (97-99)-(2-4) portfolio is -2.35, with both being significant at the 1% level, once again lending support to the notion that predictive power of the current premium for bearing negative tail risk should be directly related to its magnitude.

3.5   Robustness Checks

We perform a variety of robustness checks in order to ensure that our results are not being driven by other factors.

3.5.1   Monte Carlo Analysis of Regression Beta

The first robustness check is on the beta of the baseline regression model in Equation 7. There may be a concern that the beta may not be statistically different from zero both cross-sectionally and in the time series. The standard unidimensional t-test cannot capture this possibility. Instead, to capture both the time series and cross-sectional properties of the regression beta, we use Monte Carlo simulation to test whether beta is statistically different from zero. Monte Carlo simulation has two advantages. First, it is a distributional-free approach. Second, it allows us to make statistical inferences on both the cross-sectional and time-series dynamics of the regression beta in Equation 7.

In our sample, the number of firms that have more than 17 trading days in a given month ranges from 3626 to 7471. We denote sample size as S and number of random draws as N. For a given month in a given year, we perform the following simulation,

1) Random draw (with placement) $\beta_1, \beta_2, \beta_3, \cdots, \beta_S$ and compute the mean of $\beta_1, \beta_2, \beta_3, \cdots, \beta_S$, denote $\overline{\beta_n}$.
2) Repeat 1) N times and get $\overline{\beta_1}, \overline{\beta_2}, \overline{\beta_3}, \cdots, \overline{\beta_N}$.
3) Compute t-statistic for $\overline{\beta_1}, \overline{\beta_2}, \overline{\beta_3}, \cdots, \overline{\beta_N}$.

We then compute the average of the (time series) year-month t-statistic to get the simulated t-statistics.

[Insert Plot 1 Here]

From Panels A through B in Plot 1, we observe that the bootstrapped t-statistic is statistically and significantly different from zero even if we limit the sample size to only 500 firms

in each independent random draw[30]. This indicates that the beta of the baseline regression model in Equation 7 is both statistically and economically important and that it carries important pricing information. In other words, our tail risk premium estimation methodology indeed captures the difference between the $\mathbb{P}$- and $\mathbb{Q}$- spaced expectations of tail risk variation.

3.5.2 Sensitivity to Market Aggregate Tail Risk Premium

The second robustness check is to ensure that our results are not being driven by the sensitivity of the individual stock's loadings to the market tail risk premium. To this end, we follow Ang et al. (2006), which adopts a "beta approach." They obtain an individual stock's sensitivity (beta) to innovation in market aggregate volatility (specifically ΔVIX) and then determine whether this beta has predictive power for the next-month's stock returns. We run the following regression model,

$$r_t^i = \beta_0 + \beta_{MKT}^i MKT_t + \beta_{\Delta TRP^{Market}}^i \Delta TRP_t^{Market} + \varepsilon_t^i \tag{14}$$

where MKT is the market excess return and ΔTRP_t^{Market} is estimated tail risk premium for the S&P 500, which is our proxy for innovations in the market aggregate tail risk compensation; that is, factor loading, $\beta_{\Delta TRP^{Market}}^i$, captures the sensitivity of individual stock monthly returns to the change in market aggregate tail risk premium. The results are reported in Table 8.

[Insert Table 8 Here]

The value weighted mean t+1 return for Deciles 1 and 10 are 1.16 and 1.31, respectively. This difference is not statistically significant, which implies that sensitivity to the market aggregate tail risk premium has no predictive power for these stocks. Moreover, it implies that our tail risk premium estimation methodology captures an individual stock's idiosyncratic tail risk premium, which provides pricing information beyond the individual stock's loadings to the market tail risk premium.

3.5.3 Contemporaneous Regression

Our methodology uses a 22-day lag adjustment between the risk neutral and the physical probability space measures. However, in the literature on the variance risk premium normally does not require a lag adjustment for the $\mathbb{Q}$- spaced variables calculation in the baseline regression model in Equation 7. For example, Bollerslev, Tauchen, and Zhou (2009) are among the first to document the variance risk premium's return predictability at the quarterly horizon. They compute the variance risk premium using a relatively conventional approach, where the risk premium of return variation is defined as the difference between the time series conditional expected future return variation in the (options based) risk-neutral ($\mathbb{Q}$-spaced) framework and in the physical probability ($\mathbb{P}$-) space in a contemporaneous manner; however, this approach is inherently biased in that it assumes the risk neutral measures are backward-looking.

---

[30] In Plot 1 we limit the number of random draws to 10000. We also perform the Monte Carlo simulation by varying the number of random draws from 1000 to 10000 and the sample size from 500 to 3000; results are similar.

As a robustness check, we employ the non-lagged methodology of Bollerslev, Tauchen, and Zhou (2009) and redo the analysis presented in Section 3.4.1. We sort the stocks into ten equal groups (decile portfolios) by $TRP_i^{Monthly,nolag}$ calculated on the following regression model,

$$\Delta[2(R_{i,t} - r_{i,t}) - r_{i,t}^2] = \alpha_i + \beta_i \cdot \Delta(VIX_t^2 - V_{BKM,t}^C) + \varepsilon_{i,t} \tag{15}$$

The results are reported in Table 9.

[Insert Table 9 Here]

As can be seen by comparing the results of Table 9 to those presented in Table 3, using contemporaneous rather than 22-day lagged $\mathbb{Q}$-spaced measures makes little qualitative difference.

3.6   Potential explanations for the asymmetric way positive and negative tail risk premia impact returns

Tversky and Kahneman (1992) propose the concept of prospect theory, whereby individuals perceive the utility of gains and losses differentially. Barberis and Huang (2008) apply this theory to investor behavior and argue for the existence of a lottery effect, where biases in the probability weighting of investors cause them to overvalue stocks that have a small probability of a large positive return. The lottery effect predicts that positively skewed securities will be overvalued, and thus the existence of a negative premium for bearing positive tail risk today would imply lower returns in the future. Bali, Cakici, and Whitelaw (2011) documents that when stocks are sorted by their monthly returns, those stocks with the maximum (minimum) monthly return tend to have a lower (higher) return in the following month. They interpret these results as support for the lottery effect. We control for the MAX and MIN effect in our regression analysis, and find that predictive power of both the positive and negative tail survive the inclusion of these variables. Moreover, in fact, our results for negative tail risk not only survive the inclusion of the MIN control variable, but are in stark contrast to the predictions of the lottery effect. If biases in investor's probability weightings cause them to overvalue stocks with positive skewness, then it stands to reason that these same biases should also be causing them to undervalue stocks with negative skewness. In which case, the existence of a premium for bearing negative tail risk today, should predict a higher return in the future. However, we find just the opposite: the existence of a premium for bearing negative tail risk predicts lower future returns, not higher ones. Thus, the lottery effect, although consistent with our findings for positively skewed stocks, cannot explain our results regarding negatively skewed ones.

