A Model-Free Tail Risk Index and Its Return Predictability

Jinji Hao*

Victoria University of Wellington

First Version: August 19, 2016
This Version: August 22, 2017

Abstract

This paper proposes a tail risk index, TIX, as the growth rate of the model-free cumulant generating function of market risk calculated from index option prices. It captures the power law decay rate of the left tail of future return distributions and thus reflects market beliefs about the chance of a market crash. The change in these beliefs strongly predicts market returns both in and out of sample, with monthly $R^2$ statistics of 3.69% and 6.60%, respectively. It can generate utility gains of 9.50% per annum for a mean-variance investor. Evidence at the market and industry levels indicates informational frictions between options and spot markets with about one-third of the information content of this change in market beliefs being delayed to be incorporated into spot prices. A decomposition of VIX into its normal risk and tail risk components shows that only the change in tail risk, which accounts for on average 5% of risks measured by VIX, has return predictability.

Keywords: Tail Risk Index, VIX, SVIX, Return Predictability
JEL Classification Code: G12, G13, G14

*I would like to thank my advisor Werner Ploberger for bringing the stable distribution to my attention which inspired the idea of this paper. I thank Paul Karehnke, Ravi Sastry, Bohui Zhang, Guofu Zhou, and seminar participants at SoFiE 2017, Colorado Finance Summit 2016, Australasian Finance and Banking Conference and PhD forum 2016, University of Otago, Victoria University of Wellington and its Finance Workshop, and Washington University in St. Louis for useful comments and suggestions. All errors are mine. Email: jinji.hao@vuw.ac.nz. Phone: +64 21 388 829.
1 Introduction

Fat tails of stock return distributions have been noticed since at least Mandelbrot (1963) and Fama (1963). This tail risk has been documented in the literature to have significant implications for risk premia.1 But tail risk is typically difficult to measure due to the lack of observations of extreme events. This paper quantifies tail risk by recovering from index option prices exactly the notion of tail fatness advanced by Mandelbrot (1963) and Fama (1963) which is the power law decay rate. The power law in stock returns is among a series of power laws which have been regarded as actually nontrivial and true laws in economics (Gabaix, 2016). Besides this clean interpretation, this tail risk measure also has the advantages that it is forward-looking and it can be easily calculated as a ratio of two statistics at any time for any given horizon.

Moreover, isolating tail risk from aggregate risk measures may provide new insights on the risk-return relationship. By estimating the left jump variation, Bollerslev, Todorov, and Xu (2015) show that it is the tail risk premium component that drives the return predictability of variance risk premium. The tail risk measure proposed in this paper makes it possible to have a closed form decomposition of VIX into its normal risk and tail risk components. Although the tail risk component accounts for on average only 5% of the risk measured by VIX, it is found that the innovation in this tail risk component strongly predicts monthly market and industry portfolio returns but this predictability is lost at the aggregate level of VIX. This reveals a new channel through which tail risk has an implication for asset prices.

I start by showing that the cumulant generating function (cgf), $m(\lambda)$, of market risk, the innovation in log returns, for any horizon can be calculated from index option prices with the corresponding maturity in a model-free manner. I focus on the half piece of cgf on the positive half real line,2 which is a positive, increase, and convex curve starting with zero at the origin. It turns out that the rate at which this cgf of market risk increases reveals information about the rate at which the left tail of future return distributions decays.3 The

---

1Gabaix (2012), Wachter (2013), and Gourio (2008, 2012) show that variation in investors’ concerns regarding rare disasters drives the time-varying equity risk premium. Bollerslev and Todorov (2011) find that fears of rare events account for roughly two-thirds of the total expected excess return and the left tail risks account for 88.4% of the total variance risk premium.

2To avoid assuming the existence of moments of log returns for any order, not just the third or fourth order, only the cgf on the positive half real line is extracted. So I will not take derivatives at 0 to get moments of log returns.

3In a nutshell, the intuition can be illustrated as following. We are probably more familiar with the property that for a standard normal variable $x \sim N(0,1)$, $E[e^{\lambda x}] = e^{\frac{\lambda^2}{2}}$. Indeed, the power 2 in the
tail risk index is defined as the growth rate of the cgf from 1 to 2

\[ \text{TIX} = \log_2 \left( \frac{m(2)}{m(1)} \right). \]  

(1)

Interestingly, when the market risk is normal as in the Black-Scholes setting, TIX equals 2. When the market risk is stable as in Carr and Wu (2003) more generally, the leading example of this paper, TIX equals the stability parameter \( \alpha \in (1, 2) \) of the distribution. Since the stability parameter captures the power law decay rate of the left tail of a stable distribution, it is in this sense that TIX is a tail risk index. A lower TIX indicates a fatter tail of the distribution of market risk and, therefore, a higher tail risk. Economically, TIX reveals the market perceptions of the chance of a market crash.

Although the market risk might not follow a stable distribution exactly, a significant literature has argued that stable distributions describe stock returns very well (Fama, 1963). Carr and Wu (2003) also document empirically that the option pricing with stable distributions delivers superior performance than other popular models such stochastic volatility or jump diffusion models in the literature. Moreover, modeling index returns by a stable distribution might be justified by the generalized central limit theorem which states that the sum of any sequence of independent and identical random variables, a reasonable assumption for index returns driven by random arrival of news at high frequencies, converges to a stable distribution. Therefore, we may expect that TIX still approximately captures the power law decay rate of left tails of market risk and is a tail risk index in general.

The cgf of market risk also provides a unifying framework helping to link TIX to an explicit function of two other leading measures of market risk in the literature, VIX and the SVIX proposed by Martin (2013, 2017).\(^4\) It turns out that VIX just measures (twice) the height of the cgf of market risk at 1, while SVIX measures the convexity of the cgf over the interval \([0, 2]\). The sum of VIX and SVIX measures exactly the height of the function nominator has a particular meaning and can be generalized. It is the power law decay rate of a normal distribution. When the power law decay rate is \( \alpha \) for a more general class of stable distributions, for which normal distributions are special cases, the power in the nominator will be \( \alpha \) instead of 2. Then when comparing the function at different values of \( \lambda \), we have the chance to recover \( \alpha \).

\(^4\)The SVIX in Martin (2013) and Martin (2017) are slightly different. SVIX in Martin (2013), equation (18), is motivated as the strike price for a variance swap contract with the dividend yield taken into account. SVIX in Martin (2017), equation (12), is obtained from replicating the variance of simple return with options which does not take into account the dividend and is not a strike price for a variance swap contract anymore. These two versions of SVIX are equivalent only if the dividend yield is zero. This paper refers to the SVIX in Martin (2013).
at 2. Therefore, for any horizon $\tau$ we have

$$TIX = \log_2 \left( \frac{\log(1 + \tau \cdot SVIX^2) + \tau \cdot VIX^2}{\tau \cdot VIX^2/2} \right).$$

(2)

It should be emphasized that these graphical interpretations hold independent of any assumption regarding the distribution of market risk or its underlying process.

Empirically, using the S&P 500 index option prices, I construct a daily time series of TIX for the 30-day market risk born by the index from 1996:01 to 2016:04 as displayed in Figure 1. It experienced significant drops during the episodes of market crashes such as the mini crash on Oct 27, 1997, the Russian default and LTCM collapse in 1998, the financial tsunami in 2008, the “Flash Crash” on May 6, 2010, the U.S. debt ceiling crisis in 2011, and China’s economic slowdown in late August 2015. The innovations of this TIX time series also display an asymmetry of more frequent and larger drops than rises. It is not surprising that the large drops took place amid these episodes of crisis while the large rises appeared during the economic recoveries.

Since TIX is the market conditional expectations of tail risk, an increase in TIX indicates that the market revises downwards its beliefs about tail risk which may potentially contains fresh and positive information about future market returns. This hypothesis is supported by the superior predictability of the innovation in TIX for future monthly market returns in the sample period 1996:02 to 2016:04 compared with 14 popular predictor variables from Goyal and Welch (2008) and the short interest variable from Rapach, Ringgenberg, and Zhou (2016), which is arguably the best predictor in the literature so far. Among the in-sample predictive regressions, the innovation in TIX has the largest impact with a one-standard-deviation increase in the innovation associated with an 86 basis point increase in next month’s market excess return. Its predictive coefficient is also the most significant at the 1% level and its in-sample $R^2$ statistic is 3.69% much bigger than the second best 2.31% from stock variance and the 2.03% from the short interest variable. For the more challenging out-of-sample test, only the innovation in TIX and the short interest can reduce the mean squared forecast error (MSFE) compared to the mean forecast benchmark, which assumes non-predictability. However, the innovation in TIX reduces the MSFE by 6.60% which is significant at the 5% level for a one-sided test while the short interest reduces it by only 0.76% with a significance level of 10%. Economically, this predictability can generate utility gains of over 9.50% per annum for a mean-variance investor.

\footnote{For their sample period from 1973:01 to 2014:12.}
These tests show that changes in market beliefs about tail risk contain significant additional information about future market returns. But we should expect the market efficiency holds to a large extent so that a substantial proportion of the information has been incorporated into the prices immediately. Indeed, a contemporaneous regression shows that a one-standard-deviation increase in the innovation is associated with a 157 basis point increase in current month’s market excess return, which is highly significant, and the $R^2$ statistic is 12.21%. Therefore, about one-third of the information content of changes in market beliefs about tail risk is delayed to be incorporated into the spot prices. These results are robust for the Fama and French 17 industry portfolio returns. This indicates informational frictions between options and spot markets.

This paper contributes to three strands of literature. First, the TIX proposed in this paper is related to a few fear, rare disaster, or tail indices developed in the literature. Kelly and Jiang (2014) also take a stand on the power law in tails of stock returns and estimate a common tail risk measure from the cross-section of stock returns. But this estimator relies on pooling daily individual stock returns within a month and thus is a historical measure while TIX is forward-looking and can be calculated for any horizon.6 There are also several existing option-based indices. Bollerslev and Todorov (2011) and Bollerslev, Todorov, and Xu (2015) estimate fear indices which are proxies for the special compensation for jump tail risk. Du and Kapadia (2012) construct a jump and tail index as the difference between the risk-neutral variance of holding period log return and VIX which, in a jump diffusion model, is related to jump intensity and higher moments of jump size distribution.7 These indices are interpreted in the context of a jump-diffusion process while TIX focuses directly on the tail distribution of returns for a given horizon. Martin (2017) also studies return distributions but is able to recover the physical probability of a crash from index option prices by assuming a form of pricing kernel. Finally, Manela and Moreira (2016) construct a news implied volatility as a proxy for disaster concerns using phrase counts of front-page articles of the Wall Street Journal.

Second, this paper adds to the literature on the return predictability of option-based measures and several tail risk indices mentioned above. Martin (2013, 2017) theoretically establishes a lower bound of expected market returns based on SVIX and argues that the lower bound is empirically so tight that it can serve as an estimate of expected market

---

6 Another difference is that Kelly and Jiang’s (2014) estimator is under the physical measure while TIX is under the risk-neutral measure.

7 Based on this idea, Gao, Gao, and Song (2016) develop a rare disaster index using only the out-of-money put options written on six economic sector indices.
returns. For the monthly horizon, it achieves an out-of-sample $R^2$ of 0.74% for the full sample, for which the test is feasible since the approach has the advantage of not requiring an initial estimation of parameters, but a negative out-of-sample $R^2$ for the post-financial crisis sample. This paper improves its predictability substantially by showing that the innovation in TIX, which is a function of SVIX and VIX, can achieve a monthly out-of-sample $R^2$ of 6.60% for the post-financial crisis sample. It has also been documented that the variance risk premium (VRP) is able to predict market returns (Bollerslev, Tauchen, and Zhou, 2009; Drechsler and Yaron, 2011). Bollerslev, Todorov, and Xu (2015) estimate the left jump variation (LJV) and find that it can largely drives the predictability of variance risk premium. But the predictability of the innovation in TIX is almost unaffected by controlling either VRP or LJV. Indeed, the correlation between the innovation in TIX and VRP is only 15%. Several tail indices have been documented to have market return predictability for relative longer horizons. Kelly and Jiang’s (2014) tail index is a significant predictor from one year to five years horizons. Du and Kapadia’s (2012) jump and tail index and Manela and Moreira’s (2016) news implied volatility are significant predictors from six months to two years horizons. This paper complements this literature by showing that the innovation in TIX has a superior return predictability for the short one month horizon.

Finally, this paper enriches the literature on developing option implied measures of market risk (Britten-Jones and Neuberger, 2000; Jiang and Tian, 2005; Bakshi, Kapadia, and Madan, 2003). The model-free cgf of market risk provides a framework for studying and developing alternative measures of market risk. Indeed, TIX, VIX, and SVIX are all descriptions of particular shape of the cgf. One may also get model-free moments or cumulants of market risk, if they exist, from taking right derivatives the cgf at zero.

The paper is organized as following. In Section 2, the framework of cgf of market risk is laid down first and illustrated with two examples. Then a model-free extraction of the cgf from index option prices is provided, based on which a new tail risk index, TIX, is defined. Finally, TIX is linked to a function of VIX and SVIX through graphical interpretations of them with the cgf. Section 3 contains the empirical implementation with the S&P 500 index options. The market return predictability of the change in TIX is tested in Section 4.

---

8 Other predictors for stock market returns, individual equity returns or Treasury bill returns include the risk-neutral moments (Conrad, Dittmar, and Ghysels, 2013), the implied volatility smirk or skew (Xing, Zhang, and Zhao, 2010; Ratcliff, 2013), and forward variances (Bakshi, Panayotov, and Skoulakis, 2011).

9 The in-sample monthly $R^2$ is 0.7% and the out-of-sample monthly $R^2$ is 0.3% and insignificant at the 5% level.
Section 5 contains evidence of gradual information incorporation at the index and industry level. Section 6 provides decompositions of VIX and Section 7 concludes.

