Market Risk Premium Expectation: Combining Option Theory with Traditional Predictors*

Hong Liu
Washington University in St. Louis
Email: liuh@wustl.edu

Yueliang (Jacques) Lu
Clemson University
Email: yuelial@clemson.edu

Weike Xu
Clemson University
Email: weikex@clemson.edu

Guofu Zhou†
Washington University in St. Louis
Email: zhou@wustl.edu

First draft: December 2022

Current version: August 2023

*We greatly thank Tyler Beason, Sina Ehsani, Bing Han, Yufeng Han, Weidong Tian, Robert Van Ness, Chuanhai Zhang, conference and seminar participants at 2023 China International Risk Forum, 2023 Financial Markets and Corporate Governance Conference, 2023 Hong Kong Conference for Fintech, AI, and Big Data in Business, 2023 PKU-NUS Annual International Conference on Quantitative Finance and Economics, Fudan University, Hunan University, Hunan Normal University, Jiangxi University of Finance and Economics, Kyung Hee University, Renmin University of China, Shanghai Advanced Institute of Finance (SAIF), Tsinghua University, Washington University in St. Louis, and Xi'an Jiaotong University for their insightful comments and suggestions.

†Corresponding author.

Market Risk Premium Expectation: Combining Option Theory with Traditional Predictors

Abstract

The market risk premium is central in finance, and has been analyzed by numerous studies in the time-series predictability literature and by growing studies in the options literature. In this paper, we provide a novel link between the two literatures. Theoretically, we derive a lower bound on the equity risk premium in terms of option prices and state variables. Empirically, we show that combining information from both options and investor sentiment significantly improves the out-of-sample predictability of the market risk premium versus using either type of information alone, and that adding an economic upper bound raises predictability further.

Keywords: Out-of-sample predictability, equity risk premium, index options, sentiment, recovery

JEL Classification: G1, G11, G12, G17

1 Introduction

The expected equity market excess return, or the market risk premium, is one of the central quantities in finance and macroeconomics. Going as far back as Dow (1920), the literature on time-series predictability attempts to shed light on what economic and financial variables drive the market risk premium. For example, Fama and French (1988, 1989), Campbell and Shiller (1988a,b, 1998), and Huang, Jiang, Tu, and Zhou (2015) find that variables such as dividend-price ratio, earnings-price ratio, and investor sentiment can predict market returns.[1] Breaking new ground, Martin (2017) shows that option prices prove useful on the future market return, sparking a wealth of related research such as Kremens and Martin (2019); Martin and Wagner (2019); Kadan and Tang (2020); Chabi-Yo and Loudis (2020); Back, Crotty, and Kazempour (2022). However, the out-of-sample predictability uncovered by both strands of literature is still small.

In this paper, we provide the first study that combines two important lines of literature—time-series predictability and option recovery theory—to predict the market risk premium. We derive a new bound (a combined predictor) that incorporates the risk-neutral volatility computed from option prices and the traditional financial and macroeconomic state variables. We show that the new predictor performs well in out-of-sample forecasts and generates substantial economic gains consistently over time. In particular, it outperforms substantially than those when the method of each of the literature is used alone.

Theoretically, we follow Martin's (2017) procedure but without explicitly assuming the hypothesis of negative correlation condition (NCC). Instead, we incorporate the state variables into informative bounds on the market excess return expected in the future. In contrast to all existing extensions of Martin (2017), our study is the first to consider the role of state variables, making it possible to link the bounds to the broad classic time-series predictability literature that identifies various economic risks that impact the market.

Empirically, we construct the combined predictors using several sentiment indices in light

[1] Rapach and Zhou (2022) provide a recent survey of the literature.

of behavioral finance. To gauge forecast performance, we calculate out-of-sample R^2 statistics (R^2_{OS}) relative to the historical market mean, as suggested by Campbell and Thompson (2008). We find that neither Martin's (2017) bounds nor sentiment variables deliver consistent out-of-sample outerperformance when utilized independently: the R^2_{OS} statistics are mostly negative or positive but insignificant based on Clark and West's (2007) tests. Moreover, we find that the 'bound + past mean slackness' strategy proposed by Back, Crotty, and Kazempour's (2022) fails to improve out-of-sample performance, either.