An alternative explanation for the relationship between positive skewness and lower returns is presented in Brunnermeier, Gollier, and Parker (2007). The paper presents a general equilibrium model where individuals optimally balance a bias toward optimism with the real costs of making bad decisions. The result is that investors prefer heterogeneous, under-diversified portfolios that overweight assets with positive skewness, so they can obtain skewed portfolio returns. This preferential weighing scheme, in turn, raises the prices of and lowers the returns of positively skewed assets. However, it is reasonable to expect that, if investors are overweighting

positively skewed stocks in their portfolios, they would also be underweighting negatively skewed ones. If this were the case, then – similar to the predictions of the lottery effect – stocks with negative skewness should be undervalued and thus be associated with higher future returns. Again, the opposite is observed in the data.

Although Barberis and Huang (2008) and Brunnermeier, Gollier, and Parker (2007) offer well accepted explanations for the association between positive tail risk and negative future returns, neither can explain our findings regarding stocks with negative tail risk. A potential explanation, however, may come from psychology literature: unrealistic optimism.

Harris and Guten (1979) and Weinstein (1980, 1982, 1984, 1987, and 1989) document the existence of unrealistic optimism, a phenomenon whereby human beings have an optimistic bias about their personal risk; specifically, they perceive their own future as more optimistic compared to others. People believe that extreme negative future events are less likely to happen to themselves than to the average person, and extreme positive future events are more likely to happen to themselves than to others. In other words, humans believe that negative (positive) tail events have a lower (higher) probability of occurring to themselves than occurring to others. Consequently, when people determine the expected impact of extreme tail events on value, the results are exaggerated and optimistically biased. In addition, the more extreme the event, the greater the exaggeration of reality.

Sharot, Guitart-Masip, Korn, Chowdhury, and Dolan (2012), Sharot, Korn and Dolan (2011) and Sharot, Kanai, Marston, Korn, Rees, and Dolan (2012) provide evidence that the human memory process actually reinforces the distortions associated with unrealistic optimism. People update their beliefs more frequently in response to information that is better than expected compared to information that is worse than expected. In addition, Moutsiana, Garett, Clarke, Lotto, Blakemore, and Sharot (2013) show that humans possess a natural tendency to discount bad news while incorporating good news into beliefs.

Taken as a whole, the literature on unrealistic optimism yields interesting predictions for asset pricing. If unrealistic optimism causes investors to overestimate the likelihood of positive tail events and to simultaneously underestimate the likelihood of negative tail events, then investors will tend to pay too much for securities with exposure to either kind of tail risk. Consequently, the existence of a premium for bearing tail risk, irrespective of whether it is negative or positive tail risk, will be associated with lower future returns as the overpricing is eventually corrected. Moreover, the more extreme the tail events the greater the exaggeration/overpricing will be, and the predictive power of the associated tail risk premium on future returns should be longer lasting. We are not claiming that unrealistic optimism is the only explanation consistent with our empirical results; we merely show that our results are consistent with its predictions.

3.7   Conclusion

Bollerslev, Todorov, and Xu (2015) and Kelly and Jiang (2014) find evidence of pricing for negative tail risk, the first in an index and the second for individual securities. The current chapter extends this literature by introducing a novel methodology to directly calculate the tail risk premium for individual stocks, and then employs this measure to examine the impact of equity tail risk in the cross-section of stock returns. The methodology allows for the explicit examination of positive and negative tail risk, and finds a differential impact on return predictability based upon which type of tail risk is being priced. The current chapter controls for a variety of variables associated with positive and negative return skewness and finds at the monthly level that the existence of a premium for bearing positive tail risk today holds no statistically significant power for predicting future returns. In contrast, its counterpart for bearing negative tail risk does have significant predictive power for predicting future returns. The relationship between the premium for bearing negative tail risk in one-month future returns is negative and significant. The monthly results are confirmed by those at the daily level. The predictive power associated with a premium for bearing negative tail risk lasts for 10 trading days. In addition, an analysis at the daily level yields even deeper insights. The size of the current premium for bearing negative tail risk matters significantly for the prediction of future returns. The larger the premium associated with exposure to negative tail risk, the more negative and longer lasting its impact is on expected future returns. This is the first academic work to establish a link between the magnitude of the tail risk premium and the length of its predictive power.

The chapter discusses several potential explanations for our results including the lottery effect, selective probability weighting, crash risk, and momentum, among others. The chapter introduces the concept of unrealistic optimism, and discusses its consequences for asset pricing. Unrealistic optimism is the well documented psychological phenomenon whereby people believe that extreme negative future events are less likely to happen to themselves than to the average person, and extreme positive future events are more likely to happen to themselves than to the average person; in other words, they are overly optimistic about their prospects. Further, the more extreme and remote the likelihood of the event, the greater the optimistic bias is. The result of this optimistic bias is that investors will tend to pay too much for a security that has extreme positive or negative tail risk, because they will overestimate the likelihood of the extreme positive payoffs and underestimate the likelihood of the extreme negative payoffs. Thus, one would expect the existence of a tail risk premium, regardless of whether it is for bearing positive or negative tail risk, would be associated with lower future returns. Our empirical results are consistent with this finding.

The methodology in this chapter can be easily extended to other asset classes and to investor behavior in different countries, for example, bond markets, foreign exchange markets, and commodity markets in both U.S. and foreign markets. As future research, it would be interesting to investigate how tail risk is priced in these other asset classes, especially in the presence of liquidity risk.

References

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

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets*, 5(1), pp.31-56.

Ang, A., Chen, J. and Xing, Y. (2006). Downside risk. *Review of Financial Studies*, 19(4), pp.1191-1239.

Ang, A., Hodrick, R. J., Xing, Y. and Zhang, X. (2006). The cross-section of volatility and expected returns. *The Journal of Finance*, 61(1), pp.259-299.

Bakshi, G., Kapadia, N. and Madan, D. (2003). Stock return characteristics, skew laws, and the differential pricing of individual equity options. *The Review of Financial Studies*, 16(1), pp.101-143.