2 A Generic Framework of Market Risk

In this paper, I extract the market tail risk from equity market index option prices. European option prices are risk-neutral expectations of their final payoffs which are functions of only the future realized price of the underlying. Under the risk-neutral probability measure $Q$, the future spot price $S_T$ of the market index in general can be expressed in an exponential form at time $t$,

$$S_T = S_t e^{(r-q)\tau + \gamma \tau + \sigma Z_{\tau}}, \quad (3)$$

where $\tau = T - t$, and $r$ and $q$ denote, respectively, the continuously compounded risk-free rate and dividend yield, both of which are assumed to be deterministic. The term $\sigma Z_{\tau}$ represents the market risk with $Z_{\tau}$ being a generic mean zero random variable standardized by the size of risk $\sigma$. When the second moment of $Z_{\tau}$ exists, the size of risk $\sigma$ can be defined such that $Z_{\tau}$ has a standard deviation $\tau^{1/2}$. Otherwise, $\sigma$ may be defined to standardize the corresponding notion of “width” of the distribution of $Z_{\tau}$. For example, for the stable distribution which we will discuss in detail later, $\sigma$ can be chosen such that $Z_{\tau}$ has a scale parameter $\tau^{1/2}$. Moreover, in order to have $E_t^*[S_T] = S_t e^{(r-q)\tau}$ which is the forward price of the underlying,\(^{10}\) we need to add a convexity adjustment $\gamma$ such that\(^ {11}\)

$$E_t^*[e^{\gamma \tau + \sigma Z_{\tau}}] = 1. \quad (4)$$

Note that the specification in equation (3) does not rely on any assumption on the stochastic process of the index price.\(^ {12}\) For the class of diffusion processes with or without jumps in either price or volatility, which is widely used in the option pricing literature, the resulting distribution of the future spot price $S_T$ can be fully captured by a random variable $Z_{\tau}$ as well as its normalizing constant $\sigma$ and convexity adjustment $\gamma$. The effects of stochastic volatility and jumps will be reflected in the shape of the distribution of $Z_{\tau}$, including its moments and tail behavior, and the magnitudes of $\sigma$ and $\gamma$. When the price follows a finite moment log stable process as in Carr and Wu (2003), $Z_{\tau}$ is also stable distributed with the same stability and skewness parameters as the innovations.

\(^{10}\)The superscript $*$ is used to indicate the risk-neutral expectation throughout the paper.

\(^{11}\)Whenever without causing confusion, the potential dependence of $r, q, \sigma$, and $\gamma$ on the horizon $\tau$ is suppressed.

\(^{12}\)Except for the deterministic interest rate and dividend yield assumed.
Equation (3) reveals that the moment generating function (mgf) of the market risk \( \sigma Z_T \), denoted as \( M(\lambda) \equiv E_t^*\left[e^{\lambda \sigma Z_T}\right] \), plays a direct role in describing the future variations of the index price \( S_T \) or its simple return \( R_{t,T} \equiv S_T/S_t \). The moments of simple returns are connected to the mgf of \( \sigma Z_T \) evaluated at the power of the moments

\[
E_t^*[R^n_{t,T}] = e^{n(r-q)\tau + n\gamma\tau} E_t^*[e^{n\sigma Z_T}] = e^{n(r-q)\tau + n\gamma\tau} M(n). \tag{5}
\]

For example, the second moment of \( R_{t,T} \) depends on \( M(2) \). By equation (4), the convexity adjustment \( \gamma \) is also related to the mgf of \( \sigma Z_T \) evaluated at 1 by

\[
M(1) = e^{-\gamma \tau}. \tag{6}
\]

The mgf of a random variable is usually considered useful because if the mgf itself is finite in a neighborhood around 0, we can take its derivatives at 0 to get all the moments. However, the existence of mgf of the market risk would be a big assumption. We may expect the mgf \( M(\lambda) \) of the market risk \( \sigma Z_T \) to be finite for \( \lambda > 0 \). First, it will guarantee the forward price of underlying to be finite for non-degenerate size of risk. Equivalently, this will make the convexity adjustment in equation (4) feasible. More importantly, it is necessary for the European option prices, which we rely on, to be finite. However, we should be careful about assuming the finiteness of mgf for any \( \lambda < 0 \) since it would immediately imply the finiteness of mgf in a neighbourhood around 0 and thus the existence of moments of the random variable, \( Z_T \) or log return in our case, of any order, not just the third or fourth order.\(^\text{13}\) This is too strong as even the existence of second moment of log returns might be doubtful for serious econometricians. See Property 1 of mgf in Appendix and more rigorous discussions that follow. Therefore, I propose the following criterion for the market risk \( \sigma Z_T \).

**Assumption 1.** The mgf of \( \sigma Z_T \) is finite on the positive half real line, i.e., \( E_t^*[e^{\lambda \sigma Z_T}] < 0 \) for \( \lambda \in [0, \infty) \).

This assumption adds a restriction on the set of random variables \( Z_T \) that can be used to model the market risk. Property 1.5 implies that the right tail of the distribution of \( Z_T \) has to decay fast enough such that it is exponentially bounded at any rate. On the other

\(^\text{13}\)It should be pointed out that the moments of \( Z_T \) or log returns can still exist even if the mgf of \( \sigma Z_T \) does not exist for \( \lambda < 0 \). Moreover, the existence of moments of log returns is a stronger requirement than the existence of moments of simple returns. This is because the problem with the existence of moments of log returns arises from the fat left tail of its distribution. This problem is alleviated for simple returns which are exponential of log returns.
hand, I allow the mgf of $\sigma Z_\tau$ to be infinite on the negative half real line. This means that
the left tail of the distribution of $Z_\tau$ can decay slowly and its moments do not have to exist.

Property 1.4 says that the mgf of the market risk $\sigma Z_\tau$ increases extremely fast. After
taking logarithm, the cgf increases at a slower rate but is still convex. Since the cgf behaves
less wildly, it is the more appropriate object for investigation. Let $m(\lambda) \equiv \log M(\lambda)$ be the
cgf of the market risk $\sigma Z_\tau$. It necessarily has the following properties.

**Lemma 1.** The cgf $m(\lambda)$ of the market risk $\sigma Z_\tau$ is positive, increasing and convex on
$[0, \infty)$ with $m(0) = 0$.

**Proof.** By the Jensen’s inequality and with the zero mean assumption on $Z_\tau$, $E_t[e^{\lambda \sigma Z_\tau}] >
eq 1$. So $M(\lambda) > 1$ and $m(\lambda) > 0$ for $\lambda > 0$. $m(0) = 0$ by Property 1.1, and it is
convex by Property 1.4. Then $m(\lambda)$ being increasing follows.

The convexity adjustment $\gamma$ depends on the cgf of $\sigma Z_\tau$ as following

**Lemma 2.** The convexity adjustment $\gamma$ is the negative of the annualized cgf of the market
risk $\sigma Z_\tau$ evaluated at 1. It is negative.

**Proof.** By equation (6),

$$ \gamma = -\frac{1}{\tau} \log M(1) = -\frac{1}{\tau} m(1). \quad (7) $$

By Lemma 1, $\gamma < 0$ in general.

It is worth pointing out that the cgf is connected to a dispersion measure of random
variables, entropy, recently employed in the finance literature. The entropy of a random
variable is defined as $L(X) \equiv \log E[X] - E [\log X]$. Alvarez and Jermann (2005), Bansal and
Lehmann (1997), and Backus, Chernov, and Martin (2011) use it to measure the dispersion
of pricing kernels. Martin (2013, Result 2) shows that VIX measures the entropy of simple
returns, which is subsumed as a special case in this lemma.

**Lemma 3.** Under the risk-neutral measure, the cgf of the market risk $\sigma Z_\tau$ evaluated at $n$
is the entropy of the simple return raised to the power $n$, i.e.,

$$ m(n) = L^*(R^n_{t,T}). \quad (8) $$

**Proof.** Taking logarithm of both sides in equation (5), we have

$$ m(n) = \log E_t^*[R^n_{t,T}] - n(r - q + \gamma)\tau. \quad (9) $$

Note that $n(r - q + \gamma)\tau = E_t^*[\log R^n_{t,T}]$, so we have the result.
2.1 The Model-free Cumulant Generating Function of Market Risk

Having established the cgf of the market risk as the object to study, I now show that the whole curve of the cgf of $\sigma Z_\tau$ can be extracted from prices of European index options with a maturity $\tau$ in a model-free manner. This gives us the flexibility to study any feature of the cgf and, therefore, build alternative meaningful and informative measures of market risk.

The intuition behind this model-free extraction is that the power functions can be replicated by European option payoff functions.\(^{14}\) So the moments of $S_T$ or simple return $R_{t,T}$ can be replicated by option prices. Because of the connection between these moments and the mgf of the market risk as shown in equation (5), we are able to extract the mgf and cgf from option prices as well. To extract the convexity adjustment $\gamma$, I apply a similar replication strategy for the log function of the final price as the construction of the Chicago Board Options Exchange (CBOE) VIX.

**Proposition 1.** The cgf of the market risk for any maturity $\tau$ can be extracted from the option prices as

$$m(n) = \log \left[ e^{n(r-q)\tau} + \frac{n(n-1)e^{\tau r}}{S_t^n} \left\{ \int_{F_{t,T}}^{F_{t,T}} K^{n-2} \text{call}_{t,T}(K)dK + \int_{F_{t,T}}^{\infty} K^{n-2} \text{call}_{t,T}(K)dK \right\} \right]$$

$$- n(r-q + \gamma)\tau,$$

where $F_{t,T}$ is the futures price of the underlying and call$_{t,T}(K)$ and put$_{t,T}(K)$ are the prices of European call and put options with a strike $K$, respectively, at time $t$ for an expiry date $T$. Note that $\gamma$ itself is replicated by European option prices as

$$\gamma = - \frac{e^{\tau r}}{\tau} \left\{ \int_0^{F_{t,T}} \frac{1}{K^2} \text{put}_{t,T}(K)dK + \int_{F_{t,T}}^{\infty} \frac{1}{K^2} \text{call}_{t,T}(K)dK \right\}. \tag{11}$$

**Proof.** See the Appendix. \hfill \Box

We have already obtained the model-free extractions of the cgf of the market risk $\sigma Z_\tau$ and its convexity adjustment $\gamma$. There is one parameter remaining to be extracted from the option prices, the normalizing constant $\sigma$. If the second moment of market risk or log

\(^{14}\)Namely, $x^n = \bar{x}^n + n\bar{x}^{n-1}(x-\bar{x}) + n(n-1)\left\{ \int_{\bar{x}}^{x} k^{n-2} \max(k-x,0)dk + \int_{x}^{\infty} k^{n-2} \max(x-k,0)dk \right\}$ for any $\bar{x} > 0$.\n
9
return exists and \( \sigma \) is chosen to be the standard deviation, then \( \sigma^2 = \text{Var}^*[\log(S_T/S_t)] \). The estimator developed by Bakshi, Kapadia, and Madan (2003) can be used to obtain \( \sigma^2 \). If the second moment of market risk or log return does not exist, we need alternative estimators for the corresponding notions of \( \sigma \), which is not the focus of this paper.

### 2.2 Illustrating Examples

I illustrate the framework laid down so far with two examples. The leading example is a finite moment log stable process for the index price as in Carr and Wu (2003) where the market risk is stable distributed. It subsumes the diffusion process as a special case. The second example is the jump diffusion process of Merton (1976) for which, although we do not have an explicit expression for the distribution of market risk, we can obtain its cgf analytically. To save space, I defer the second example to Appendix and discuss the first one in more detail here.

The family of stable distributions is described by a stability parameter or tail index \( \alpha \in (0, 2] \), a skewness parameter \( \beta \in [-1, 1] \), a scale parameter and a location parameter.\(^{15}\) The family of stable distributions subsumes the normal distribution as a special case when \( \alpha = 2 \). It features fat tails when \( \alpha < 2 \), which is a stylized fact for equity returns, so that its second moment does not even exist. To guarantee the option prices are finite, Carr and Wu (2003) propose to use a maximal negative parameter \( \beta = -1 \) such that the right tail of the distribution of log returns decays sufficiently fast. The tail index governs the power law decay rate of the left tail of the distribution.\(^{16}\)

Suppose the equity index price follows a finite moment log stable process under the risk neutral measure,

\[
dS_t/S_t = (r - q)dt + \sigma dL_t^{\alpha,-1}
\]

\(^{15}\)The characteristic function is given by

\[
E \exp(itX) = \begin{cases} 
\exp[it\mu - |ct|^{\alpha}(1 - i\beta \text{sign}(t) \tan \frac{\pi \alpha}{2})] & \alpha \neq 1, \\
\exp[it\mu - |ct|(1 + i\beta \text{sign}(t) \frac{2}{\pi} \log |c|)] & \alpha = 1,
\end{cases}
\]

where \( c \) and \( \mu \) are the scale and location parameters, respectively.

\(^{16}\)The stability parameter governs the power law tail decay rate in the following sense: for \( \alpha \in (0, 2) \)

\[
\begin{align*}
\lim_{x \to \infty} x^\alpha P(X > x) &= C_\alpha \frac{1+\beta}{2} e^\alpha, \\
\lim_{x \to \infty} x^\alpha P(X < -x) &= C_\alpha \frac{1-\beta}{2} e^\alpha,
\end{align*}
\]

where \( C_\alpha \) is a constant depending only on \( \alpha \) (Samorodnitsky and Taqqu, 1994, Property 1.2.15). When \( \beta = -1 \) as in Carr and Wu (2003), the right tail decays faster than a power law. Indeed, if \( \alpha < 1 \), the support of the distribution is only \( (-\infty, \mu] \). But the relevant case for this paper is \( \alpha \in (1, 2) \) so the distribution still has a full support on the real line.
where $dL_t^{\alpha,-1}$ has a stable distribution with tail index $\alpha \in (1,2)$, zero drift, scale of $dt^{1/\alpha}$, and a maximal negative skew parameter $\beta = -1$. $\sigma > 0$ is the scale per unit of time. We can then write the index price at time $T$ in an exponential form

$$S_T = S_t e^{(r-q)\tau + \gamma \tau + \sigma L_t^{\alpha, -1}}$$

(12)

where $\sigma L_t^{\alpha, -1}$ represents the market risk which is also stable distributed with a tail index $\alpha$, due to the stability property of the distribution, and a skewness $-1$. The scale instead of volatility of the log return distribution, $\sigma$, serves as the normalizing constant in our framework. The convexity adjustment $\gamma$ is added such that

$${\mathbb E}_t [e^{\gamma \tau + \sigma L_t^{\alpha, -1}}] = 1.$$ 

By the property of stable distribution (Samorodnitsky and Taqqu, 1994, Property 1.2.4, Proposition 1.2.12)

$${\mathbb E}_t [e^{\sigma L_t^{\alpha, -1}}] = e^{-\tau \sigma^\alpha \sec(\pi \alpha/2)}.$$

So the cgf of the market risk is

$$m(\lambda) = -\tau (\lambda \sigma)^\alpha \sec(\pi \alpha/2).$$

(13)

Hence, we also have an explicit expression for the convexity adjustment

$$\gamma = \sigma^\alpha \sec \frac{\pi \alpha}{2},$$

(14)

which depends on both the normalizing constant $\sigma$ and the tail index $\alpha$. Note that $\alpha \in (1,2)$, so $\gamma$ is negative.\textsuperscript{17}

### 2.3 A Tail Risk Index

Now I propose a tail risk index of the market risk based on its model-free cgf extracted from index option prices. It turns out that the speed at which the cgf increases gives up information about the rate at which the tail of distribution of market risk decays. I define a tail risk index in general as following.