In contrast, by combining the sentiment variables with the option bounds, we observe a substantial improvement in out-of-sample R^2_{OS} statistics. The R^2_{OS} statistics become mostly positive and are statistically significant. For instance, combing Martin's (2017) option bound with Rapach, Ringgenberg, and Zhou's (2016) short interest index generates a significant R^2_{OS} statistic of 0.695% compared with -1.450% for Martin's (2017) bound and -0.776% for short interest index. We further pool the individual forecasts from three combined predictors as Rapach, Strauss, and Zhou (2010) argue that pooling can better regularize forecast variability. Indeed, we find that pooling generates substantial forecasting gains with R^2_{OS} statistics reaching 1.417%.

We next consider imposing economic priors to further improve the forecast. In the predictability literature, Campbell and Thompson (2008) and Pettenuzzo, Timmermann, and Valkanov (2014) are the pioneering examples that incorporate economic constraints into the forecasts. Since the option theory essential provides lower bounds, we hence examine upper bounds only. For simplicity, based on MacKinlay (1995) and Cochrane and Saa-Requejo (2000), we impose an upper bound on the Sharpe ratio varying from 0.6 to 1. This upper bound notably contributes to further enhancing the predictability of the market risk premium, resulting in an R^2_{OS} as high as 2.199%, about 55% increase from the R^2_{OS} without the upper bound (which is around 1.417%).

The statistically superb performance by combining option theory with sentiment-based variables is also economically valuable. We show that the combined predictors, on average, generate higher average returns, larger out-of-sample Sharpe ratios, and greater certainty equivalent returns. By computing Fleming, Kirby, and Ostdiek's (2001) performance fee, we find that a mean-variance investor would be willing to pay more than 300 basis points per year to switch from the historical mean to acquire the forecast from the combined predictors. For example, relative to the historical mean benchmark, the performance fees are around -442 and 189 basis points if Martin's (2017) option bound and Rapach, Ringgenberg, and Zhou's (2016) short interest index are used separately, but jump to 385 basis points if the two are combined.

We have conducted a series of robustness checks to validate our findings. Firstly, we present consistent outperformance of our combined predictors over option bounds and investor sentiment in both expansion and recession periods. Secondly, our results remain robust across longer-horizon forecasts, alternative functional forms for state variables, and an extended sample period. For instance, combining option and stock market information yields significantly positive R^2_{OS} statistics of 6.065%, 14.105%, and 26.975%, for 3-month, semi-annual, and annual return forecasts, respectively. Finally, the forecast encompassing test and stabilization test lend additional support to the notion that combining information from option prices and stock market variables significantly enhances out-of-sample predictability compared to using either type of information in isolation.

Our paper makes a significant contribution to two important strands of literature on market risk premium. The first line, with Martin (2017) as a notable example, investigates how elusive it is to estimate the (conditional) expected return, which dates back to Merton (1980); Black (1993); Elton (1999). The topic is cutting-edge and invigorates a sequence of research complementarities, including Chabi-Yo and Loudis (2020) on the aggregate market; Martin and Wagner (2019); Kadan and Tang (2020); Chabi-Yo, Dim, and Vilkov (2022) on individual stocks; Heston (2021) on variance premium; Kremens and Martin (2019) on the foreign currency; Bakshi, Gao, and Xue (2022) on the treasury market; and Liu, Tang, and Zhou (2022) on the Federal Open Market Committee risk premium. The second line, as emphasized by Spiegel (2008) in *The Review of Financial Studies*, challenges researchers on whether our empirical model forecasts the equity premium any better than the historical mean. The studies in this line focus on the out-of-sample predictability of the aggregate stock market return via extensions of the conventional predictive regression approach, including Welch and Goyal (2008) on macroeconomic variables; Neely, Rapach, Tu, and Zhou (2014) on technical indicators; Dong, Li, Rapach, and Zhou (2022) on long-short anomaly portfolio returns; and Engelberg, McLean, Pontiff, and Ringgenberg (2023) on firm-level variables.

The rest of this paper is structured as follows: Section 2 presents the theory. In Section 3, we discuss the empirical framework, followed by the out-of-sample test in Section 4. Robustness checks are provided in Section 5. Finally, we conclude in Section 6.