Bali, T. G., Cakici, N. and Whitelaw, R. F. (2011). Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics*, 99(2), pp.427-446.

Bali, T. G., Cakici, N. and Whitelaw, R. F. (2014). Hybrid tail risk and expected stock returns: When does the tail wag the dog? *Review of Asset Pricing Studies*, p.rau006.

Barberis, N. and Huang, M. (2008). Stocks as lotteries: The implications of probability weighting for security prices. *The American Economic Review*, 98(5), pp.2066-2100.

Bekaert, G. and Hoerova, M. (2014). The VIX, the variance premium and stock market volatility. *Journal of Econometrics*, 183(2), pp.181-192.

Bollerslev, T., Tauchen, G. and Zhou, H. (2009). Expected stock returns and variance risk premia. *Review of Financial studies*, 22(11), pp.4463-4492.

Bollerslev, T. and Todorov, V. (2011a). Estimation of jump tails. *Econometrica*, 79(6), pp.1727-1783.

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

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

Breeden, D. T., Gibbons, M. R. and Litzenberger, R. H. (1989). Empirical tests of the consumption-oriented CAPM. *The Journal of Finance*, 44(2), pp.231-262.

Brunnermeier, M. K., Gollier, C. and Parker, J. A., (2007). Optimal beliefs, asset prices, and the preference for skewed returns. *American Economic Review*, 97(2), pp.159-165.

Carr, P. and Wu, L. (2008). Variance risk premiums. *The Review of Financial Studies*, 22(3), pp.1311-1341.

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), pp.1059-1100.

Chang, B. Y., Christoffersen, P. and Jacobs, K. (2013). Market skewness risk and the cross section of stock returns. *Journal of Financial Economics*, 107(1), pp.46-68.

Chow, V., Jiang, W. and Li, J. (2014). Does VIX truly measure return volatility? *Handbook of Financial Econometrics, Mathematics, Statistics, and Technology Forthcoming*.

Cremers, M., Halling, M. and Weinbaum, D. (2015). Aggregate jump and volatility risk in the cross-section of stock returns. *The Journal of Finance*, 70(2), pp.577-614.

Demeterfi, K., Derman, E., Michael, K., and Zhou, J. (1999). More than you ever wanted to know about volatility swaps. Quantitative Strategies Research Notes, Goldman Sachs.

Dimson, E. (1979). Risk measurement when shares are subject to infrequent trading. *Journal of Financial Economics*, 7(2), pp.197-226.

Du, J. and Kapadia, N. (2012). Tail and volatility indices from option prices. Unpublished working paper. University of Massachusetts, Amhurst.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), pp.3-56.

Fu, F. (2009). Idiosyncratic risk and the cross-section of expected stock returns. *Journal of Financial Economics*, 91(1), pp.24-37.

Gao, G. P., Gao, P. and Song, Z. (2018). Do hedge funds exploit rare disaster concerns? *The Review of Financial Studies*, 31(7), pp.2650-2692.

Gao, G. P., Lu, X. and Song, Z. (2018). Tail risk concerns everywhere. *Management Science*.

Harris, D. M. and Guten, S. (1979). Health-protective behavior: An exploratory study. *Journal of Health and Social Behavior*, pp.17-29.

Jegadeesh, N. (1990). Evidence of predictable behavior of security returns. *The Journal of Finance*, 45(3), pp.881-898.

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

Jiang, G. J. and Oomen, R. C. (2008). Testing for jumps when asset prices are observed with noise– a "swap variance" approach. *Journal of Econometrics*, 144(2), pp.352-370.

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

Lamont, O. A. (2001). Economic tracking portfolios. *Journal of Econometrics*, 105(1), pp.161-184.

Lehmann, B. N. (1990). Fads, martingales, and market efficiency. *The Quarterly Journal of Economics*, 105(1), pp.1-28.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, pp.13-37.

Lu, Z. and Murray, S. (2017). Bear Beta.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1-2), pp.125-144.

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica: Journal of the Econometric Society*, pp.768-783.

Moutsiana, C., Garrett, N., Clarke, R. C., Lotto, R. B., Blakemore, S. J. and Sharot, T. (2013). Human development of the ability to learn from bad news. *Proceedings of the National Academy of Sciences*, 110(41), pp.16396-16401.

Newey, W. K. and West, K. D. (1986). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix.

Sharot, T., Guitart-Masip, M., Korn, C. W., Chowdhury, R. and Dolan, R. J. (2012). How dopamine enhances an optimism bias in humans. *Current Biology*, 22(16), pp.1477-1481.

Sharot, T., Kanai, R., Marston, D., Korn, C. W., Rees, G. and Dolan, R. J. (2012). Selectively altering belief formation in the human brain. *Proceedings of the National Academy of Sciences*, 109(42), pp.17058-17062.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3), pp.425-442.

Scholes, M. and Williams, J. (1977). Estimating betas from nonsynchronous data. *Journal of Financial Economics*, 5(3), pp.309-327.

Todorov, V. and Tauchen, G. (2011). Volatility jumps. *Journal of Business & Economic Statistics*, 29(3), pp.356-371.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4), pp.297-323.
.
Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, 39(5), p.806.

Weinstein, N. D. (1982). Unrealistic optimism about susceptibility to health problems. *Journal of Behavioral Medicine*, 5(4), pp.441-460.

Weinstein, N. D. (1984). Why it won't happen to me: perceptions of risk factors and susceptibility. *Health Psychology*, 3(5), p.431.

Weinstein, N. D. (1987). Unrealistic optimism about susceptibility to health problems: Conclusions from a community-wide sample. *Journal of Behavioral Medicine*, 10(5), pp.481-500.

Weinstein, N. D. (1989). Optimistic biases about personal risks. *Science*, 246(4935), pp.1232-1234.

Appendix   Variable Definitions

Tail risk premium: We compute tail risk premium as in Equation 10 in Section 3.1.

Log (Size): Following Fama and French (1993), size is computed each June as stock price times number of shares outstanding (in hundreds). Size is measured in hundred thousand. We control for size effect by taking the natural logarithm of Size.