\textsuperscript{17}When $\alpha = 2$, the stable process turns into a diffusion process. But due to the parameterization of the stable distribution adopted here, the standard deviation of the normal distribution is the scale parameter of the stable distribution multiplied by $\sqrt{2}$. By the Itô’s lemma, we would have the convexity adjustment $\gamma = -\sigma^2$, where $\sigma$ is the scale parameter.
Definition 1. The tail risk index, or “TIX”, is defined as logarithm to base 2 of the ratio of the cgf of the market risk $\sigma Z_\tau$ evaluated at 2 over that evaluated at 1, namely

$$TIX = \log_2\left(\frac{m(2)}{m(1)}\right).$$  \hfill (15)

Graphically, TIX is represented in Figure 2 by the ratio of $CE$ over $AB$. The reason I define TIX as the ratio at 2 and 1 is that TIX will be related to the CBOE VIX and the SVIX proposed by Martin (2013, 2017) in a closed form. Why is this a tail risk index? In the previous example of a stable distributed market risk, we obtain the cgf analytically in equation (13). It is interesting that if we take the ratio of the cgf at 2 and 1, the scale parameter $\sigma$ will cancel out with only the stability parameter $\alpha$ left. This tail risk index then captures $\alpha$ exactly.

Proposition 2. If the market risk is stable distributed with a stability parameter $\alpha$, then

$$TIX = \alpha.$$

Proof. By equation (13), we have $m(2)/m(1) = 2^\alpha$.

Since the stability parameter captures the power law decay rate of the left tail of a stable distribution

$$P(X < -x) \sim C_\alpha \times x^{-\alpha} \quad \text{as} \quad x \to +\infty,$$

where $C_\alpha$ is a constant depending only on $\alpha$, it is in this sense that TIX is a tail risk index.

It might be a concern that this insight is valid only for a stable distributed market risk. But there are good reasons that a stable distribution approximates the distribution of market risk very well. First, there is a significant literature going back to at least Mandelbrot (1963) and Fama (1963) arguing that stock returns follow stable distributions. Second, modeling index returns by a stable distribution might be justified by the generalized central limit theorem\footnote{It does not require the finite variance assumption as in the classical central limit theorem where the limit is a normal distribution.} which states that the sum of any sequence of independent and identical random variables, a reasonable assumption for index returns driven by random arrival of news at high frequencies, converges to a stable distribution. Third, Carr and Wu (2003) document empirically that the option pricing with stable distributions delivers superior performance than other more complicated models, such as stochastic volatility or jump diffusion models, in the option pricing literature.
Overall, it is expected that TIX still approximately captures the power law decay rate of the left tail of distribution of market risk and thus is a tail risk index in general. Given the equivalence between the cgf of the market risk and the entropy of the simple returns raised to appropriate powers in Lemma 3, this definition can also be expressed as \( TIX_t = \log_2(L^*(R_t^2)/L^*(R_{t,T})) \).

### 2.4 Graphical Interpretations of TIX, VIX, and SVIX

In this section, I show that the CBOE VIX and the SVIX proposed by Martin (2013, 2017) also have nice graphical interpretations from the cgf of the market risk. See a short introduction to them in Appendix. This makes it possible to find that TIX is indeed a function of VIX and SVIX.

VIX is calculated from option prices as following

\[
VIX_t^2 = \frac{2e^{r\tau}}{\tau} \left\{ \int_0^{F_{t,T}} \frac{1}{K^2} \text{put}_{t,T}(K) dK + \int_{F_{t,T}}^{\infty} \frac{1}{K^2} \text{call}_{t,T}(K) dK \right\}. \tag{16}
\]

Although VIX is designed to be a measure of risk-neutral expectation of future realized variance when initially launched, I will not take this for granted since a diffusion process for the underlying is assumed in the construction of its formula (Demeterfi et al., 1999). From now on I will treat equation (16) as the definition of VIX and see what it really measures in general.

Combining the definition of VIX in equation (16) and the replication of \( \gamma \) in equation (11), we can immediately see that VIX measures the convexity adjustment of the market risk, i.e., \( VIX_t^2 = -2\gamma \). As Martin (2013, Result 2) points out,\(^{19}\) VIX as defined by equation (16) is a measure of the entropy of simple return of the S&P 500 index, i.e., \( VIX_t^2 = \frac{2}{\gamma} L^*(R_{t,T}) \). In our general framework of market risk, these results can be summarized as

**Proposition 3.** Up to a constant scale, the CBOE VIX as defined by (16) is equivalent to the following measures

1. the cgf of the market risk \( \sigma Z_\tau \) at 1;
2. the convexity adjustment of the market risk;
3. the entropy of the market simple return.

\(^{19}\)It is implied by the last two lines in equation (43). Note that \( E_t^*(S_T/S_t) = (r - q)\tau \). So the second last line in equation (43) equals \( 2L^*(R_{t,T}) \).
To be specific,
\[ VIX_t^2 = \frac{2}{\tau} m(1) = -2\gamma = \frac{2}{\tau} \mathcal{L}^*(R_{t,T}). \]  
(17)

Proof. With the discussions above, the results follow by taking into account either Lemma 3 for the case \( n = 1 \) or Lemma 2.

It should be emphasized that these results do not rely on any assumption on the process followed by the index price. As depicted in Figure 2, half of the (unannualized squared) VIX measures exactly the height \( AB \) of the convex cgf of the market risk at 1. As the market risk evolves over time, the curve of its cgf shifts upwards or downwards, in particular, at 1 on the horizontal axis, and VIX fluctuates as well. Therefore, VIX reveals only a piece of information about the market risk that is contained locally by \( m(1) \). We might potentially extract more information about the market risk by looking at the global shape of the curve.

Perhaps the simplest interpretation of VIX is the convexity adjustment \( \gamma \). The convexity adjustment essentially gauges the gap, on a log scale, in the Jensen’s inequality when we take expectation of the convex exponential function of a mean zero market risk. When the market risk is normal, VIX would reveal its standard deviation. Otherwise, VIX would contain information about the higher moments, if any, of market risk (Martin, 2013, Result 2). But the convexity adjustment interpretation is more general and it holds even if these moments do not exist. Using this interpretation of VIX, it is straightforward to obtain what VIX measures when the underlying price follows alternative processes (see VIX formula for a jump diffusion process in Appendix).

Example 1. (Finite Moment Log Stable Process Continued) For this class of processes, we have the convexity adjustment in equation (14). By Proposition 3, we have
\[ VIX_t^2 = -2\sigma^\alpha \sec \frac{\pi\alpha}{2}. \]  
(18)

It can be seen that when the underlying price is logstable VIX combines the tail risk \( \alpha \) and the normal risk \( \sigma \). It is not clear why this specific functional form of combination is meaningful. I will decompose VIX into normal risk and tail risk components later.

Analogous to VIX, Martin (2013, 2017) proposes an index, SVIX, as the annualized strike of a simple variance swap for the S&P 500 index,
\[ SVIX_t^2 = \frac{2e^{\tau T}}{\tau \cdot F_{t,T}^2} \left\{ \int_0^{F_{t,T}} \text{put}_{t,T}(K)dK + \int_{F_{t,T}}^\infty \text{call}_{t,T}(K)dK \right\}. \]  
(19)
From now on equation (19) is treated as the definition of SVIX. Martin (2013, Result 4) further shows that it measures the risk-neutral variance of simple (excess) return of the S&P 500 index$^{20}$

$$SVIX^2_t = \frac{1}{\tau} \text{var}^*_t(R_{t,T}/\bar{R}), \quad (20)$$

where $\bar{R} \equiv e^{(r-q)\tau}$. Within our general framework of market risk, we have the following interpretation of SVIX.

**Proposition 4.** The SVIX measures the change in slope or the convexity of the cdf of the market risk $\sigma Z_\tau$ over the interval $[0, 2]$,

$$\log(1 + \tau \cdot SVIX^2_t) = m(2) - 2m(1) + m(0). \quad (21)$$

**Proof.** Note that $\text{var}^*_t(R_{t,T}/\bar{R}) = E^*_t[R_{t,T}/\bar{R}]^2 - (E^*_t[R_{t,T}/\bar{R}])^2$ and by the specification of price in equation (3) and the correction condition in equation (4)

$$E^*_t[R_{t,T}/\bar{R}] = 1,$$

and

$$E^*_t[R_{t,T}/\bar{R}]^2 = e^{2\gamma \tau} E^*_t[e^{2\sigma Z_\tau}] = e^{2\gamma \tau} M(2).$$

Then by the expression of the convexity adjustment in equation (7)

$$e^{2\gamma \tau} = \frac{1}{M^2(1)}.$$

So

$$SVIX^2_t = \frac{1}{\tau} \left( \frac{M(2)}{M^2(1)} - 1 \right).$$

Taking logarithm of an affine transformation of SVIX and with $m(0) = 0$, we obtain the result. \hfill \Box

As shown graphically in Figure 2, $m(1) - m(0)$ represents the slope of $OB$ or $BD$. $m(2) - m(1)$ represents the slope of $BE$. So $m(2) - 2m(1) + m(0)$ represents the change in slope when we move along the curve of the mgf from $\lambda = 0$ to $\lambda = 2$. In this sense, SVIX measures the convexity of the cdf of the market risk over the interval $[0, 2]$. Alternatively, the difference in the slope between $BD$ and $BE$ or SVIX can be represented by $DE$ in Figure 2. This interpretation of SVIX makes it easy to obtain what SVIX measures when the underlying price follows alternative processes (see SVIX formula for a jump diffusion process in Appendix).

$^{20}$Result 4 in Martin (2013) assumes zero dividend yield but it can be generalized.
Example 2. (Finite Moment Log Stable Process Continued) Based on the cgf of the market risk in equation (13) and by Proposition 4, the SVIX under a finite moment log stable process is given by

$$\log(1 + \tau \cdot SVIX^2_t) = \tau(2 - 2^\alpha)\sigma^\alpha \sec(\pi\alpha/2).$$

(22)

Note that the VIX for a finite moment log stable process is given by equation (18). So in this example VIX and SVIX are related in the following way$^{21}$

$$\log(1 + \tau \cdot SVIX^2_t) = (2^{\alpha-1} - 1) \cdot \tau \cdot VIX^2_t.$$  

(24)

Interestingly, in the same Figure VIX is represented by CD. So the sum of VIX and SVIX measures the height, CE, of the cgf of the market risk at 2,

$$\tau \cdot VIX^2_t + \log(1 + \tau \cdot SVIX^2_t) = m(2).$$

(25)

So we have the following alternative interpretation of TIX as a function of VIX and SVIX.

Proposition 5. The TIX for market risk at any horizon $\tau$ can be equivalently computed as

$$TIX_t = \log_2\left(\frac{\log(1 + \tau \cdot SVIX^2_t) + \tau \cdot VIX^2_t}{\tau \cdot VIX^2_t/2}\right).$$

(26)

3 Data Description

The general framework of market risk proposed in this paper is empirically implemented for the S&P 500 index in the U.S. equity market. The S&P 500 index option data for 1996:01 to 2016:04 used to extract the cgf’s of the market risk is from the OptionMetrics. The zero coupon yield data provided in the OptionMetrics is also used as the risk-free rate for extracting the cgf’s. The dividend yield is backed out from the at-the-money option prices using put-call parity. I follow the CBOE’s practice in constructing VIX as closely

$^{21}$In particular, when $\alpha = 2$, i.e., the price $S_T$ is lognormal, we have

$$\log(1 + \tau \cdot SVIX^2_t) = \tau \cdot VIX^2_t.$$  

(23)

This is the Result 5 obtained in Martin (2013) under the assumption that the stochastic discount factor (SDF) and $S_T$ are conditionally jointly lognormal. Here we obtain the same result even without any assumption on the SDF. When the market risk features fat left tails, i.e., $\alpha \in (1, 2)$, VIX is higher,

$$\tau \cdot VIX^2_t > \log(1 + \tau \cdot SVIX^2_t).$$

16
as possible when recovering the cgf’s and computing TIX.\textsuperscript{22} The detailed procedures are described in Appendix E. To clean the option data, I first apply the filters used by the CBOE. I exclude options that have a zero bid price. Then for call options as we move to successively higher strike prices, once two consecutive call options are found to have zero bid prices, no calls with higher strikes are considered. This rule is applied for put options similarly. Secondly, I delete all replicated entries as in Martin (2013, 2017). Since only the highest bid and the lowest ask are available for each option in OptionMetrics, the mid of bid and ask is used as the option price as in the literature.

To test the market return predictability of the predictor to be constructed in this paper, 15 other popular predictor variables are considered for comparison. 14 of them are studied in Goyal and Welch (2008) and can be obtained from Amit Goyal’s website.\textsuperscript{23} The data on the short interest variable SII proposed in Rapach, Ringgenberg, and Zhou (2016), which is arguably the strongest known predictor of aggregate stock returns, is available from David Rapach’s website.\textsuperscript{24}

\subsection*{3.1 Empirical Cumulant Generating Functions}

In this paper, I focus on the market risk with a horizon of 30 calendar days. This is consistent with the horizon targeted by the CBOE VIX and is also in align with the attempt to examine the return predictability at the monthly horizon later. At the end of each trading day, two cross sections of index option prices, the near-term and the next-term, are obtained with their maturities being the closest to and also bracketing the target of 30 calendar days. Proposition 1 is then applied for the two terms respectively to get their cgf’s. Finally, a linear interpolation is employed to obtain the cgf for the market risk with a horizon of 30 calendar days.\textsuperscript{25} To obtain a continuous curve, I extract the cgf’s for every 0.2 along the horizontal axis starting from the origin. Figure 3 displays the monthly cgf’s from Jan 1996 to April 2016 with each subfigure containing the twelve curves for the year labeled. To save space, the support of these cgf’s is restricted to the interval [0, 4]. To visualize the evolution of these cgf’s, the same vertical scale is used across these subfigures.