2 Theory

In this section, based on Martin (2017), we provide the theoretical underpinnings for the important role of state variables. To facilitate our discussion, we use the following notations, where a discrete time subscript is denoted by t,

- $S_T =$ price of a stock market index (inclusive of dividends) at future time T;
- $R_T \equiv \frac{S_T}{S_t} =$ gross market return over the period t to T. We assume that $R_T > 0$;
- $\mathbb{P} =$ the real-world probability measure, and the information set at time t is $\mathscr{F}_t$;
- $\mathbb{Q} =$ the risk-neutral probability measure;
- $M_T =$ stochastic discount factor (SDF) with $\mathbb{E}_t^{\mathbb{P}}(M_T R_T) = 1$ holding;
- $R_{f,t} = \frac{1}{\mathbb{E}_t^{\mathbb{P}}(M_T)} = \mathbb{E}^{\mathbb{Q}}(R_T) =$ gross risk-free return over the period t to T (known at time t);
- $E_t^{\mathbb{P}}(x) =$ conditional expectation of a random variable under $\mathbb{P}$;
- $Cov_t^{\mathbb{P}}(x,y) =$ conditional covariance between two random variables under $\mathbb{P}$; and
- $Var_t^{\mathbb{Q}}(x,y) =$ conditional variance under $\mathbb{Q}$.

2.1 Negative correlation condition

Martin (2017) decomposes the market risk premium into two components

$$\begin{aligned}\mathbb{E}_t^{\mathbb{P}}(R_T) - R_{f,t} &= \left[\mathbb{E}_t^{\mathbb{P}}(M_T R_T^2) - R_{f,t}\right] - \left[\mathbb{E}_t^{\mathbb{P}}(M_T R_T^2) - \mathbb{E}_t^{\mathbb{P}}(R_T)\right], \\ &= \frac{1}{R_{f,t}} Var_t^{\mathbb{Q}}(R_T) - Cov_t^{\mathbb{P}}(M_T R_T, R_T). \end{aligned} \tag{1}$$

The first component, the risk-neutral variance, can be computed directly from time-t prices of index options, as known from the work of Breeden and Litzenberger (1978). The second component is a covariance term. We use the superscript $\mathbb{P}$ to highlight the fact that those quantities are under the real-world probability measure. Henceforth, we drop the superscript and use $\mathbb{E}_t(.)$ to represent the mean conditional expectation under the $\mathbb{P}$-measure.

Martin (2017) imposes a weak restriction that is termed *negative correlation condition* (NCC). He further shows that NCC holds theoretically under mild conditions in a variety of asset pricing settings, and it also holds empirically when a typical factor structure for the SDF is assumed.

Definition 1. *The negative correlation condition (NCC) holds if*

$$Cov_t(M_T R_T, R_T) \leq 0,$$

for all M_T under the real-world probability measure.

By NCC, the risk-neutral variance can be viewed as a lower bound of the equity risk premium (the expected market excess return), which is

$$\mathbb{E}_t(R_T) - R_{f,t} \geq \frac{1}{R_{f,t}} Var_t^{\mathbb{Q}}(R_T). \tag{2}$$

Martin (2017) provides the first test on whether the implied risk-neutral volatility bounds could be related directly to the equity premium, going beyond early related studies by Merton (1980); Black (1993); Elton (1999). However, the academic study on the lower bound of the equity premium remains controversial due to the fact that the NCC is pivotal to obtaining the lower bound and yet there is no direct quantification of the premise of the NCC. For instance, Bakshi, Crosby, Gao, and Zhou (2021) exploits theoretical and empirical constructions to challenge the hypothesis of the NCC. They use options on the S&P 500 index and STOXX 50 equity index and conclude that the overall tests favor the rejection.

Back, Crotty, and Kazempour (2022) recently test those lower bounds at different horizons conditionally and reject the hypothesis that they are tight for market risk premium. Therefore, using the lower bounds as forecasts of market risk premium appears insufficient in many cases due to their high slackness. Goyal, Welch, and Zafirov (2021) also examine those option bounds and demonstrate that the out-of-sample performance is never statistically significant. As a result, Back, Crotty, and Kazempour (2022) propose to add past mean slackness to Martin's (2017) option bounds as a potential solution but are impeded by the lack of enough data to estimate mean slackness. However, they stress that 150 years of data is necessary for the 'bound + mean slackness' strategy to achieve a substantial improvement in out-of-sample performance.