Log (B/M): Following Fama and French (1993), book-to-market is computed as the ratio of book common equity over market capitalization (size). Book common equity is calculated using Compustat's book value of stockholders' equity plus balance-sheet deferred taxes and investment tax credit minus the book value of preferred stock. The ratio is computed as the book common equity at the end of fiscal year over size as the December end of fiscal year end.[31]

Market beta: We follow Scholes and Williams (1977) and Dimson (1979) to address nonsynchronous trading in beta estimation. We run regression including lag, current and lead market risk premium as independent variables as in Equation 16,

$$R_{i,t} - r_{f,t} = \alpha_i + \beta_{1,i}(R_{market,t-1} - r_{f,t-1}) + \beta_{2,i}(R_{market,t} - r_{f,t}) + \beta_{3,i}(R_{market,t+1} - r_{f,t+1}) + \varepsilon_{i,t} \tag{16}$$

where $R_{i,t}$ is return for stock i on day t. $R_{market,t}$ is market return on day t and $r_{f,t}$ is risk free rate on day t. We estimate the above equation for each stock using daily returns within each month. For each month, the market beta is estimated as follows in Equation 17 for each stock i,

$$\widehat{\beta_\iota} = \widehat{\beta_{1,\iota}} + \widehat{\beta_{2,\iota}} + \widehat{\beta_{3,\iota}} \tag{17}$$

Idiosyncratic Volatility: Following Ang, Hodrick, Xing and Zhang (2006), idiosyncratic volatility is calculated as

$$Idiosyncratic\ volatility = \sqrt{var(\varepsilon_{t,t})} \tag{18}$$

where $\varepsilon_{t,t}$ is the error term from the three-factor Fama and French (1993) regression. The regression is estimated monthly with more than 17 daily observations in a month.

Lagged 1-month return: Following Jegadeesh (1990) and Lehmann (1990), we use lagged 1-month return to account for short-term return reversal effect; the reversal variable for each stock in month m is defined as the return on the stock over the previous month, i.e., the return in month $m-1$.

Lagged 12-month return: As Jegadeesh and Titman (1993) documented intermediate-term momentum effect, we use lagged 12-month return to account for momentum effect; it is defined as return $m-12$ for each stock in month m.

Illiquidity: Following Amihud (2002), we compute stock illiquidity for each stock i in each month m as the ratio of the absolute monthly stock return to its dollar trading volume:

$$Illiquidity_{i,m} = \frac{|R_{i,m}|}{|Volumn_m \times Price_m|} \tag{19}$$

Maximum (Minimum) monthly return: Following Bali, Cakici and Whitelaw (2011), we control for maximum (minimum) monthly return for each stock i in month m as the maximum (minimum) daily return within month m.

$$Maximum\ monthly\ return_m = \max\{R_{i,t}\}, t = 1, \cdots, T \tag{20}$$

$$Minimum\ monthly\ return_m = \min\{R_{i,t}\}, t = 1, \cdots, T \tag{21}$$

---

[31] To avoid issues with extreme values, the book-to-market ratios are winsorized at the 1% and 99% levels.

where T is the maximum number of daily observations in month m. These are estimated monthly with more than 17 daily observations in a month.

Log (trading volume): Trading volume is the sum of the trading volumes during that month. We control for size effect by taking natural logarithm of Size.

Price: The price on the last trading date of the month.

Tables and Plots

Table 1. Characteristics of Portfolios Sorted by Tail Risk Premium
Each week, stocks in the CRSP database are ranked by their respective tail risk premium. The equal-weighted characteristics of each quintile are computed over the same week. The procedure is repeated for every month from January 1990 to September 2014. Tail risk premium and illiquidity are in 10^{-6}. Lagged 1-month return, Lagged 12-month return, Maximum monthly return, and Minimum monthly return are in percentages. Log (Size), Log (B/M), Market beta, Idiosyncratic Volatility, Log (trading volume), and Price are in absolute values. See Appendix for variable definitions.

Characteristics of Portfolios Sorted by Tail Risk Premium

Deciles	1	2	3	4	5	6	7	8	9	10
Tail risk premium	-88.01	-4.33	-1.27	-0.44	-0.08	0.16	0.53	1.42	4.53	89.05
Log (Size)	4.00	4.87	5.56	6.10	6.49	6.54	6.14	5.59	4.88	3.95
Log (B/M)	-0.53	-0.59	-0.60	-0.59	-0.57	-0.57	-0.60	-0.61	-0.59	-0.53
Market beta	0.91	0.96	0.91	0.81	0.71	0.74	0.82	0.99	0.94	0.81
Idiosyncratic volatility	6.15	3.73	2.76	2.10	1.66	1.66	2.12	2.77	3.76	6.35
Lagged 1-month return	2.47	2.05	1.97	1.88	1.61	1.48	0.88	0.54	-0.23	-2.44
Lagged 12-month return	0.18	0.96	1.51	1.20	1.50	1.34	1.34	1.27	1.11	0.37
Illiquidity	123.92	28.31	19.43	9.67	7.85	5.32	10.36	15.29	37.74	125.65
Maximum monthly return	15.11	8.93	6.64	5.05	4.04	4.10	5.17	6.81	9.30	16.71
Minimum monthly return	-12.19	-7.59	-5.74	-4.46	-3.60	-3.60	-4.48	-5.71	-7.51	-11.63
Log (trading volume)	9.02	9.31	9.58	9.71	9.80	9.84	9.75	9.60	9.34	8.99
Price	7.92	16.95	28.76	40.00	58.04	48.50	44.57	25.91	16.46	8.13
Number of stocks	13749	14237	13977	13307	12515	12635	13221	14023	14273	13792

Table 2. Pearson Correlation Matrix

Pearson correlation coefficients reported as percent for firm characteristics of all CRSP stocks from January 1990 to September 2014. See Appendix for variable definitions.