There are several immediate observations from this figure. First, these curves are all positive, increasing and convex which is consistent with the theoretical properties of the

\begin{footnotesize}
\begin{enumerate}
  \item See the white paper for VIX on the CBOE’s website.
  \item http://www.hec.unil.ch/agoyal/.
  \item http://sites.slu.edu/rapachde/home/research.
  \item Occasionally, we need to interpolate the 30 calendar days with two next-terms.
\end{enumerate}
\end{footnotesize}
cgf of the market risk established in Lemma 1. Second, there are salient time variations in the shape of these curves. The cgf spikes during the periods of the 1998 Russian default and LTCM collapse, the 2002 burst of dot-com bubble, the 2008-2009 financial crisis, and the 2011 Euro sovereign debt crisis.

3.2 TIX

Given the time series of the cgf of the 30-day market risk, we can easily obtain the time series of TIX by computing the logarithm (to base 2) of the ratio of the height of each curve at 2 over that at 1. Panel A in Figure 1 plots the daily time series of TIX from Jan 4, 1996 to April 29, 2016. The TIX has been below 2 all the time during this period indicating fatter tails of the distributions of market risk than a normal distribution. Note that a lower TIX implies a slower tail decay rate of the distribution of market risk and, hence, a higher tail risk.

Within this sample period, the first significant drop in TIX followed the mini crash in the global stock market on October 27, 1997 caused by the asian financial crisis. The second notable decline in TIX in the second half of 1998 was due to the Russian default and the LTCM collapse. The third and the largest slump in TIX was triggered by the bailout of Fannie and Freddie on September 7, 2008 as well as the takeover, bankruptcy, and bailout of Merrill Lynch, Lehman Brothers, and AIG on September 14, 15, and 16, 2008, respectively. The fourth bottom in TIX was recorded on May 7, 2010 immediately following the “Flash Crash” in the stock market on May 6, 2010 amidst the concerns about the euro sovereign debt crisis. The fifth remarkable decline in TIX happened in the summer of 2011 due to the U.S. debt ceiling crisis and later that year as eurozone fears intensified. Another plunge in TIX was witnessed when the equity markets fell sharply in late August, 2015 propelled by fears that China’s economic slowdown was turning out to be worse than feared. For all these episodes, the TIX fell below 1.9 or even 1.85.

There are also periods of milder declines in TIX. It fell to around 1.92 after September 11 attacks in 2011, after the Internet bubble bursting in 2002, in October, 2014 on concerns about global growth and worries about the spread of the Ebola virus, and at the end of the year due to mounting concerns about the currency crisis in Russia, the turmoil in oil markets and Greek elections.
3.3 Innovations in TIX

To recover the information contained in TIX, I obtain its innovations by taking the log difference of the time series of TIX. To avoid the noise in data at higher frequencies, I construct the monthly innovations in TIX. To be specific, I take the log difference of TIX at the end of two consecutive months. Since TIX is forward-looking with a 30 calendar day horizon, each innovation then tells us how the market updates its beliefs about the following 30-day extreme market risk. The way the market updates its beliefs could potentially reveal the fresh information flows in the market and, therefore, give rise to the return predictability. As TIX increases, the tail of the distribution of market risk gets thinner. So it indicates that the market adjusted downwards the expected extreme risk and the market conditions are improving. Therefore, I conjecture that the innovation in TIX is positively related to the future market returns. Figure 4 plots the non-overlapping monthly time series of innovations in TIX, $d \log TIX$. The series has a mean of -0.01%, a median of 0.07%, a standard deviation of 0.74%, a skewness of -1.17, a kurtosis of 7.40, and a first order autocorrelation of -0.21.

It can be seen that there is an asymmetry in the innovations in the sense that negative innovations of large magnitude are more frequent than positive ones. The largest negative shocks occurred in October 1997 (-2.35%), August 1998 (-2.78%), September and October 2008 (-1.91% and -3.43%), and August 2015 (-3.64%). The largest positive shocks were recorded in November 1998 (2.19%), November 2008 (1.59%), and September 2015 (1.50%).

To relate the information content of innovation in TIX to the large literature on market return predictability, I compare its predictive ability to that of 14 monthly predictor variables from Goyal and Welch (2008), which constitute a set of popular predictors in the literature, and that of the short interest variable SII from Rapach, Ringgenberg, and Zhou (2016), which is arguably the strongest predictor so far. Specifically, I include the following predictors:


2. Log dividend yield (DY): log of a 12-month moving sum of dividends minus the log of lagged stock prices.


7. Net equity expansion (NTIS): ratio of a 12-month moving sum of net equity issues by NYSE-listed stocks to the total end-of-year market capitalization of NYSE stocks.

8. Treasury bill rate (TBL): interest rate on a three-month Treasury bill (secondary market).


14. Inflation (INFL): calculated from the Consumer Price Index (CPI) for all urban consumers.

15. SII: (standardized) detrended log of equal-weighted short interest.

Following the practice in the literature on predicting market returns, I focus on predicting the excess return on a value-weighted market portfolio. I measure the market excess return as the log return on the S&P 500 index minus the log return on a one-month Treasury bill.
3.4 Sample Properties

Table 1 shows the summary statistics for \(d\log\) TIX and 14 popular predictor variables from Goyal and Welch (2008) over the 1996:02 to 2016:04 sample period and for the short interest variable SII from Rapach, Ringgenberg, and Zhou (2016) over the 1996:02 to 2014:12 sample period.

Table 2 displays correlation coefficients for \(d\log\) TIX and the other 15 popular predictor variables in the literature. It can be seen that many of the popular predictors from the literature exhibit strong correlations with each other. The correlation between short interest variable SII and NTIS is -0.51 in this sample. However, the newly proposed predictor \(d\log\) TIX appears largely unrelated to these predictors. The strongest correlations (in magnitude) between \(d\log\) TIX and these popular predictors occur with SVAR and DFR, which have correlations of only -0.24 and 0.22, respectively. The correlation between \(d\log\) TIX and the short interest variable SII is merely -0.01. In other words, the predictor \(d\log\) TIX seems to contain substantially different information from many of the stock return predictors used in the existing literature.

4 Predictability of Market Returns

In this section, I conduct both in-sample and out-of-sample tests of the return predictability of the innovation in TIX. During the sample period 1996:02-2016:04, the annualized equity risk premium is 5.55%, the volatility is 15.43%, and the Sharpe ratio is 0.36.

4.1 In-sample Test

To examine the return predictability of innovation in TIX and to compare it with other popular predictors in the literature, I first run the following in-sample predictive regression for the sample from 1996:02 to 2016:12

\[
rt+1 = \alpha + \beta xt + \epsilon_{t+1},
\]

(27)

where \(rt+1\) is the market excess return in month \(t+1\), \(xt\) is the value of a particular predictor in month \(t\), and \(\epsilon_{t+1}\) is the residual.

To have the comparison of coefficient estimates meaningful, I standardize each predictor to have a standard deviation of one. I also take the negative of SVAR, TBL, LTY, TMS, DFR, and SII before running the regressions for these predictors so that their coefficient
estimates are of a positive sign as the other predictors. I use a heteroskedasticity- and autocorrelation-robust $t$-statistic and compute the a wild bootstrapped $p$ value to test the null hypothesis $\beta = 0$ against the alternative $\beta > 0$ as in Rapach, Ringgenberg, and Zhou (2016). For this sample period, after taking into account the lags, I have 243 observations for estimating equation (27).

Table 3 displays the predictive regression results. The second column shows the coefficient estimates $\hat{\beta}$ for predictors. In brackets under the coefficient estimates are their $t$-statistics as well as their significant levels. The third column presents the in-sample $R^2$ statistics of the OLS regressions. The new predictor $d\log TIX$ outperforms the others in all dimensions. First, its predicting coefficient estimate is the largest, meaning that the innovation in TIX has the largest impact on the predicted market returns. For a one-standard-deviation increase in the innovation in TIX, the predicted monthly market return decreases by 0.86%. It is followed by SVAR with a coefficient estimate of 0.68, SII with a coefficient estimate of 0.64, and DY with a coefficient estimate of 0.62. Second, its $t$-statistic is the highest with a significance level of 1%. Finally, its $R^2$ statistic, 3.69%, is much bigger than the other predictors which achieve at most an $R^2$ statistic of 2.31% by SVAR. The short interest variable SII has an $R^2$ statistic of 2.03%.

4.2 Out-of-sample Test

To check the robustness of the in-sample results and exclude the possibility of overfitting by the innovation in TIX, I look at the out-of-sample predictability of $d\log TIX$ as well as other predictors. To be specific, at the end of month $t$, I predict the market return next month $t + 1$ by

$$\hat{r}_{t+1} = \hat{\alpha}_t + \hat{\beta}_t x_t$$

where $\hat{\alpha}_t$ and $\hat{\beta}_t$ are OLS estimates of $\alpha$ and $\beta$ in equation (27), respectively, based on the data from the beginning of the sample to month $t$. I consider two alternative out-of-sample performance evaluation periods. The first one is 2008:10 to 2016:04 which is after the peak of the Global Finance Crisis in September 2008. The second one is 2007:01 to 2016:04 which includes the whole cycle of the Global Finance Crisis. I use the sample before each out-of-sample period for the initial in-sample estimation of the predictive regression. In particular, the mean forecast, the average excess return from the beginning of the sample to month $t$, serves as the benchmark. This forecast corresponds to the case of $\beta = 0$ in the predictive regression so that it assumes no predictability of market returns. The out-
of-sample $R^2$ statistic, as defined in Campbell and Thompson (2008) as the proportional reduction in the mean squared forecast error (MSFE) of each predicting variable compared to the benchmark of mean forecast, is employed as a measure of out-of-sample predictability. To test if the out-of-sample $R^2$ statistic of a particular predictor variable is significant, I follow Rapach, Ringgenberg, and Zhou (2016) to use the Clark and West (2007) statistic to test the null hypothesis that the MSFE of mean forecast is less than or equal to that of the predictive forecast against the alternative that the MSFE of mean forecast is greater than that of the predictive forecast.

The out-of-sample $R^2$ statistics for the innovation in TIX as well as other popular predictor variables are shown in Table 4. In the second column when the period after the peak of the Global Finance Crisis is used for forecast evaluation, it can be seen that the $R^2$ statistics are negative for 13 out of the 14 variables from Goyal and Welch (2008), meaning they are outperformed by the mean forecast. INFL is an exception but its $R^2$ statistic of 0.51% is still insignificant. The short interest variable SII performs slightly better than the mean forecast with an $R^2$ statistic of 0.76% which is significant at the 10% level. However, the $R^2$ statistic is 6.60% for the innovation in TIX which is much larger and also more significant at the 5% level. Therefore, the innovation in TIX outperforms the mean forecast as well as other popular predictor variables by a large margin. In the third column when the full cycle of the Global Finance Crisis is used for forecast evaluation, we can see only the innovation in TIX can significantly outperform the mean forecast by reducing 5.11% of its MSFE. Although NTIS and INFL are able to reduce the MSFE of mean forecast, they are statistically insignificant. All other predictor variables including SII are outperformed by the mean forecast.

4.3 Utility Gains

Given the evidence of superior return predictability of the change in TIX both in and out of sample, how much can investors benefit from it when making portfolio choice decisions? In this section, I investigate this question from the perspective of a mean-variance investor as in Campbell and Thompson (2008) and Rapach, Ringgenberg, and Zhou (2016). Suppose the investor allocates money between the market portfolio and the risk-free asset and chooses the optimal weight for the market portfolio each period based on the mean and variance forecasts

$$w_t = \frac{1}{\phi} \frac{\hat{r}_{t+1}}{\hat{\sigma}_{t+1}^2}. \quad (29)$$

23
where $\phi$ is the investor’s coefficient of relative risk aversion, $\hat{r}_{t+1}$ is the predictive regression return forecast, and $\hat{\sigma}^2$ is the variance forecast. As in Rapach, Ringgenberg, and Zhou (2016), I set $\phi = 3$.

If this investor repeats this asset allocation strategy during the out-of-sample period, there will be a realized mean return and variance of the portfolio. The mean-variance utility or certainty equivalent return (CER) of this investor then is given by

$$\text{CER} = \bar{r} - \phi \bar{\sigma}^2,$$

where $\bar{r}$ and $\bar{\sigma}^2$ are the mean and variance, respectively, of the portfolio return over the forecast evaluation period. It is the risk-free rate which brings the same utility as the allocation strategy in (29).

Alternatively, the investor may adopt the mean forecast and allocate the portfolio accordingly. This will give the investor a utility from the corresponding realized mean return and variance of the portfolio. We measure the utility gains of the change in TIX as a predictor as the difference between the CER using the change in TIX as a predictor and the CER using the mean forecast.

Table 5 reports the CER gains of a mean-variance investor from using the change in TIX as a predictor of monthly market excess returns. I use the period from 2008:10 to 2016:04 in the second column and the period from 2007:01 to 2016:04 in the third column as the forecast evaluation periods. I also include the Buy and Hold strategy in the last row for comparison. We can see after the peak of the Global Financial Crisis, the investor gains a return of 9.50% per annum from the return predictability of change in TIX. In contrast, SII and the buy and hold strategy generate utility gains of 3.69% and 1.58% per annum, respectively. The 14 predictor variables from Goyal and Welch (2008) generate less utility gains than SII or even utility losses. From the beginning of the Global Financial Crisis, the investor gains a return of 6.34% per annum from the return predictability of change in TIX. The second best is from DFY with utility gains of 3.61% per annum. SII and the buy and hold strategy generate utility gains of 0.93% and 1.00% per annum, respectively.

### 4.4 The Relation to other Option Based Measures as Predictors

There has been a literature on option based risk measures which may have predictability for market returns. Now I examine several of them which are closely related to TIX to see if the return predictability of the innovation in TIX is due to a novel piece of information extracted from option prices.
4.4.1 Martin’s Lower Bound

Martin (2013, 2017) theoretically establishes a lower bound of expected market returns as $\text{LB} = \tau e^{r_T} \cdot \text{SVIX}^2$ under a non-negative correlation assumption. He further argues that the lower bound is empirically so tight that it can serve as an estimate of expected market returns. This paper achieves a similar goal but our approaches are different in three aspects. First, I extract the information from option prices through a tail risk index which uses SVIX as one of two inputs. Second, I focus on the change in market beliefs about tail risk instead of the level of risk itself. Third, I estimate expected market returns by a regression on the proposed predictor while Martin’s lower bound approach has the advantage of not requiring an in-sample estimation of parameters. Given that these are two ways of extracting information about expected market returns from option prices. A natural question would be which approach is more effective for this purpose. This section answers this question by comparing the out-of-sample return predictability of Martin’s lower bound and the innovation in TIX.