2.2 A Generalization of Martin (2017) Bound

The significant slackness in the lower bound derived by Martin (2017) results in the limited importance of the bound. We generalize Martin (2017) bound for an economy where asset prices depend on a vector x_t of state variables (possibly non-fundamentals such as sentiment (e.g. Asriyan, Fuchs, and Green, 2019; Hore, 2015)). In such an economy, for any security with a return process R_t (not just the market portfolio), in equilibrium we have $E[R_T] = f(x)$, for some function $f(.)$ and that $E[M_T R_T^2] = g(x)$, for some function $g(.)$. Therefore, we have

$$E[R_T] = k(x)E(M_T R_T^2), \tag{3}$$

where $k(x) \equiv f(x)/g(x)$. It is straightforward to show that the NCC assumption in Martin (2017) (and thus Martin's lower bound) is equivalent to assuming $k(x) \geq 1$.

It follows from Equation (3) that we can obtain a new relation for the market risk premium:

$$\mathbb{E}_t(R_T) - R_{f,t} = k(x)\left(\frac{1}{R_{f,t}}Var_t^{\mathbb{Q}}(R_T) + R_{f,t}\right) - R_{f,t}, \tag{4}$$

which links the risk-neutral option bound to the state variable vector. When $k(x) \geq 1$, Equation (4) reduces to Martin's bound.

Equation (4) essentially combines the forward-looking feature (option prices) and backward-looking feature (state variables). The function form for $k(x_t)$ can be either linear or non-linear. Compared with Back, Crotty, and Kazempour (2022), instead of adding the past mean values as a correction for slackness, here we use the state variables as a real-time correction to the option bounds. In the next section, we will empirically test the efficacy of Equation (4) in out-of-sample forecasting of market risk premium. We also compare both the statistical and economic performances with those using option bounds and traditional predictors alone.

3 Econometric Methodology

In this section, we first discuss three categories of predictors used to forecast the market return, including option bounds, traditional predictors, and the predictors that combine the first two. We next discuss the forecast construction and the criteria used to evaluate the out-of-sample forecasts.

3.1 Predictors

We consider three categories of predictors used for market risk premium forecast, namely, option bounds, stock market predictors, and the combined predictors that incorporate the stock market predictor into the risk-neutral option bounds.

3.1.1 Option bounds

We follow Martin (2017) to compute the option bounds of different horizons $T-t$,

$$b_t \equiv \frac{1}{R_{f,t}} Var_t^{\mathbb{Q}}(R_T) = (T-t)R_{f,t}\text{SVIX}^2_{t\to T}, \tag{5}$$

where $\text{SVIX}^2_{t\to T}$ is defined via the formula,

$$\text{SVIX}^2_{t\to T} = \frac{2}{(T-t)R_{f,t}S_t^2}\left[\int_0^{F_{t,T}} \text{put}_{t,T}(K)dK + \int_{F_{t,T}}^{\infty} \text{call}_{t,T}(K)dK\right], \tag{6}$$

where $\text{put}_{t,T}(K)$ ($\text{call}_{t,T}(K)$) denotes the market price of a put (call) option with strike K and maturity $T-t$, and $F_{t,T}$ is the forward price of the underlying.

We also consider Back, Crotty, and Kazempour (2022) and compute the slackness-adjusted option bounds

$$b_t + \text{mean slackness}, \tag{7}$$

where the slackness is simply the realized market excess return minus the Martin's (2017) bound.

We use option price data from OptionMetrics to construct time series of option bounds at time horizons $T-t = 1,3,6$, and 12 months, from January 4, 1996 to December 31, 2020. We interpolate the bound linearly to match maturities of 30, 90, 180, and 360 days. To compute the slackness-adjusted bounds, we follow Back, Crotty, and Kazempour (2022) by matching these option bounds with realized market excess returns on the S&P 500 index compounded over the 21, 63, 126, and 252 trading days. The daily return data for the S&P 500 are obtained from CRSP.

3.1.2 Stock market predictors

Studies investigating the time-series return predictability attempt to shed light on a variety of economic and financial variables that can affect the market risk premium. However, Welch and Goyal (2008) found that most macro variables fail to outperform the historical mean benchmark in out-of-sample tests, including the dividend-to-price ratio, book-to-market ratio, inflation, and others.

In the spirit of behavioral finance, we consider three sentiment-related variables as potential predictors since investor sentiment can generate return predictability (see, for example, De Long, Shleifer, Summers, and Waldmann, 1990). Rapach and Zhou (2022) provide a recent survey on the use of sentiment in return prediction. Specifically, we consider the sentiment index by Baker and Wurgler (2006), the sentiment index by Huang et al. (2015), and the short interest index by Rapach, Ringgenberg, and Zhou (2016).