	Tail risk premium	Log (Size)	Log (B/M)	Market beta	Idiosyncratic volatility	Lagged 1-month return	Lagged 12-month return	Illiquidity	Maximum monthly return	Minimum monthly return	Log (trading volume)
Tail risk premium	100.00	-0.09	-0.33	0.28	1.56	-2.55	0.26	-0.35	3.74	0.78	0.01
Log (Size)		100.00	-27.87	12.54	-49.70	5.59	4.67	-2.82	-37.25	39.00	76.96
Log (B/M)			100.00	-6.32	4.90	3.91	-7.55	1.62	3.72	-1.22	-28.95
Market beta				100.00	-0.12	0.58	1.29	-0.38	3.44	-2.71	17.09
Idiosyncratic volatility					100.00	-10.41	-5.05	3.27	89.72	-80.77	-19.94
Lagged 1-month return						100.00	0.06	-0.58	-10.29	9.54	2.65
Lagged 12-month return							100.00	-0.20	-4.20	4.19	1.61
Illiquidity								100.00	2.53	-3.26	-2.46
Maximum monthly return									100.00	-60.25	-11.99
Minimum monthly return										100.00	11.32
Log (trading volume)											100.00

Table 3. Portfolio Returns Sorted into Deciles by Tail Risk Premium
We form value-weighted decile portfolios every month by sorting stocks based on tail risk premium in Equation 10. Portfolios are formed every month, based on tail risk premium in Equation 10 computed using daily data over the previous month. Panel A displays the 1/0/1 portfolio strategy (sort in one month, examine the one-month return for the following month); Panel B displays the 1/1/1 portfolio strategy strategy (sort in one month, examine the one-month return starting two months from now); and Panel C displays the 1/2/1 portfolio strategy (sort in one month, examine the one-month return starting three months from now). Portfolio 1 (10) is the portfolio of stocks with the lowest (highest) previous month tail risk premium. The statistics in the columns labeled Mean and Std. Dev. are measured in monthly percentage terms and apply to the total, not excess, and simple returns. Size reports the average log market capitalization for firms within the portfolio, and B/M reports the average book-to-market ratio. The row "10-1" refers to the difference in monthly returns between portfolio 10 and portfolio 1. NW t-stat refers to robust Newey-West (1986) t-stat. Pre-formulation TRP is reported in basis points. The sample period is January 1990 to September 2014.

Panel A: Portfolios Sorted by Tail Risk Premium, 1/0/1 Portfolio Strategy						Factor Loadings (bps)
Rank	**Mean**	**Std. Dev.**	**%Mkt Share**	**Size**	**B/M**	**Pre-Formation TRP**
1	2.02	7.19	1.76%	4.03	0.94	-1.023
2	1.70	6.25	4.45%	4.92	0.79	-0.042
3	1.61	5.31	8.72%	5.53	0.73	-0.013
4	1.49	4.52	14.48%	6.03	0.70	-0.004
5	1.43	3.98	19.92%	6.40	0.70	-0.001
6	1.31	4.10	19.73%	6.39	0.70	0.002
7	1.22	4.71	15.20%	6.07	0.70	0.005
8	1.09	5.48	9.44%	5.56	0.73	0.014
9	1.02	6.47	4.40%	4.93	0.78	0.044
10	0.27	7.39	1.88%	4.03	0.93	1.063
10-1	-1.75					
t-stat	(-11.26)					
NW t-stat	(-7.30)					

						Factor Loadings (bps)
Panel B: Portfolios Sorted by Tail Risk Premium, 1/1/1 Portfolio Strategy						
Rank	**Mean**	**Std. Dev.**	**%Mkt Share**	**Size**	**B/M**	**Pre-Formation *TRP***
1	0.76	7.51	1.76%	4.05	0.94	-0.982
2	0.97	6.36	4.45%	4.93	0.79	-0.042
3	1.11	5.54	8.75%	5.53	0.73	-0.013
4	1.12	4.62	14.46%	6.04	0.70	-0.004
5	1.06	4.13	19.97%	6.40	0.70	-0.001
6	1.06	4.19	19.75%	6.40	0.70	0.002
7	1.06	4.66	15.22%	6.07	0.70	0.005
8	1.06	5.48	9.41%	5.56	0.72	0.014
9	0.97	6.52	4.37%	4.93	0.78	0.044
10	0.84	7.28	1.85%	4.04	0.94	0.997
10-1	0.08					
t-stat	(0.73)					
NW *t*-stat	(0.76)					

Panel C: Portfolios Sorted by Tail Risk Premium, 1/2/1 Portfolio Strategy | **Factor Loadings (bps)**

Rank	Mean	Std. Dev.	%Mkt Share	Size	B/M	Pre-Formation *TRP*
1	0.96	7.58	1.72%	4.06	0.95	-0.934
2	1.24	6.60	4.44%	4.93	0.79	-0.042
3	1.16	5.60	8.78%	5.53	0.73	-0.013
4	1.24	4.70	14.48%	6.04	0.70	-0.004
5	1.16	4.18	19.96%	6.41	0.70	-0.001
6	1.18	4.11	19.79%	6.40	0.70	0.002
7	1.18	4.67	15.21%	6.08	0.70	0.005
8	1.21	5.38	9.43%	5.57	0.73	0.014
9	1.14	6.41	4.37%	4.94	0.78	0.044
10	0.99	7.33	1.82%	4.06	0.94	0.969
10-1	0.03					
t-stat	(0.18)					
NW *t*-stat	(0.18)					

Table 4. Fama-MacBeth Cross-Sectional Regression
Results of a Fama-MacBeth cross-sectional regression of stock returns for the following:

$$R_{i,t+1} = \gamma_{0,t+1} + \gamma_{1,t+1} \times TRP_{i,t}^{Monthly} + \phi'_{t+1} \times Z_{i,t} + \varepsilon_{i,t+1} \tag{11}$$

$R_{i,t+1}$ is monthly stock return for stock i in month $t + 1$. $TRP_{i,t}^{Monthly}$ is individual stock tail risk premium for stock i in month t, calculated in Equation 10. $Z_{i,t}$ represents a vector of characteristics and controls for firm i at the end of month t such as size, B/M ratio, market beta, illiquidity, etc. The control variables are described in the Appendix. The Newey-West (1976) HAC robust t-statistic is reported in parentheses. Specification (1) is univariate regression; Specification (2) is multiple regression adding control variables. The sample period is from January 1990 to September 2014.