The second column in Table 6 reports the out-of-sample $R^2$ statistics in forecasting monthly market returns over the period 2008:10-2016:04 achieved by Martin’s lower bounds as direct estimates and the innovation in TIX by predictive regressions. Two versions of SVIX are considered for robustness when computing the lower bounds. SVIX$_{wq}$ represents the one used in this paper with the dividend yield taken into account. SVIX$_{woq}$ represents the one used in Martin (2017) assuming a lump sum dividend payment but a zero dividend yield. It can be seen that the lower bound constructed with either version of SVIX has an out-of-sample $R^2$ statistic about -1.90%, meaning it is outperformed by the mean benchmark forecast. This is in contrast to the $R^2$ statistic of 6.60% achieved by the innovation in TIX.

Since the lower bound approach does not require an initial estimation of parameters, we now have a chance to consider a longer out-of-sample period 1996:02-2016:04 (column 3). I use the monthly returns in previous two years to calculate the initial mean forecast and then use an expanding window onwards. It can be seen that the lower bounds reduce the MSFE of mean benchmark forecast by about 0.74% but it is statistically insignificant.

4.4.2 Variance Risk Premium and Left Jump Variation

A literature since Bollerslev, Tauchen, and Zhou (2009) has documented that variance risk premium is a significant predictor of market returns. It is defined as the difference between
VIX$^2$ and the realized variance. Given that VIX$^2$ is another input for calculating TIX, now I examine if there is any overlap between the information about market returns contained in variance risk premium and the innovation in TIX.

Table 7 reports the in-sample predictive regression results for variance risk premium and the innovation in TIX. Column 2 and 3 show the results for predictive regressions with a single predictor. As can be seen that variance risk premium is indeed a much stronger return predictor in this sample. Its in-sample $R^2$ statistic is 6.49% which is much higher than that of the innovation in TIX, 3.69%. Its t-statistic for the predictive coefficient, 4.49, is also much larger than that for the innovation in TIX. However, in a predictive regression including both of them as predictors, the results in column 4 show that both variance risk premium and the innovation in TIX are still significant. Both of their predictive coefficients slightly decrease with the t-statistic for the innovation in TIX largely unchanged and that for variance risk premium decreasing from 4.49 to 3.36. Moreover, the $R^2$ statistic is now 8.90% implying that neither of these two predictors is subsumed by another one. This is consistent with the fact that the correlation coefficient between them is merely 15%. Overall, this confirms that the innovation in TIX contains largely different information about market returns from variance risk premium.

5 Informational Frictions

5.1 Industry Evidence

The strong return predictability of innovation in TIX documented in the previous section indicates that changes in market beliefs about tail risk contain information about future market returns. This interpretation implies that these changes in market beliefs should predict returns of other portfolios as well. Moreover, given that TIX is constructed from S&P 500 index options and the market returns I employed so far are the S&P 500 index excess returns, it is worth examining if this return predictability is simply driven by the linkage between these options and their underlying security. For a robustness check, in this section I examine the predictability for industry returns of the change in TIX. If the previous evidence of return predictability carries over to these portfolios, we would be more comfortable to conclude that changes in market beliefs about tail risk contain information about future equity returns in general.

I use the 17 industry portfolio returns from Kenneth French’s website. In particular, these portfolios contain all NYSE, AMEX, and NASDAQ stocks and are not limited to
constituents of the S&P 500 index. I run the in-sample predictive regression (27) and conduct the out-of-sample test for excess returns of each industry.

The second to fourth columns in Table 8 report the in-sample predictive regression results for the 17 industry portfolios as well as the S&P 500 index in the first row for comparison. It can be seen that for 15 out of the 17 industry portfolios the change in TIX is a significant predictor of their excess returns. The predicted excess return that is associated with a one-standard-deviation decrease in tail risk ranges from 41 base points for the Food industry to 203 base points for the Steel industry. These coefficients are significant at least at the 5% level except for the Food industry. The proportion of variation of future monthly industry excess returns that can be explained by changes in TIX ranges from 1.08% for the Food industry to 7.08% for the Chemicals industry. The two industries for which the change in TIX does not have return predictability are Consumption and Utilities.\textsuperscript{26} This is reasonable given that the Consumption and Utilities industries are less sensitive to disaster risk, similar to the Food industry for which I find weaker predictability.

Table 9 reports the out-of-sample test results for industry excess returns. For the same 15 industries, the change in TIX is able to significantly reduce the MSFE compared to the mean benchmark forecasts. The out-of-sample $R^2$ ranges from 1.88% for the Food industry to 10.88% for the Chemicals industry. For the Consumption and Utilities industries, we still do not see return predictability.

Overall, it seems that a decline in market tail risk, although it is extracted from the S&P 500 index options, contains positive information about future equity returns in general.

### 5.2 Contemporaneous regressions

The strong return predictability of innovation in TIX documented in the previous section indicates that the information contained in changes in market beliefs about tail risk has not been fully incorporated into spot prices instantaneously. The return predictability for market and industry returns implies potential informational frictions between options and spot markets. Nevertheless, we should expect that the market efficiency holds to a large extent with a substantial fraction of this information content reflected into market returns in the same period. To check this conjecture, I run contemporaneous regressions

$$r_t = \alpha + \beta x_t + \epsilon_t, \quad t = 1, 2, \ldots, T - 1$$

\textsuperscript{26}Consumption industry includes drugs, soap, perfumes, and tobacco stocks. Utilities includes Utilities stocks.
where $r_t$ is the asset excess return in month $t$, $x_t$ is the innovation in TIX in month $t$, and $\epsilon_t$ is the residual. I consider both market excess return and 17 industry excess returns.

The fourth to sixth columns in Table 8 report the contemporaneous regression results. In the first row we can see that a one-standard-deviation decrease in tail risk is accompanied by a 157 base point increase in the market returns in the same month, which is about twice the magnitude of predicted increase in next month’s returns. As expected, the coefficient estimate is highly significant with a $t$ statistics 4.40, and the $R^2$ is 12.21% which is about four times that in the predictive regression. For industry excess returns, the coefficient estimates are highly significant at the 1% level for all industries. The $R^2$ ranges from 3.74% for the Utilities industry to 9.44% for the Mines industry and 11.81% for the Other. We can see the $\hat{\beta}$ estimates and $R^2$ are much higher in contemporaneous regressions than in predictive regressions except for the Clothes industry, for which their magnitudes are very close.

Figure 5 visualizes the incorporation of the information content of changes in tail risk into market and industry returns in the current month and next month. In the top figure, the blue bar indicates for each industry the increase in predicted return next month $t + 1$ for a one-standard-deviation decrease in tail risk in month $t$. The red bar indicates for each industry the increase in return in month $t$ that is associated with a one-standard-deviation decrease in tail risk in month $t$. In the bottom figure, I plot the faction of information content of changes in tail risk that is delayed to be incorporated into the current market prices. It is defined as the ratio of the blue bar over the sum of blue and red bars for the index and each industry. It shows that 35.48% of the information content is not incorporated into the index prices immediately. Across industries, this fraction ranges from 12.16% for Cnsum industry to 49.71% for Clothes industry. The average across these 17 industries is 37.09%. Overall, except for the Consumptions and Utilities industries this plot is quite flat, implying similar degrees of informational frictions across the industries and more than one-third of the information content of changes in tail risk is delayed to be incorporated into the market prices.

6 Decompositions of VIX

It has been documented in the literature that VIX does not predict market returns. Indeed, the change in VIX does not predict market returns either, in contrast to the strong return predictability of change in TIX found in this paper. In this section, I look deep into the
sources of return predictability by decomposing VIX into different components. The closed form VIX formula for the log stable process as shown in equation (18) provides a clear picture that VIX comprises both normal risk as captured by the scale parameter $\sigma$ and tail risk as captured by the tail parameter $\alpha$. I use VIX($\sigma, \alpha$) to explicitly indicate these two types of risk incorporated in VIX.

There are two ways of decomposing VIX. The first decomposition is conducted to isolate these components. The normal risk component is defined as VIX($\sigma, 2$) so that the tail risk is fixed to be the same as that of a normal distribution and the normal risk $\sigma$ drives the variation of VIX. The tail risk component is then defined as the difference between VIX and its normal risk component. Namely

$$\text{Normal Risk Component} = \sqrt{2\sigma}. \quad (32)$$
$$\text{Tail Risk Component} = \text{VIX} - \sqrt{2\sigma}. \quad (33)$$

Figure 6 plots VIX and its components. The second to fifth columns in Table 10 report their pairwise correlations. It can be seen that VIX, its normal risk and tail risk components, and TIX are highly correlated. The normal risk component (subfigure b) literally has a correlation coefficient 1 with VIX. Moreover, it also accounts for on average 95.28% of the risk measured by VIX (subfigure d). The tail risk component (subfigure c) is of a much smaller magnitude accounting for on average 4.72% of the risk measured by VIX but is also highly correlated with VIX and the normal risk component with correlation coefficients of 85% and 82%, respectively. It is not surprising that the tail risk component has a correlation coefficient -97% with TIX due to the construction. The correlations between TIX and VIX and its normal risk component are relatively lower, -70% and -67%, respectively.

In terms of the information contained in the log change of these risk measures, the sixth to ninth columns in Table 10 report their pairwise correlations, and Table 11 and Table 12 report their in-sample predictive regression and out-of-sample test results, respectively. First, it is found that the changes in tail risk component and TIX are highly correlated with a correlation coefficient of -92%. Moreover, the change in tail risk component is significant in predicting monthly market returns. The predictive coefficient is 0.68 which is significant at the 5% level and the $R^2$ statistic is 2.30%. On the one hand, this confirms that the change in tail risk has a predictive power. On the other hand, this indicates that the accuracy in measuring tail risk is important for capturing the information about future market returns. The information loss due to an imperfect correlation as high as
-91% substantially reduced the predictive power of change in tail risk component to an $R^2$ of 2.22% compared to the 3.69% for the change in TIX. Second, although the change in normal risk component is largely correlated with that in TIX (-0.67%), it is not significant in predicting monthly market return either in or out of sample. Therefore, this decomposition exercise shows that the normal risk component, which is the dominant component of VIX, contaminates the return predictability of the tail risk component such that no evidence of return predictability is found for the change in VIX.

The second decomposition of VIX looks directly at the two constituent terms for log VIX. Taking logarithm of both sides of equation (18), we have

$$\log VIX(\sigma, \alpha) = \frac{\alpha}{2} \log \sigma + \frac{1}{2} \log \left(-2 \sec \frac{\pi \alpha}{2}\right).$$

I call the first term on the right hand side the interaction term since it is the product of normal risk $\sigma$ and tail risk $\alpha$. I call the second term the tail risk term since it depends only on $\alpha$.

Figure 7 plots log VIX and its two constituent terms. The first observation is that the variation in the tail risk term is relatively small (subfigure c). It fluctuates within a band of 0.02. So the variation in log VIX is mostly driven by the interaction term (subfigure b). It is now clear to see that the strong return predictability of change in TIX is lost when it is interacted with the normal risk log $\sigma$ the change in which does not have return predictability.

7 Conclusion

This paper constructs a tail risk index, TIX, by recovering the stability parameter of market risk from index option prices. It is defined as the growth rate of the cgf of market risk on the positive half real line which can be replicated by cross sections of index option prices in a model-free manner. TIX measures the left tail decay rate of the distribution of market risk and thus reveals the market perceptions of disaster risk. Since the cgf of market risk provides a unifying framework for interpreting the two leading forward-looking measures of market risk in the literature, the CBOE VIX and the SVIX proposed by Martin (2013, 2017), TIX is explicitly linked to a function of these two indices.

I construct a daily time series of TIX for the 30-day market risk born by the S&P 500 index. To examine the information content of the innovation in TIX, I test its predictability for future monthly market returns. The innovation in TIX strongly predicts market returns
both in and out of sample, with monthly $R^2$ statistics of 3.33% and 6.15%, respectively, outperforming the popular return predictors in Goyal and Welch (2008) and the short interest variable in Rapach, Ringgenberg, and Zhou (2016). This is also confirmed by evidence using industry excess returns. Decompositions of VIX into its normal risk and tail risk components show that only the innovation in its tail risk component has return predictability.

Contemporaneous regressions at either the index or industry level show that only about two-thirds of the information content of changes in market beliefs about tail risk has been incorporated into the spot prices immediately with another one-third being reflected into prices next month. This indicates informational frictions between the options and spot markets.
Appendix

A Properties of Moment Generating Function

The following properties help to formalize the concept of mgf of the market risk.

Property 1. (Moment Generating Function) Let \( M(t) = E[e^{tX}] \) be the mgf of a random variable \( X \) with a cdf \( F \) and \( m(t) \equiv \log M(t) \) be its cgf, we have

1. \( M(0) = 1 \) and \( m(0) = 0 \).

2. If for some \( t_1 < t_2 \), \( M(t_1) < \infty \) and \( M(t_2) < \infty \), then for any \( t \in [t_1, t_2] \), \( M(t) < \infty \).

3. If for some \( t_1 < 0 < t_2 \), \( M(t_1) < \infty \) and \( M(t_2) < \infty \), then the moments of \( X \) exist for any order, i.e., \( E(X^n) < \infty \) for any \( n > 0 \).

4. (Convexity) On any interval on the real line where \( M(t) \) is finite, it is superconvex, i.e., \( m(t) \) is convex.

5. (Tail Decay Rates) If \( M(t_0) \) is finite for some \( t_0 > 0 \) \((t_0 < 0, \text{ respectively})\), then the right (left, respectively) tail of \( F \) is exponentially bounded,

\[
P(X > x) \leq m(t_0)e^{-t_0x}, \quad (P(X < x) \leq m(t_0)e^{-t_0x}, \text{ respectively}).
\]

(35)

Conversely, if the right (left, respectively) tail of \( F \) is exponentially bounded, i.e., there exists \( C > 0 \) and \( b > 0 \) such that

\[
P(X > x) \leq Ce^{-bx}, \quad (P(X < x) \leq Ce^{bx}, \text{ respectively}).
\]

(36)

then for any \( t \in (0, b) \) \((t \in (-b, 0), \text{ respectively})\), \( M(t) < \infty \).