3.1.3 Combined predictors

Following Equations (4) and (5), we construct a combined predictor such that,

$$b_t\left[k(x_t)\right] = k(x_t)\left[(T-t)R_{f,t}\mathrm{SVIX}^2_{t\to T} + R_{f,t}\right] - R_{f,t}, \tag{8}$$

where $k(x_t) = \exp(a + bx_t)$, and x_t denotes one of the three sentiment-based predictors above. We consider alternative function forms for $k(.)$ in Section 5.

One thing to note is that option bounds with maturities of 30, 90, 180, and 360 days are computed at a daily frequency, whereas the stock market predictor variables are at a monthly frequency. We merge the stock market predictors with the option bounds computed at the last trading day of each month, resulting in a time series of combined predictors at a monthly frequency.

Taken together, we have three categories of predictors. In the next step, we will test whether these predictors can successfully predict the market excess returns out of sample.

3.2 Forecast construction

We employ the out-of-sample tests since such tests provide the most rigorous and relevant evidence regarding stock return predictability (Welch and Goyal, 2008; Martin and Nagel, 2022;

Dong et al., 2022). Because of the horizon-matching feature in Martin's (2017) theory, we cannot focus on monthly, quarterly, or annual regressions as is common in the conventional literature (See Goyal, Welch, and Zafirov, 2021). Instead, our main results are based on market risk premium forecast over the next 21 trading days horizon, thus with the 30-day option bound. This approach is also adopted by Back, Crotty, and Kazempour (2022). For brevity, we refer to it as the 1-month forecast. We will consider longer-horizon forecasts in Section 5.

For stock market predictors and combined predictors, we begin with a standard predictive regression model,

$$r_{t+1} = \alpha_t + \beta_t Z_{i,t} + \varepsilon_{t+1}, \tag{9}$$

where r_{t+1} is the market excess return over 21 trading days, $Z_{i,t}$ is a predictor, and ε_{t+1} is a disturbance term. As in Welch and Goyal (2008); Campbell and Thompson (2008); Rapach, Strauss, and Zhou (2010), we generate out-of-sample forecasts using a recursive estimation window and obtain the one-step ahead forecast at time t as

$$\hat{r}_{t+1} = \hat{\alpha}_t + \hat{\beta}_t Z_{i,t}, \tag{10}$$

where $\{\hat{\alpha}_t, \hat{\beta}_t\}$ are ordinary least squares (OLS) estimates using the data up to time t.

For option bounds, according to the theory by Martin (2017), those bounds are already meaningful expected market returns. Therefore, we directly use them as forecasts. The same argument can be applied to the slackness-adjusted option bound by Back, Crotty, and Kazempour (2022). Thus, we have

$$\hat{r}_{t+1} = b_t, \quad \text{or} \quad \hat{r}_{t+1} = b_t + \text{mean slackness}. \tag{11}$$

Despite the theoretical framework above, we can still treat the option bound, b_t, as a predictor and run a predictive regression to obtain the forecasted return as in Equation (10),

$$\hat{r}_{t+1} = \hat{\alpha}_t + \hat{\beta}_t b_t. \tag{12}$$

We assess the forecast accuracy with an out-of-sample R^2 statistic, namely, R^2_{OS}, relative to the prevailing historical average.[2] Given T forecasts in the out-of-sample evaluation period, R^2_{OS} statistic essentially measures the relative reduction in mean square prediction error (MSPE),

$$R^2_{OS} = 1 - \frac{\sum_{t=1}^{T}(r_t - \hat{r}_t)^2}{\sum_{t=1}^{T}(r_t - \bar{r}_t^{HA})^2}. \tag{13}$$

where $\bar{r}_t^{HA}$ and $\hat{r}_t$ denote the predicated market excess returns based on the historical mean and a competing model, respectively. When $R^2_{OS} > 0$, the $\hat{r}_t$ forecast generates a lower MSPE than the prevailing mean forecast, delivering out-of-sample evidence of return predictability.

To assess the statistical significance of R^2_{OS}, we use the Clark and West's (2007) *MSPE-adjusted* statistic. Traditional predictors typically perform poorly in out-of-sample forecasts with negative R^2_{OS} values, suggesting that beating the historical average forecast is difficult (Welch and Goyal, 2008). Additionally, because market excess return consists of a large unpredictable component, R^2_{OS} is usually small. Notwithstanding, Campbell and Thompson (2008) suggest that a monthly R^2_{OS} statistic of 0.5% is the threshold for economic significance for a mean-variance investor. In Section 4.4, we will assess the economic values of various market risk premium predictors by measuring their economic values to an investor.