	(1)	(2)
Intercept	0.009	0.021
	(2.42)	(6.16)
Tail risk premium	-0.801	-1.155
	(-5.60)	(-5.50)
Log (Size)		0.000
		(0.19)
Log (B/M)		0.003
		(3.52)
Market beta		0.000
		(0.74)
Illiquidity		19.715
		(1.11)
Idiosyncratic volatility		-0.001
		(-1.90)
Lagged 1-month return		0.003
		(0.85)
Lagged 12-month return		-0.001
		(-0.40)
Maximum monthly return		-0.032
		(-2.02)
Minimum monthly return		-0.019
		(-1.06)
Log (trading volume)		-0.001
		(-0.80)
Adjusted R^2 (%)	0.133	4.765

Table 5. Fama-MacBeth Regression including dummy variables for Decile 1 and Decile 10 Tail Risk Premia

Results of a Fama-MacBeth cross-sectional regression of stock returns for the following:

$$R_{i,t+n} = \gamma_{0,t+n} + \gamma_{1,t+n} \times TRP_{i,t}^{Monthly} \times I_{[Decile\ 1\ TRP_{i,t}^{Monthly}]} + \gamma_{2,t+n} \times TRP_{i,t}^{Monthly} \times I_{[Decile\ 10\ TRP_{i,t}^{Monthly}]} + \phi'_{t+n} \times Z_{i,t} + \varepsilon_{i,t+n} \quad (12)$$

Where $I_{[Decile\ 1\ TRP_{i,t}^{Monthly}]}$ is a dummy variable that equals 1 if $TRP_{i,t}^{Monthly}$ is in Decile 1 and equals 0 otherwise, and $I_{[Decile\ 10\ TRP_{i,t}^{Monthly}]}$ is the corresponding dummy variable for Decile 10. $R_{i,t+n}$ is monthly stock return for stock i in month $t+n$, where $N = 1,2$. $TRP_{i,t}^{Monthly}$ is individual stock tail risk premium for stock i in month t, calculated in Equation 10. $Z_{i,t}$ represents a vector of characteristics and controls for firm i at the end of month t such as size, B/M ratio, market beta, illiquidity, etc. The control variables are described in the Appendix. The Newey-West (1976) HAC robust t-statistic is reported in parentheses. Specification (1) is the multiple regression adding control variables for monthly return $R_{i,t+1}$; Specification (2) is the multiple regression adding control variables for monthly return $R_{i,t+2}$. The sample period is from January 1990 to September 2014.

	(1)	(2)
Return	t+1	t+2
Intercept	0.020	0.022
	(5.92)	(6.51)
$TRP_{i,t}^{Monthly} \times I_{[Decile\ 10\ TRP_{i,t}^{Monthly}]}$	-1.971	-0.022
	(-5.09)	(-0.05)
$TRP_{i,t}^{Monthly} \times I_{[Decile\ 1\ TRP_{i,t}^{Monthly}]}$	-0.803	-0.668
	(-1.86)	(-1.39)
Log (Size)	0.000	0.001
	(0.00)	(0.63)
Log (B/M)	0.003	0.003
	(3.46)	(3.19)
Market beta	0.000	0.000
	(0.55)	(0.23)
Illiquidity	21.788	14.994
	(1.73)	(1.33)
Idiosyncratic volatility	-0.002	-0.002
	(-2.73)	(-2.73)
Lagged 1-month return	0.003	0.014
	(0.91)	(4.47)
Lagged 12-month return	-0.002	-0.004
	(-0.97)	(-1.72)
Maximum monthly return	-0.026	0.005
	(-1.98)	(0.40)
Minimum monthly return	-0.025	0.020
	(-1.62)	(1.58)
Log (trading volume)	-0.000	-0.001
	(-0.52)	(-1.41)
Adjusted R^2 (%)	4.063	3.947

Table 6. Fama-MacBeth Regression of Positive and Negative Tails using Daily Returns
Results of daily Fama-MacBeth cross-sectional regression of stock returns for the following:

$$R_{i,t+n}^{Daily} = \gamma_{0,t+n} + \gamma_{1,t+n} \times TRP_{i,t}^{Daily} \times I_{[Decile\ 10\ TRP_{i,t}^{Daily}]} + \phi_{t+n}^{'} \times Z_{i,t} + \varepsilon_{i,t+n} \qquad (13)$$

Where $I_{[Decile\ 10\ TRP_{i,t}^{Daily}]}$ is a dummy variable that equals 1 if $TRP_{i,t}^{Daily}$ is in Decile 10 and equals 0 otherwise. $R_{i,t+n}^{Daily}$ is one-day holding period return for stock i from day $t + n - 1$ to day $t + n$, where $n = 1,2 \cdots, 18$. A maximum of 18 days is used since we require stocks to have a minimum of 18 days to be included in the sample. $TRP_{i,t}^{Daily}$ is individual stock tail risk premium for stock i in month t, calculated in Equation 10. $Z_{i,t}$ represents a vector of characteristics and controls for firm i at the end of month t such as size, B/M ratio, market beta, illiquidity, etc. The control variables are described in the Appendix. The table reports predictive regression using next-month daily return (Day 1, 7, 10, 14, 18) as a dependent variable. The Newey-West (1976) HAC robust t-statistic is reported in parentheses. Decile portfolio sort section, similar to Table 3, reports Decile 10 return minus Decile 1 return difference in basis points, as well as t-statistic and Newey-West (1976) HAC robust t-statistic associated with it (see Table 3 for detailed testing methodology). For ease of reading, significant coefficients are highlighted in bold.

	(1)	(2)	(3)	(4)	(5)
Return	Day 1	Day 7	Day 10	Day 14	Day 18
Intercept	-0.005	0.000	-0.000	0.001	0.001
	(-5.92)	(0.74)	(-0.07)	(2.05)	(1.34)
$TRP_{i,t}^{Monthly} \times I_{[Decile\ 10\ TRP_{i,t}^{Monthly}]}$	**-23.931**	**-5.609**	**-5.910**	0.531	4.511
	(-5.80)	**(-2.34)**	**(-2.68)**	(0.16)	(1.53)
Log (Size)	0.001	-0.000	0.000	-0.000	-0.000
	(2.95)	(-1.07)	(1.68)	(-1.88)	(-0.24)
Log (B/M)	0.001	0.000	0.000	0.000	0.000
	(4.69)	(1.54)	(2.21)	(2.59)	(1.05)
Market beta	0.000	-0.000	-0.000	-0.000	0.000
	(0.30)	(-0.64)	(-0.06)	(-2.26)	(1.05)
Illiquidity	12.796	5.576	12.852	4.756	-1.524
	(3.12)	(1.76)	(2.67)	(1.78)	(-1.47)
Idiosyncratic volatility	0.001	0.000	0.000	0.000	0.000
	(1.62)	(2.15)	(2.02)	(1.85)	(1.52)
Lagged 1-month return	-0.001	-0.000	0.001	0.002	0.000
	(-0.56)	(-0.02)	(0.82)	(2.07)	(0.58)
Lagged 12-month return	-0.001	-0.001	-0.001	-0.000	-0.000
	(-1.11)	(-1.73)	(-1.15)	(-0.03)	(-0.16)
Maximum monthly return	-0.010	-0.004	-0.008	-0.008	-0.002
	(-1.52)	(-0.81)	(-1.95)	(-1.88)	(-0.48)
Minimum monthly return	-0.017	0.002	0.001	0.001	0.001
	(-2.37)	(0.39)	(0.12)	(0.28)	(0.21)
Log (trading volume)	0.000	0.000	-0.000	-0.000	-0.000
	(0.95)	(0.02)	(-0.98)	(-0.49)	(-0.10)
Adjusted R^2 (%)	3.989	2.360	2.115	2.095	1.832