Properties 1.1 and 1.2 imply that if we assume the mgf \( M(\lambda) \) is finite for a \( \bar{\lambda} > 0 \), then it is finite for all \( \lambda \in [0, \bar{\lambda}] \). It is enough for our purpose if we only consider a bounded range of size of risk \( \sigma \) and a bounded horizon \( \tau \). I choose not to make this restriction and thus assume \( M(\lambda) \) is finite for all \( \lambda > 0 \). Property 1.3 tells us to be careful about assuming the finiteness of mgf for any \( \lambda < 0 \) since it immediately implies the finiteness of mgf in a neighbourhood around 0 and thus the existence of moments of \( Z_\tau \) or log returns of any order.
B CGF of Market Risk for Jump Diffusion Processes

When the underlying price follows a jump diffusion process (Merton, 1976)
\[
dS_t/S_t = (r - q)dt + \delta dW_t + (e^x - 1)dN(t) - \phi E^*[e^x - 1]dt
\]
where \(N(t)\) is a Poisson process with intensity \(\phi\) and \(x\) is the i.i.d. jump size, the final price \(S_T\) can be expressed as
\[
S_T = S_t \exp \left[ (r - q - \frac{1}{2}\delta^2 - k)\tau + \delta W_\tau + \sum_{n=1}^{N(\tau)} x - \phi E^*[x]\tau \right]
\]
where \(k = \phi E^*[e^x - 1 - x]\) and the term labeled as \(\sigma Z_\tau\) represents the mean zero market risk. Clearly the convexity adjustment is
\[
\gamma = -\frac{1}{2}\delta^2 - k.
\]

Although we do not have an explicit expression for the distribution of market risk, its cgf can be obtained as
\[
m(\lambda) = \tau \left[ \frac{1}{2}(\lambda\delta)^2 + \phi E^*[e^{\lambda x} - 1 - \lambda x]\right],
\]
which is the sum of two terms coming from the diffusive and jump components, respectively.

B.1 VIX for Jump Diffusion Process

When the underlying price follows a jump diffusion process, we have the convexity adjustment in equation (38) and, hence, VIX is given by
\[
VIX_t^2 = \delta^2 + 2\phi E^*[e^x - 1 - x].
\]
Therefore, for a jump diffusion process, VIX is a biased estimator for the return variance contributed by diffusive innovations. The error term depends on the second and higher order moments of the jump size distribution as well as the jump intensity.

B.2 SVIX for Jump Diffusion Process

Based on the cgf of the market risk in equation (39) and by Proposition 4, the SVIX under a jump diffusion process can be obtained as
\[
\log(1 + \tau \cdot SVIX_t^2) = \tau \left( \delta^2 + \phi E^*[e^x - 1]^2 \right).
\]
As for VIX, the SVIX depends on both diffusive risk and jump risk. The difference between them is

$$\tau \cdot \text{VIX}^2_t - \log(1 + \tau \cdot \text{SVIX}^2_t) = \tau \phi E^* \left[ 2(e^x - 1 - x) - (e^x - 1)^2 \right] = \tau \phi E^* \left[ -\frac{2}{3} x^3 + o(x^3) \right], \quad (41)$$

which depends mainly on the third moment of the jump size distribution. In particular, if the magnitude of negative jumps is larger than that of positive jumps, which is consistent with the empirical evidence, VIX is higher than SVIX, $\tau \cdot \text{VIX}^2_t > \log(1 + \tau \cdot \text{SVIX}^2_t)$.

C Proofs

C.1 Proof of Proposition 1

Proof. I first replicate the log function of the underlying price at time $T$ by the payoff functions of a bond, a futures contract, and the out-of-the-money put and call options

$$\log \frac{S_T}{S_t} = \log \frac{F_{t,T}}{S_t} + \frac{S_T - F_{t,T}}{F_{t,T}} - \int_{0}^{F_{t,T}} \frac{1}{K^2} \max(K - S_T, 0) dK - \int_{F_{t,T}}^{\infty} \frac{1}{K^2} \max(S_T - K, 0) dK. \quad (42)$$

Note that $F_{t,T}/S_t = e^{(r-q)\tau}$ and $E_t^*(S_T - F_{t,T}) = 0$. Then equation (11) follows by taking the risk-neutral expectation of both sides and employing the risk-neutral pricing of options.

Next, for $n \neq 1$, I can also replicate the power function by the same set of payoff functions

$$S^n_T = F^n_{t,T} + n F_{t,T}^{n-1}(S_T - F_{t,T}) + n(n - 1) \left\{ \int_{0}^{F_{t,T}} K^{n-2} \max(K - S_T, 0) dK \right\}.$$

Dividing both sides by $S^n_t$ and taking risk-neutral expectation of them, we have

$$E_t^*[R^n_{t,T}] = e^{n(r-q)\tau} + \frac{n(n - 1) e^{\tau}}{S^n_t} \left\{ \int_{0}^{F_{t,T}} K^{n-2} \text{put}_{t,T}(K) dK + \int_{F_{t,T}}^{\infty} K^{n-2} \text{call}_{t,T}(K) dK \right\}.$$

Then by equation (9), we have equation (10).

For $n = 1$, recall $m(1) = -\gamma \tau$ in Lemma 2, so equation (10) is true in general. \qed

D A Short Introduction to VIX and SVIX

The CBOE VIX which is often referred to as the “fear index” in the market. The motivation of VIX comes from a variance swap contract which is an agreement at time $t$ on the exchange
of the sum of squared log returns
\[
\left( \log \frac{S_{t+\Delta}}{S_t} \right)^2 + \left( \log \frac{S_{t+2\Delta}}{S_{t+\Delta}} \right)^2 + \cdots + \left( \log \frac{S_T}{S_{T-\Delta}} \right)^2
\]
for a fixed strike price \( V \). Based on Neuberger (1994), under a relatively strong assumption of a diffusion process \( d \log S_t = r - q - \frac{1}{2} \sigma_t^2 + \sigma_t dZ_t \), in the continuous time limit as \( \Delta \) shrinks to zero, the strike price can be heuristically expressed as
\[
V = E_t^* \left[ \int_t^T \sigma_t^2 dt \right]
= 2E_t^* \left[ \int_t^T \frac{1}{S_t} dS_t - \int_t^T d \log S_t \right]
= 2(r - q)\tau - 2E_t^* \log \frac{S_T}{S_t}
= 2e^{r\tau} \left\{ \int_0^{F_{t,T}} \frac{1}{K^2} \text{put}_{t,T}(K) dK + \int_{F_{t,T}}^\infty \frac{1}{K^2} \text{call}_{t,T}(K) dK \right\}.
\]

The last step follows by plugging in the replication of the log payoff in equation (42) in the proof of Proposition 1. Note in particular that this step is true in general independent of the assumption of a diffusion process.

Due to the unlimited risk exposure to extreme price movements, the variance swap market experienced turmoil during the 2008 financial crisis. Martin (2013, 2017) proposes a novel simple variance swap contract which is an agreement on exchange of the sum of squared simple returns instead of log returns
\[
\left( \frac{S_{t+\Delta} - S_t}{F_{t,t}} \right)^2 + \left( \frac{S_{t+2\Delta} - S_{t+\Delta}}{F_{t,t+\Delta}} \right)^2 + \cdots + \left( \frac{S_T - S_{T-\Delta}}{F_{t,T-\Delta}} \right)^2
\]
in which \( F_{t,s} \) is the forward price at time \( t \) of the underlying asset to time \( s \). Under some mild assumptions and, in particular, allowing for jumps in the underlying process, Martin (2013, 2017) shows that in the continuous time limit as \( \Delta \) shrinks to 0, the strike price of the simple variance swap contract can be replicated by options
\[
V = \frac{2e^{r\tau}}{F_{t,T}^2} \left\{ \int_0^{F_{t,T}} \text{put}_{t,T}(K) dK + \int_{F_{t,T}}^\infty \text{call}_{t,T}(K) dK \right\}.
\]

E Constructions of the Cumulant Generating Function of the Generic Market Risk and TIX

I use the option quotes from the OptionMetrics. Although this dataset may be different from the one the CBOE uses to compute VIX, when constructing the cgf and TIX, I try
to follow as closely as possible the practice for constructing the CBOE VIX as described in the VIX white paper.

I clean the data in two ways. First, as the CBOE does for VIX, I exclude options that have a bid price of zero. Moreover, for call options as we move to successively higher strike prices, once two consecutive call options are found to have zero bid prices, no calls with higher strikes are considered. I do the same cleaning for put options as we move to successively lower strike prices. Second, I delete all replicated entries as in Martin (2013, 2017).

E.1 Construction of $\gamma_t$

I first compute $VIX_t^2$ following exactly the same procedure as that for the CBOE VIX. Then we have $\gamma_t = -\frac{VIX_t^2}{2}$.

A special note is that in OptionMetrics the strikes at which we have quotes for call options may not match those for put options. When determining the forward SPX level, $F$, I first obtain the strikes at which we have quotes for both call and put options. Across these strikes, I identify the one at which the absolute difference between the call and put prices is smallest. Then by put-call parity, $F = \text{Strike Price} + \exp(rT) \times (\text{Call Price} - \text{Put Price})$. I then determine $K_0$, the strike price immediately below the forward index level at which we at least have the quote for either the call or put option. If the price for one of the two options is missing, I apply the put-call parity to recover it from the other one. Then the procedure for the CBOE VIX can be followed exactly.

E.2 Construction of cgf for a given maturity $T$

To construct the cgf using option prices as in equation (10), we need to deal with the two terms on the right hand. For the second term, we have already obtained $\gamma$ in the first step. We can also get the risk-free rate minus dividend yield implied by option prices as $r - q = \log(F/S_t)/\tau$.

For the first term, the integration

$$e^{r\tau}\left\{ \int_0^{F_{t,T}} K^{-2n} \text{put}_{t,T}(K)dK + \int_{F_{t,T}}^{\infty} K^{-2n} \text{call}_{t,T}(K)dK \right\}$$

is similar to the integration in the VIX formula (16) except that the weights for options at strike $K_i$ are now $1/K_i^{2-n}$ instead of $1/K_i^2$. So we can follow the same procedure as I did
in the first step for constructing VIX\(^2\). In particular, the first term can be computed as

\[
\log \left( \frac{n(n - 1)}{S_t^n} \sum_i \frac{\Delta K_i}{K_i^{2-n}} e^{rt} Q(K_i) + \frac{K_0^{n-1}}{S_t^n} ((1 - n)K_0 + nF) \right).
\]

(44)

E.3 Construction of cgf for the 30-day maturity \(T_{30}\)

As for the CBOE VIX, I use the near-term and the next-term options to target the 30-day maturity. Specifically, I compute the cgf implied by the near- and next-term options, respectively, following the first two steps above. This will give us the cgf \(m_{\text{near}}(n)\) and \(m_{\text{far}}(n)\) for the maturities \(T_{\text{near}}\) and \(T_{\text{far}}\), respectively. Then I interpolate the cgf of the 30-day market risk by

\[
m(n) = \frac{m_{\text{near}}(n) * (T_{\text{far}} - T_{30}) + m_{\text{far}}(n) * (T_{30} - T_{\text{near}})}{T_{\text{far}} - T_{\text{near}}}. \tag{45}
\]

E.4 Construction of TIX for the 30-day maturity \(T_{30}\)

Having the cgf of the 30-day market risk, I construct TIX as

\[
\text{TIX} = \log_2 \left( \frac{m(2)}{m(1)} \right). \tag{46}
\]
Figure 1: Daily Time Series of TIX and S&P 500 Index: 1996:01-2016:04

Panel A plots the time series of daily TIX. TIX measures the extreme risk of the market risk for next 30 calendar days. It is calculated as $\log_2(m(2/m(1)))$ where $m(\cdot)$ is the cgf of the market risk and is extracted from the option prices of the S&P 500 index. Panel B plots the time series of daily close price of the S&P 500 index.
This figure draws the cumulant generating function $m(\lambda)$ of the market risk $\sigma Z_\tau$. The unannualized VIX, $\tau \cdot \text{VIX}^2$, measures twice the height of $m(\cdot)$ at 1 which is represented by $CD$ in the figure. The transformed unannualized SVIX, $\log(1 + \tau \cdot \text{SVIX}^2)$, measures the difference between the height of $m(\cdot)$ at 2 and two times the height of $m(\cdot)$ at 1 which is represented by $DE$ in the figure. Alternatively, it measures the convexity of the curve, i.e., the change in the slope between $OB$ (or $BD$) and $BE$. TIX measures the ratio of $CE$ over $AB$. 
This figure plots the cgf’s of the 30-day market risk in the S&P 500 index extracted from its option prices. The cgf curves are grouped by year. Each subfigure displays the 12 curves representing the cgf’s at the end of each month within the year.
Panel A displays the monthly time series of innovations in TIX which are computed as the log difference in TIX at the end of each month. Panel B displays the monthly S&P 500 value-weighted excess returns.
The top figure displays the coefficients of in-sample predictive regression (blue bar) and contemporaneous regression (red bar) of index and industry excess returns on the innovation in TIX. The bottom figure plots the proportion of information content of innovations in TIX that is delayed to be incorporated into the spot price, i.e., the ratio of blue bar over the sum of blue bar and red bar.
This figure decomposes VIX based on its formula for a log stable process, \( VIX^2 = -2\sigma^\alpha \sec \frac{\pi \alpha}{2} \). I use \( VIX(\sigma, \alpha) \) to indicate the dependence of VIX on normal risk \( \sigma \) and tail risk \( \alpha \). I define the normal risk component of VIX as \( VIX(\sigma, 2) \) where the tail risk is fixed to be the same as a normal distribution which equals \( \sqrt{2}\sigma \). The tail risk component of VIX is defined as \( VIX - \sqrt{2}\sigma \). VIX and its components are scaled by 100 in this figure. The last subfigure plots the percentage of tail risk component of VIX. VIX and TIX, which equals \( \alpha \), are computed from S&P 500 index options prices in OptionMetrics as shown in the paper.
This figure decomposes log of VIX based on its formula for a log stable process, \( \log VIX(\sigma, \alpha) = \frac{\alpha}{2} \log \sigma + \frac{1}{2} \log(-2 \sec \frac{\pi \alpha}{2}) \). I define the first term on the right hand side as the interaction term between normal risk \( \sigma \) and tail risk \( \alpha \) and the second term as the tail risk term which depends only on \( \alpha \). VIX used in this figure is not scaled by 100. VIX and TIX, which equals \( \alpha \), are computed from S&P 500 index options prices in OptionMetrics as shown in the paper.
## Table 1: Sample statistics, 1996:02-2016:04