Table 7. Extreme Tail: Percentile Portfolio Returns Sorted by Tail Risk Premium
We form value-weighted percentile portfolios every month by sorting stocks based on tail risk premium in Equation 10. Portfolios are formed every month, based on tail risk premium in Equation 10 computed using daily data over the previous month. The table displays the 1/0/1 portfolio strategy (sort in one month, examine the one-month return for the following month). Portfolio 1 (100) is the portfolio of stocks with the lowest (highest) previous month tail risk premium. The statistics in the columns labeled Mean and Std. Dev. are measured in monthly percentage terms and apply to the total, not excess, simple returns. Size reports the average log market capitalization for firms within the portfolio, and B/M reports the average book-to-market ratio. The row "100-1" refers to the difference in monthly returns between portfolio 100 and portfolio 1; the row "99-2" refers to the difference in monthly returns between portfolio 99 and portfolio 2; and the row "98-3" refers to the difference in monthly returns between portfolio 98 and portfolio 3. The row "98-100 minus 1-3" stands for the difference between mean monthly returns of portfolio 98 through 100 and mean monthly returns of portfolio 1 through 3. The row "97-99 minus 2-4" represents the difference between mean monthly returns of portfolio 97 through 99 and mean monthly returns of portfolio 2 through 4. NW t-stat refers to robust Newey-West (1986) t-stat. Pre-formulation TRP is reported in basis points. The sample period is January 1990 to September 2014.

				Factor Loadings (bps)		
Percentile Portfolio Returns Sorted by Tail Risk Premium						
Rank	**Mean**	**Std. Dev.**	**%Mkt Share**	**Size**	**B/M**	**Pre-Formation *TRP***
1	2.74	9.19	6.18%	3.28	1.10	-8.330
2	2.20	8.54	6.68%	5.53	1.06	-1.072
3	2.14	8.21	6.70%	3.74	1.01	-0.570
4	2.32	7.97	8.25%	3.91	0.97	-0.368
5	1.83	7.81	8.63%	4.04	0.93	-0.264
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
96	0.44	8.39	8.23%	4.02	0.93	0.268
97	0.20	7.91	8.03%	3.89	0.96	0.373
98	-0.30	7.84	7.05%	3.74	1.00	0.572
99	-0.30	8.56	6.58%	3.54	1.04	1.074
100	-1.17	8.81	6.29%	3.26	1.12	8.504
100-1	-3.91					
t-stat	(-8.34)					
NW *t*-stat	(-7.21)					
99-2	-2.50					
t-stat	(-7.14)					
NW *t*-stat	(-7.24)					
98-3	-2.44					
t-stat	(-7.01)					
NW *t*-stat	(-6.35)					
98-100 minus 1-3	-2.95					
t-stat	(-13.02)					
NW *t*-stat	(-9.14)					
97-99 minus 2-4	-2.35					
t-stat	(-12.11)					
NW *t*-stat	(-9.58)					

Table 8. Portfolios Returns Sorted into Deciles by Sensitivity to Market Aggregate Tail Risk Premium

We form value-weighted decile portfolios every month by sorting stocks based on sensitivity to market tail risk premium, β^i_{MKT}, in Equation 14.

$$r^i_t = \beta_0 + \beta^i_{MKT}MKT_t + \beta^i_{\Delta TRP^{Market}}\Delta TRP^{Market}_t + \varepsilon^i_t \qquad (14)$$

Portfolios are formed every month, based on sensitivity to market tail risk premium, β^i_{MKT}, in Equation 14 computed using daily data over the previous month. The table displays the 1/0/1 portfolio strategy (sort in one month, examine the one-month return for the following month). Portfolio 1 (10) is the portfolio of stocks with the lowest (highest) previous month tail risk premium. The statistics in the columns labeled Mean and Std. Dev. are measured in monthly percentage terms and apply to the total, not excess, simple returns. Size reports the average log market capitalization for firms within the portfolio and B/M reports the average book-to-market ratio. The row "10-1" refers to the difference in monthly returns between portfolio 10 and portfolio 1. NW t-stat refers to robust Newey-West (1986) t-stat. Pre-formulation TRP is reported in basis points. The sample period is January 1990 to September 2014.

Rank	Mean	Std. Dev.	%Mkt Share	Size	B/M	Factor Loadings Pre-Formation *TRP*
1	1.16	7.05	2.31%	4.20	0.90	-24.86
2	1.34	5.71	6.73%	5.17	0.76	-10.00
3	1.26	5.01	11.29%	5.65	0.73	-5.73
4	1.35	4.66	14.23%	5.92	0.72	-3.08
5	1.31	4.42	14.98%	6.03	0.72	-1.03
6	1.33	4.37	15.28%	6.00	0.73	0.83
7	1.33	4.58	14.54%	5.95	0.72	2.86
8	1.38	5.06	11.41%	5.68	0.74	5.47
9	1.45	5.83	6.89%	5.20	0.77	9.71
10	1.31	6.84	2.34%	4.20	0.92	24.64
10-1	0.15					
t-stat	(1.08)					
NW t-stat	(0.95)					

Table 9. Portfolios Returns Sorted into Deciles by Contemporaneous Tail Risk Premium

We form value-weighted decile portfolios every month by sorting stocks by tail risk premium estimated based on Equation 15. Portfolios are formed every month, based on tail risk premium in Equation 15 computed using daily data over the previous month. Panel A displays the 1/0/1 portfolio strategy (sort in one month, examine the one-month return for the following month), Panel B displays the 1/1/1 portfolio strategy (sort in one month, examine the one-month return starting two month from now) and Panel C displays the 1/2/1 portfolio strategy (sort in one month, examine the one-month return starting three month from now). Portfolio 1 (10) is the portfolio of stocks with the lowest (highest) previous month tail risk premium. The statistics in the columns labeled Mean and Std. Dev. are measured in monthly percentage terms and apply to the total, not excess, simple returns. Size reports the average log market capitalization for firms within the portfolio, and B/M reports the average book-to-market ratio. The row "10-1" refers to the difference in monthly returns between portfolio 10 and portfolio 1. NW t-stat refers to robust Newey-West (1986) t-stat. Pre-formulation TRP is reported in basis points. The sample period is January 1990 to September 2014.