<table>
<thead>
<tr>
<th>Predictor</th>
<th>mean</th>
<th>median</th>
<th>1st percentile</th>
<th>99th percentile</th>
<th>std. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>DP</td>
<td>-4.02</td>
<td>-4.01</td>
<td>-4.48</td>
<td>-3.38</td>
<td>0.22</td>
</tr>
<tr>
<td>DY</td>
<td>-4.02</td>
<td>-4.00</td>
<td>-4.48</td>
<td>-3.40</td>
<td>0.22</td>
</tr>
<tr>
<td>EP</td>
<td>-3.16</td>
<td>-3.04</td>
<td>-4.81</td>
<td>-2.66</td>
<td>0.40</td>
</tr>
<tr>
<td>DE</td>
<td>-0.87</td>
<td>-0.99</td>
<td>-1.24</td>
<td>1.28</td>
<td>0.45</td>
</tr>
<tr>
<td>SVAR</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>BM</td>
<td>0.27</td>
<td>0.28</td>
<td>0.12</td>
<td>0.41</td>
<td>0.08</td>
</tr>
<tr>
<td>NTIS</td>
<td>0.00</td>
<td>0.01</td>
<td>-0.05</td>
<td>0.03</td>
<td>0.02</td>
</tr>
<tr>
<td>TBL (%)</td>
<td>2.30</td>
<td>1.65</td>
<td>0.01</td>
<td>6.09</td>
<td>2.15</td>
</tr>
<tr>
<td>LTY (%)</td>
<td>4.67</td>
<td>4.80</td>
<td>2.17</td>
<td>7.17</td>
<td>1.35</td>
</tr>
<tr>
<td>LTR (%)</td>
<td>0.63</td>
<td>0.77</td>
<td>-6.72</td>
<td>8.69</td>
<td>3.06</td>
</tr>
<tr>
<td>TMS (%)</td>
<td>2.37</td>
<td>2.45</td>
<td>-0.26</td>
<td>4.43</td>
<td>1.30</td>
</tr>
<tr>
<td>DFY (%)</td>
<td>1.01</td>
<td>0.91</td>
<td>0.55</td>
<td>3.09</td>
<td>0.44</td>
</tr>
<tr>
<td>DFR (%)</td>
<td>-0.00</td>
<td>0.04</td>
<td>-6.25</td>
<td>5.95</td>
<td>1.82</td>
</tr>
<tr>
<td>INFL (%)</td>
<td>0.18</td>
<td>0.19</td>
<td>-1.01</td>
<td>0.98</td>
<td>0.37</td>
</tr>
<tr>
<td>SII</td>
<td>0.00</td>
<td>-0.06</td>
<td>-1.55</td>
<td>2.38</td>
<td>1.00</td>
</tr>
<tr>
<td>$d \log TIX$ (%)</td>
<td>-0.01</td>
<td>0.07</td>
<td>-2.83</td>
<td>1.50</td>
<td>0.74</td>
</tr>
</tbody>
</table>

The database contains 243 monthly observations for February 1996 to April 2016 for all variables except SII for which it contains 227 monthly observations for February 1996 to December 2014. The table displays summary statistics for 14 predictor variables from Goyal and Welch (2008), the short interest variable SII from Rapach, Ringgenberg, and Zhou (2016), and the innovation in $TIX$. DP is the log dividend-price ratio, DY is the log dividend yield, EP is the log earnings-price ratio, DE is the log dividend-payout ratio, SVAR is stock variance, BM is the book-to-market value ratio for the Dow Jones Industrial Average, NTIS is net equity expansion, TBL is the interest rate on a three-month Treasury bill, LTY is the long-term government bond yield, LTR is the return on long-term government bonds, TMS is the long-term government bond yield minus the Treasury bill rate, DFY is the difference between Moody’s BAA- and AAA-rated corporate bond yields, DFR is the long-term corporate bond return minus the long-term government bond return, and INFL is inflation calculated from the CPI for all urban consumers. SII is the (standardized) detrended log of equal-weighted short interest. $d \log TIX$ is the log difference in $TIX$ at the end of each month.
Table 2: Correlation matrix, 1996:02-2016:04

<table>
<thead>
<tr>
<th>Predictor</th>
<th>DP</th>
<th>DY</th>
<th>EP</th>
<th>DE</th>
<th>SVAR</th>
<th>BM</th>
<th>NTIS</th>
<th>TBL</th>
<th>LTY</th>
<th>LTR</th>
<th>TMS</th>
<th>DFY</th>
<th>DFR</th>
<th>INFL</th>
<th>SII</th>
<th>d log TIX</th>
</tr>
</thead>
<tbody>
<tr>
<td>DP</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DY</td>
<td>0.98</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EP</td>
<td>-0.01</td>
<td>-0.01</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DE</td>
<td>0.48</td>
<td>0.48</td>
<td>-0.88</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SVAR</td>
<td>0.29</td>
<td>0.22</td>
<td>-0.28</td>
<td>0.39</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BM</td>
<td>0.70</td>
<td>0.68</td>
<td>0.42</td>
<td>-0.03</td>
<td>0.07</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NTIS</td>
<td>-0.49</td>
<td>-0.46</td>
<td>0.09</td>
<td>-0.31</td>
<td>-0.31</td>
<td>-0.27</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBL</td>
<td>-0.59</td>
<td>-0.58</td>
<td>-0.03</td>
<td>-0.26</td>
<td>-0.10</td>
<td>-0.80</td>
<td>0.27</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LTY</td>
<td>-0.55</td>
<td>-0.54</td>
<td>-0.24</td>
<td>-0.05</td>
<td>-0.02</td>
<td>-0.76</td>
<td>0.47</td>
<td>0.81</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LTR</td>
<td>0.03</td>
<td>0.01</td>
<td>0.04</td>
<td>-0.02</td>
<td>0.18</td>
<td>0.05</td>
<td>0.02</td>
<td>0.00</td>
<td>-0.09</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TMS</td>
<td>0.40</td>
<td>0.40</td>
<td>-0.20</td>
<td>0.37</td>
<td>0.14</td>
<td>0.36</td>
<td>0.04</td>
<td>-0.80</td>
<td>-0.31</td>
<td>-0.10</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DFY</td>
<td>0.62</td>
<td>0.59</td>
<td>-0.51</td>
<td>0.74</td>
<td>0.58</td>
<td>0.34</td>
<td>-0.58</td>
<td>-0.45</td>
<td>-0.36</td>
<td>0.02</td>
<td>0.37</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DFR</td>
<td>0.01</td>
<td>0.11</td>
<td>-0.19</td>
<td>0.17</td>
<td>-0.25</td>
<td>-0.03</td>
<td>0.03</td>
<td>-0.08</td>
<td>-0.01</td>
<td>-0.47</td>
<td>0.11</td>
<td>0.12</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INFL</td>
<td>-0.15</td>
<td>-0.14</td>
<td>0.07</td>
<td>-0.14</td>
<td>-0.24</td>
<td>-0.09</td>
<td>0.05</td>
<td>0.15</td>
<td>0.18</td>
<td>-0.12</td>
<td>-0.06</td>
<td>-0.27</td>
<td>-0.05</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SII</td>
<td>0.20</td>
<td>0.17</td>
<td>-0.05</td>
<td>0.13</td>
<td>0.13</td>
<td>0.07</td>
<td>-0.51</td>
<td>0.07</td>
<td>0.06</td>
<td>0.02</td>
<td>-0.04</td>
<td>0.27</td>
<td>-0.10</td>
<td>0.13</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>d log TIX</td>
<td>-0.02</td>
<td>0.05</td>
<td>-0.05</td>
<td>0.04</td>
<td>-0.24</td>
<td>-0.01</td>
<td>0.00</td>
<td>-0.02</td>
<td>-0.04</td>
<td>0.00</td>
<td>0.06</td>
<td>0.22</td>
<td>0.00</td>
<td>-0.01</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

The table displays correlation coefficients for 14 predictor variables from Goyal and Welch (2008), the short interest variable SII from Rapach, Ringgenberg, and Zhou (2016), and the innovation in TIX. 0.00 indicates less than 0.005 in absolute value. The sample used for SII is 1996:02-2014:12.
Table 3: In-sample predictive regression estimation results, 1996:02-2016:04

<table>
<thead>
<tr>
<th>Predictor</th>
<th>( \hat{\beta} )</th>
<th>( R^2(%) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>DP</td>
<td>0.55</td>
<td>1.50</td>
</tr>
<tr>
<td></td>
<td>[1.32]</td>
<td></td>
</tr>
<tr>
<td>DY</td>
<td>0.62</td>
<td>1.93</td>
</tr>
<tr>
<td></td>
<td>[1.65]*</td>
<td></td>
</tr>
<tr>
<td>EP</td>
<td>0.26</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>[0.62]</td>
<td></td>
</tr>
<tr>
<td>DE</td>
<td>0.03</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>[0.07]</td>
<td></td>
</tr>
<tr>
<td>SVAR (-)</td>
<td>0.68</td>
<td>2.31</td>
</tr>
<tr>
<td></td>
<td>[1.76]**</td>
<td></td>
</tr>
<tr>
<td>BM</td>
<td>0.31</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>[1.03]</td>
<td></td>
</tr>
<tr>
<td>NTIS</td>
<td>0.55</td>
<td>1.54</td>
</tr>
<tr>
<td></td>
<td>[1.28]</td>
<td></td>
</tr>
<tr>
<td>TBL (-)</td>
<td>0.16</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>[0.58]</td>
<td></td>
</tr>
<tr>
<td>LTY (-)</td>
<td>0.27</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>[1.05]</td>
<td></td>
</tr>
<tr>
<td>LTR</td>
<td>0.16</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>[0.57]</td>
<td></td>
</tr>
<tr>
<td>TMS (-)</td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>[0.04]</td>
<td></td>
</tr>
<tr>
<td>DFY (-)</td>
<td>0.31</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>[0.62]</td>
<td></td>
</tr>
<tr>
<td>DFR</td>
<td>0.34</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>[0.69]</td>
<td></td>
</tr>
<tr>
<td>INFL</td>
<td>0.41</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>[1.19]</td>
<td></td>
</tr>
<tr>
<td>SII(-)</td>
<td>0.64</td>
<td>2.03</td>
</tr>
<tr>
<td></td>
<td>[2.01]**</td>
<td></td>
</tr>
<tr>
<td>( d\log \text{TIX} )</td>
<td>0.86</td>
<td>3.69</td>
</tr>
<tr>
<td></td>
<td>[2.44]**</td>
<td></td>
</tr>
</tbody>
</table>

The table reports the ordinary least squares estimate of \( \beta \) and \( R^2 \) statistic for the predictive regression model,

\[ r_{t+1} = \alpha + \beta x_t + \epsilon_{t+1} \quad \text{for} \quad t = 1, \cdots, T - 1, \]

where \( r_{t+1} \) is the S&P 500 log excess return, \( x_t \) is the predictor variable in the first column, and \((-)\) indicates that I take the negative of the predictor variable. Each predictor variable is standardized to have a standard deviation of one. In brackets are heteroskedasticity- and autocorrelation-robust \( t \)-statistics for testing \( H_0 : \beta = 0 \) against \( H_A : \beta > 0 \); *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively, according to wild bootstrapped \( p \)-values; 0.00 indicates less than 0.005 in absolute value. The sample used for SII is 1996:02-2014:12.
Table 4: **Out-of-sample test results**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DP</td>
<td>-2.39</td>
<td>-5.78</td>
</tr>
<tr>
<td>DY</td>
<td>-0.75</td>
<td>-3.88</td>
</tr>
<tr>
<td>EP</td>
<td>-8.55</td>
<td>-7.87</td>
</tr>
<tr>
<td>DE</td>
<td>-9.36</td>
<td>-7.25</td>
</tr>
<tr>
<td>SVAR</td>
<td>-4.17</td>
<td>-3.41</td>
</tr>
<tr>
<td>BM</td>
<td>-0.96</td>
<td>-1.24</td>
</tr>
<tr>
<td>NTIS</td>
<td>-1.33</td>
<td>0.42</td>
</tr>
<tr>
<td>TBL</td>
<td>-1.45</td>
<td>-1.12</td>
</tr>
<tr>
<td>LTY</td>
<td>-1.37</td>
<td>-1.04</td>
</tr>
<tr>
<td>LTR</td>
<td>-1.48</td>
<td>-1.33</td>
</tr>
<tr>
<td>TMS</td>
<td>-0.72</td>
<td>-0.64</td>
</tr>
<tr>
<td>DFY</td>
<td>-7.13</td>
<td>-4.75</td>
</tr>
<tr>
<td>DFR</td>
<td>-6.17</td>
<td>-5.34</td>
</tr>
<tr>
<td>INFL</td>
<td>0.51</td>
<td>0.18</td>
</tr>
<tr>
<td>SII</td>
<td>0.76*</td>
<td>-0.81</td>
</tr>
<tr>
<td>$d\log\ TIX$</td>
<td>6.60**</td>
<td>5.11**</td>
</tr>
</tbody>
</table>

The second and third columns report the proportional reduction in mean squared forecast error (MSFE) at the monthly horizon for a predictive regression forecast of the S&P 500 log excess return based on the predictor variable in the first column vis-à-vis the prevailing mean benchmark forecast, where statistical significance is based on the Clark and West (2007) statistic for testing the null hypothesis that the prevailing mean MSFE is less than or equal to the predictive regression MSFE against the alternative hypothesis that the prevailing mean MSFE is greater than the predictive regression MSFE. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The sample used for SII is 1996:02-2014:12.
### Table 5: Out-of-sample CER gains

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DP</td>
<td>0.54</td>
<td>-1.54</td>
</tr>
<tr>
<td>DY</td>
<td>1.49</td>
<td>-0.78</td>
</tr>
<tr>
<td>EP</td>
<td>0.71</td>
<td>1.01</td>
</tr>
<tr>
<td>DE</td>
<td>1.96</td>
<td>3.51</td>
</tr>
<tr>
<td>SVAR</td>
<td>0.48</td>
<td>0.77</td>
</tr>
<tr>
<td>BM</td>
<td>-0.73</td>
<td>-1.39</td>
</tr>
<tr>
<td>NTIS</td>
<td>-2.43</td>
<td>1.16</td>
</tr>
<tr>
<td>TBL</td>
<td>-0.09</td>
<td>0.30</td>
</tr>
<tr>
<td>LTY</td>
<td>0.82</td>
<td>1.14</td>
</tr>
<tr>
<td>LTR</td>
<td>-3.58</td>
<td>-3.04</td>
</tr>
<tr>
<td>TMS</td>
<td>-0.86</td>
<td>-0.67</td>
</tr>
<tr>
<td>DFY</td>
<td>0.91</td>
<td>3.61</td>
</tr>
<tr>
<td>DFR</td>
<td>-0.98</td>
<td>-0.91</td>
</tr>
<tr>
<td>INFL</td>
<td>-1.46</td>
<td>-1.82</td>
</tr>
<tr>
<td>SII</td>
<td>3.69</td>
<td>0.93</td>
</tr>
<tr>
<td>d log TIX</td>
<td>9.50</td>
<td>6.34</td>
</tr>
<tr>
<td>Buy and Hold</td>
<td>1.58</td>
<td>1.00</td>
</tr>
</tbody>
</table>

The table reports the annualized certainty equivalent return (CER) gain (in percent) for a mean-variance investor with relative risk aversion coefficient of three who allocates between equities and risk-free bills using a predictive regression excess return forecast based on the predictor variable in the first column relative to the prevailing mean benchmark forecast. The equity weight is constrained to lie between -0.5 and 1.5. Buy and hold corresponds to the investor passively holding the market portfolio. The forecast horizon and rebalancing frequency coincide and are monthly. The sample used for SII is 1996:02-2014:12.
Table 6: Return predictability: Martin’s lower bound versus TIX

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>LBwq</td>
<td>−1.89</td>
<td>0.74</td>
</tr>
<tr>
<td>LBwoq</td>
<td>−1.90</td>
<td>0.73</td>
</tr>
<tr>
<td>dlog TIX</td>
<td>6.60**</td>
<td></td>
</tr>
</tbody>
</table>

Out-of-sample $R^2$ statistics (%)

This table reports the proportional reduction in mean squared forecast error (MSFE) at the monthly horizon for a forecast of the S&P 500 log excess return based on either Martin’s lower bound or the innovation in TIX, vis-à-vis the prevailing mean benchmark forecast. For the innovation in TIX, the forecast is generated by a predictive regression. For Martin’s lower bound, which is defined as $\text{LB} = \tau e^{\tau^T} \cdot \text{SVIX}^2$, the forecast is simply the lower bound of expected returns. Two versions of SVIX are considered for robustness. SVIX$_{wq}$ represents the one used in this paper with the dividend yield taken into account. SVIX$_{woq}$ represents the one used in Martin (2017) assuming a lump sum dividend payment but a zero dividend yield. The statistical significance is based on the Clark and West (2007) statistic for testing the null hypothesis that the prevailing mean MSFE is less than or equal to the predictive regression MSFE against the alternative hypothesis that the prevailing mean MSFE is greater than the predictive regression MSFE. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.
Table 7: Return predictability: variance risk premium, left jump variation and TIX

<table>
<thead>
<tr>
<th></th>
<th>In-sample period: 1996:02-2016:04</th>
<th>In-sample period: 1996:02-2011:12</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r_{t+1}$</td>
<td>$r_{t+1}$</td>
</tr>
<tr>
<td>$d\log$ TIX</td>
<td>0.86</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>[2.44]**</td>
<td>[2.45]**</td>
</tr>
<tr>
<td>VRP</td>
<td>1.13</td>
<td>1.03</td>
</tr>
<tr>
<td>LJV</td>
<td>0.08</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>[0.17]</td>
<td>[0.50]</td>
</tr>
<tr>
<td>$R^2$(%)</td>
<td>3.69</td>
<td>6.49</td>
</tr>
</tbody>
</table>

The table reports the ordinary least squares estimate of $\beta$ and $R^2$ statistic for the predictive regression model,

$$r_{t+1} = \alpha + \beta X_t + \epsilon_{t+1} \quad \text{for} \quad t = 1, \cdots, T - 1,$$

where $r_{t+1}$ is the S&P 500 monthly log excess return, $X_t$ is a vector of predictor variables in the first column, and (−) indicates that I take the negative of the predictor variable. VRP is the variance risk premium from Hao Zhou’s website. LJV is the left jump variation shared by authors of Bollerslev, Todorov, and Xu (2015). Each predictor variable is standardized to have a standard deviation of one. Brackets below the $\hat{\beta}$ estimates report heteroskedasticity- and autocorrelation-robust $t$-statistics for testing $H_0 : \beta = 0$ against $H_A : \beta > 0$; *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively, according to wild bootstrapped $p$-values; 0.00 indicates less than 0.005 in absolute value.
Table 8: In-sample predictive regression and contemporaneous regression results at the industry level, 1996:02-2016:04

<table>
<thead>
<tr>
<th>Industry</th>
<th>$\hat{\beta}$</th>
<th>t-statistics</th>
<th>$R^2$(%)</th>
<th>$\hat{\beta}$</th>
<th>t-statistics</th>
<th>$R^2$(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P 500</td>
<td>0.86</td>
<td>2.44***</td>
<td>3.69</td>
<td>1.57</td>
<td>4.40***</td>
<td>12.21</td>
</tr>
<tr>
<td>Food</td>
<td>0.41</td>
<td>1.45*</td>
<td>1.08</td>
<td>1.17</td>
<td>3.69***</td>
<td>9.02</td>
</tr>
<tr>
<td>Mines</td>
<td>1.79</td>
<td>2.70***</td>
<td>4.15</td>
<td>2.68</td>
<td>2.95***</td>
<td>9.44</td>
</tr>
<tr>
<td>Oil</td>
<td>0.82</td>
<td>1.96**</td>
<td>2.01</td>
<td>1.29</td>
<td>2.82***</td>
<td>4.96</td>
</tr>
<tr>
<td>Clths</td>
<td>1.47</td>
<td>2.78***</td>
<td>5.52</td>
<td>1.48</td>
<td>2.93***</td>
<td>5.65</td>
</tr>
<tr>
<td>Durbl</td>
<td>1.46</td>
<td>3.05***</td>
<td>5.72</td>
<td>1.77</td>
<td>3.53***</td>
<td>8.35</td>
</tr>
<tr>
<td>Chems</td>
<td>1.65</td>
<td>2.95***</td>
<td>7.08</td>
<td>1.75</td>
<td>3.29***</td>
<td>8.02</td>
</tr>
<tr>
<td>Cnsum</td>
<td>0.17</td>
<td>0.54</td>
<td>0.17</td>
<td>1.23</td>
<td>4.90***</td>
<td>8.85</td>
</tr>
<tr>
<td>Cnstr</td>
<td>1.49</td>
<td>3.84***</td>
<td>5.87</td>
<td>1.62</td>
<td>3.47***</td>
<td>6.91</td>
</tr>
<tr>
<td>Steel</td>
<td>2.03</td>
<td>3.16***</td>
<td>4.90</td>
<td>2.80</td>
<td>3.18***</td>
<td>9.35</td>
</tr>
<tr>
<td>FabPr</td>
<td>1.31</td>
<td>3.37****</td>
<td>5.02</td>
<td>1.78</td>
<td>3.08***</td>
<td>9.33</td>
</tr>
<tr>
<td>Machn</td>
<td>1.15</td>
<td>2.25***</td>
<td>2.12</td>
<td>2.19</td>
<td>4.22***</td>
<td>7.64</td>
</tr>
<tr>
<td>Cars</td>
<td>1.39</td>
<td>2.82***</td>
<td>3.91</td>
<td>2.08</td>
<td>3.34***</td>
<td>8.77</td>
</tr>
<tr>
<td>Trans</td>
<td>0.82</td>
<td>2.09**</td>
<td>2.47</td>
<td>1.56</td>
<td>3.83***</td>
<td>8.98</td>
</tr>
<tr>
<td>Utils</td>
<td>0.13</td>
<td>0.51</td>
<td>0.10</td>
<td>0.82</td>
<td>2.42***</td>
<td>3.74</td>
</tr>
<tr>
<td>Rtail</td>
<td>0.83</td>
<td>2.34**</td>
<td>3.15</td>
<td>1.44</td>
<td>4.41***</td>
<td>9.39</td>
</tr>
<tr>
<td>Finan</td>
<td>1.28</td>
<td>2.55***</td>
<td>4.70</td>
<td>1.58</td>
<td>3.22***</td>
<td>7.13</td>
</tr>
<tr>
<td>Other</td>
<td>0.93</td>
<td>2.24**</td>
<td>3.02</td>
<td>1.84</td>
<td>4.76***</td>
<td>11.81</td>
</tr>
</tbody>
</table>

The table reports the ordinary least squares estimate of $\beta$ and $R^2$ statistic for the predictive regression model,

$$r_{t+1} = \alpha + \beta x_t + \epsilon_{t+1} \quad \text{for} \quad t = 1, \ldots, T - 1,$$

or contemporaneous regression model

$$r_t = \alpha + \beta x_t + \epsilon_t \quad \text{for} \quad t = 1, \ldots, T - 1,$$

where $r_{t+1}$ is the monthly log excess return of each industry in the first column, $x_t$ is the innovation in TIX, and (−) indicates that I take the negative of the predictor variable. Each predictor variable is standardized to have a standard deviation of one. The $t$-statistics are heteroskedasticity- and autocorrelation-robust for testing $H_0 : \beta = 0$ against $H_A : \beta > 0$; *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively, according to wild bootstrapped $p$-values; 0.00 indicates less than 0.005 in absolute value.
Table 9: **Out-of-sample test results at the industry level, 2008:10-2016:04**

<table>
<thead>
<tr>
<th>Industry</th>
<th>Out-of-sample $R^2$ statistics (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P 500</td>
<td>6.60**</td>
</tr>
<tr>
<td>Food</td>
<td>1.88*</td>
</tr>
<tr>
<td>Mines</td>
<td>7.38***</td>
</tr>
<tr>
<td>Oil</td>
<td>3.54**</td>
</tr>
<tr>
<td>Clths</td>
<td>9.04**</td>
</tr>
<tr>
<td>Durbl</td>
<td>7.97***</td>
</tr>
<tr>
<td>Chems</td>
<td>10.88***</td>
</tr>
<tr>
<td>Cnsum</td>
<td>-3.15</td>
</tr>
<tr>
<td>Cnstr</td>
<td>8.16**</td>
</tr>
<tr>
<td>Steel</td>
<td>9.98***</td>
</tr>
<tr>
<td>FabPr</td>
<td>10.23***</td>
</tr>
<tr>
<td>Machn</td>
<td>7.52***</td>
</tr>
<tr>
<td>Cars</td>
<td>6.46***</td>
</tr>
<tr>
<td>Trans</td>
<td>3.95**</td>
</tr>
<tr>
<td>Utils</td>
<td>-1.15</td>
</tr>
<tr>
<td>Rtail</td>
<td>6.53**</td>
</tr>
<tr>
<td>Finan</td>
<td>6.59**</td>
</tr>
<tr>
<td>Other</td>
<td>6.94**</td>
</tr>
</tbody>
</table>

The second column reports the proportional reduction in mean squared forecast error (MSFE) at the monthly horizon for a predictive regression forecast of the log excess returns of each industry in the first column based on the innovation in TIX vis-à-vis the prevailing mean benchmark forecast, where statistical significance is based on the Clark and West (2007) statistic for testing the null hypothesis that the prevailing mean MSFE is less than or equal to the predictive regression MSFE against the alternative hypothesis that the prevailing mean MSFE is greater than the predictive regression MSFE. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.
Table 10: VIX decomposition I: correlation matrix, 1996:01-2016:04

<table>
<thead>
<tr>
<th></th>
<th>Level</th>
<th>Log Change</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>VIX</td>
<td>Normal Risk</td>
</tr>
<tr>
<td>VIX</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Normal Risk Component</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Tail Risk Component</td>
<td>0.85</td>
<td>0.82</td>
</tr>
<tr>
<td>TIX</td>
<td>-0.70</td>
<td>-0.67</td>
</tr>
</tbody>
</table>

This table reports the correlation matrices of VIX, its components, and TIX for their levels (column 2-5) and log changes (column 6-9), respectively. VIX is decomposed based on its formula for a log stable process, \( VIX^2 = -2\sigma^2 \sec \frac{\alpha}{2} \). I use VIX(\( \sigma, \alpha \)) to indicate the dependence of VIX on normal risk \( \sigma \) and tail risk \( \alpha \). I define the normal risk component of VIX as VIX(\( \sigma, 2 \)) where the tail risk is fixed to be the same as a normal distribution which equals \( \sqrt{2}\sigma \). The tail risk component of VIX is defined as VIX - \( \sqrt{2}\sigma \). VIX and TIX, which equals \( \alpha \), are computed from S&P 500 index options prices in OptionMetrics as shown in the paper.
Table 11: **VIX decomposition I: in-sample predictive regression estimation results, 1996:02-2016:04**

<table>
<thead>
<tr>
<th>Predictor (log change)</th>
<th>$\hat{\beta}$</th>
<th>$R^2$(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VIX (-)</td>
<td>0.38</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>[1.02]</td>
<td></td>
</tr>
<tr>
<td>Normal Risk Component (-)</td>
<td>0.34</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td>[0.92]</td>
<td></td>
</tr>
<tr>
<td>Tail Risk Component (-)</td>
<td>0.68</td>
<td>2.30</td>
</tr>
<tr>
<td></td>
<td>[1.91]**</td>
<td></td>
</tr>
<tr>
<td>TIX</td>
<td>0.86</td>
<td>3.69</td>
</tr>
<tr>
<td></td>
<td>[2.44]**</td>
<td></td>
</tr>
</tbody>
</table>

The table reports the ordinary least squares estimate of $\beta$ and $R^2$ statistic for the predictive regression model,

$$r_{t+1} = \alpha + \beta x_t + \epsilon_{t+1} \quad \text{for} \quad t = 1, \cdots, T - 1,$$

where $r_{t+1}$ is the S&P 500 monthly log excess return, $x_t$ is the predictor variable in the first column, and $(-)$ indicates that I take the negative of the predictor variable. Each predictor variable is standardized to have a standard deviation of one. In brackets are heteroskedasticity- and autocorrelation-robust t-statistics for testing $H_0 : \beta = 0$ against $H_A : \beta > 0$; *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively, according to wild bootstrapped $p$-values; 0.00 indicates less than 0.005 in absolute value.
Table 12: **VIX decomposition I: Out-of-sample test results, 2008:10-2016:04**

<table>
<thead>
<tr>
<th>Predictor (log change)</th>
<th>Out-of-sample $R^2$ statistics (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VIX</td>
<td>−0.65</td>
</tr>
<tr>
<td>Normal Risk Component</td>
<td>−0.92</td>
</tr>
<tr>
<td>Tail Risk Component</td>
<td>3.34*</td>
</tr>
<tr>
<td>TIX</td>
<td>6.60**</td>
</tr>
</tbody>
</table>

The second column in this table reports the proportional reduction in mean squared forecast error (MSFE) at the monthly horizon for a predictive regression forecast of the S&P 500 log excess return based on the predictor variable in the first column vis-à-vis the prevailing mean benchmark forecast, where statistical significance is based on the Clark and West (2007) statistic for testing the null hypothesis that the prevailing mean MSFE is less than or equal to the predictive regression MSFE against the alternative hypothesis that the prevailing mean MSFE is greater than the predictive regression MSFE. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.
References