	Panel A: Portfolios Sorted by Tail Risk Premium, 1/0/1 Portfolio Strategy						**Factor Loadings (bps)**
Rank	**Mean**	**Std. Dev.**	**%Mkt Share**	**Size**	**B/M**		**Pre-Formation TRP**
1	2.00	7.23	1.83%	4.03	0.94		-1.006
2	1.69	6.31	4.46%	4.92	0.79		-0.043
3	1.59	5.23	8.89%	5.54	0.73		-0.013
4	1.52	4.38	14.92%	6.05	0.70		-0.005
5	1.40	4.02	19.56%	6.39	0.70		-0.001
6	1.30	4.14	19.84%	6.39	0.70		0.001
7	1.26	4.75	15.16%	6.07	0.70		0.005
8	1.19	5.44	9.05%	5.55	0.72		0.014
9	0.91	6.46	4.41%	4.92	0.78		0.045
10	0.34	7.45	1.89%	4.03	0.94		1.108
10-1	-1.66						
t-stat	(-9.92)						
NW t-stat	(-6.96)						

Panel B: Portfolios Sorted by Tail Risk Premium, 1/1/1 Portfolio Strategy | **Factor Loadings (bps)**

Rank	Mean	Std. Dev.	%Mkt Share	Size	B/M	Pre-Formation *TRP*
1	0.72	7.60	1.82%	4.04	0.94	-0.973
2	0.89	6.48	4.48%	4.93	0.79	-0.043
3	1.13	5.54	8.91%	5.55	0.73	-0.013
4	1.08	4.72	14.93%	6.06	0.70	-0.004
5	1.06	4.13	19.62%	6.39	0.70	-0.001
6	1.08	4.12	19.85%	6.39	0.70	0.001
7	1.14	4.55	15.14%	6.07	0.70	0.005
8	1.12	5.45	9.01%	5.55	0.72	0.014
9	0.99	6.45	4.36%	4.92	0.78	0.045
10	0.80	7.24	1.87%	4.05	0.94	1.030
10-1	0.08					
t-stat	(0.68)					
NW *t*-stat	(0.73)					

Panel C: Portfolios Sorted by Tail Risk Premium, 1/2/1 Portfolio Strategy **Factor Loadings (bps)**

| Rank	Mean	Std. Dev.	%Mkt Share	Size	B/M	Pre-Formation *TRP*
1	1.03	7.55	1.78%	4.06	0.94	-0.922
2	1.27	6.70	4.46%	4.93	0.79	-0.043
3	1.19	5.62	8.91%	5.55	0.73	-0.013
4	1.23	4.78	14.96%	6.06	0.70	-0.005
5	1.21	4.18	19.64%	6.40	0.70	-0.001
6	1.20	4.07	19.85%	6.40	0.70	0.001
7	1.18	4.62	15.15%	6.08	0.70	0.005
8	1.11	5.30	9.03%	5.56	0.72	0.014
9	1.11	6.37	4.37%	4.92	0.79	0.044
10	0.99	7.31	1.86%	4.06	0.95	1.010
10-1	-0.04					
t-stat	(-0.33)					
NW *t*-stat	(-0.37)					

Plot 1. Monte Carlo Analysis of Regression Beta

The Monte Carlo analysis is performed to test whether the baseline regression (7) beta equal to zero. Denote sample size as S, number of random draws as N. For a given month in a given year, we perform the following,

1) Random draw (with placement) $\beta_1, \beta_2, \beta_3, \cdots, \beta_S$ and compute the mean of $\beta_1, \beta_2, \beta_3, \cdots, \beta_S$, denote $\overline{\beta_n}$.
2) Repeat 1) N times and get $\overline{\beta_1}, \overline{\beta_2}, \overline{\beta_3}, \cdots, \overline{\beta_N}$.
3) Compute t-statistic for $\overline{\beta_1}, \overline{\beta_2}, \overline{\beta_3}, \cdots, \overline{\beta_N}$.

We then compute the average of the (time series) year-month t-statistic to get the simulated t-statistic. The sample period is January 1990 to September 2014. In the following graphs, Panels A through C plot the Monte Carlo simulated regression beta values for different combinations of sample size (S=500, 1000) and number of random draw values (N=10000).

Panel A: Sample Size (S) = 500, Number of Random Draw (N) = 10000. Average t-statistic=3.684.

109
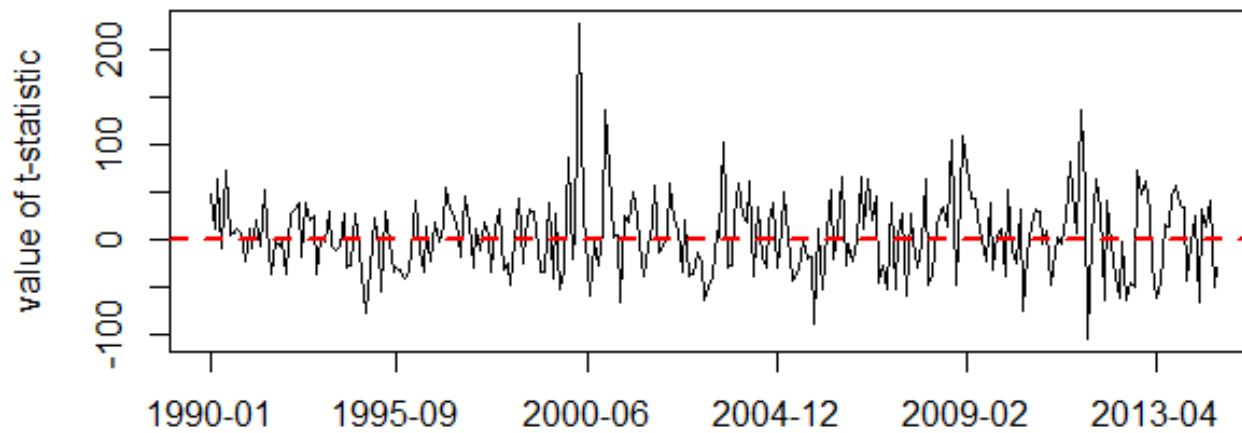
Panel B: Sample Size (S) = 1000, Number of Random Draw (N) = 10000. Average t-statistic=5.152.