Our sample spans from January 1996 to December 2020 due to the option data availability. We consider three different out-of-sample forecast evaluation periods: (i) 2001:01–2020:12 with an initial estimation window of 5 years; (ii) 2006:01–2020:12 with an initial estimation window of 10 years; and (iii) 2011:01–2020:12 with an initial estimation window of 15 years. Overall, considering multiple out-of-sample periods helps provide us with a good sense of the robustness of the out-of-sample forecasting results.

[2] The historical average of market risk premium is constructed from the daily return series of the S&P 500 index, starting from July 3, 1962 on CRSP.

4 Out-of-Sample Results

In this section, we present our main out-of-sample results, the efficacy of imposing economic restrictions, possible statistical explanations, and finally the economic values of combining option prices with traditional variables.

4.1 Statistical gains

Table 1 reports the monthly R^2_{OS} statistics (in percentage) for forecasting future 1-month market excess returns for the three evaluation periods. Panels A, B, and C present the results of option bounds, sentiment-based predictors, and the combined predictors that incorporates option prices into each of the three sentiment-based measures, respectively.

In Panel A, we find that option bounds perform poorly in out-of-sample tests. Using Martin's (2017) option bound directly as a forecast of future market risk premium produces either a negative or positive but insignificant R^2_{OS} statistics. For example, we find that option bound delivers a monthly R^2_{OS} of -1.450% during the evaluation periods between 2001 and 2020. Moreover, using option bounds as regression predictors lead to much worse out-of-sample results. For instance, we find that R^2_{OS} becomes -8.445% once we run the forecast in an OLS regression. Adding the past mean slackness as a bound correction does not improve the out-of-sample performance. For instance, during the out-of-sample period of 2001–2020, the R^2_{OS} statistic is -2.533%. As argued by Back, Crotty, and Kazempour (2022), the improvement from adding past mean slackness is limited due to the lack of available slackness data.[3] For instance, for the evaluation periods between 2011 and 2020, the R^2_{OS} statistics jump from 0.741% for Martin's (2017) option bound to 2.227% for the Back, Crotty, and Kazempour (2022) adjusted bound, though neither is statistically significant.

Panel B, Table 1 reports the results of time-series predictors. In most cases, these predictors

[3] Back, Crotty, and Kazempour (2022) argue that to achieve a significantly positive R^2_{OS}, researchers need at least 150 years of data to estimate the past mean slackness.

perform poorly with negative R^2_{OS} statistics. For example, the two investor sentiment indices, IS_{BW} and IS_{HJTZ}, and the short interest index, SSI all have negative R^2_{OS} statistics of -1.443%, -0.268%, and -0.776%, respectively, in the out-of-sample period between 2001 and 2020. We only observe a significantly positive R^2_{OS} statistic for SSI in the evaluation period 2006–2020, and a marginally significant positive R^2_{OS} for IS_{BW} in the period 2011–2020. In other words, the out-of-sample predictability of traditional predictors is not robust. The lack of consistent out-of-sample evidence indicates the need for the refinement of those stock market predictors.

Panel C, Table 1, reports the forecasting results of the combined predictors that incorporate both option market information and stock market information. Compared with Martin's (2017) options bounds in Panel A, the combined predictors show much stronger out-of-sample evidence that the market return is predictable. For instance, for the period 2001–2020, combining option prices with IS_{HJTZ} and SII yields out-of-sample R^2_{OS} statistics of 0.809% and 0.695%, both significant at the 5% level. In the last row of Panel C, Table 1, we aggregate the individual forecasts from the three combined predictors by taking an arithmetic mean. Rapach, Strauss, and Zhou (2010) show that a simple combination forecast exerts a strong shrinkage effect and generally achieves better forecasting performance. Indeed, we find that pooling generates substantial forecasting gains with significantly positive R^2_{OS} statistics in all evaluation periods, ranging from 0.649% to 1.649%. As pointed out by Campbell and Thompson (2008), an R^2_{OS} of 0.5% for monthly data can signal an economically meaningful degree of return predictability in terms of increased annual portfolio returns for a mean-variance investor. Therefore, the combination of stock market information with option market information consistently produces economically significant gains over time.

To gain a better understanding of these forecasts, Figure 1 illustrates the out-of-sample forecasts for the 2001–2020 evaluation period. The blue line in each panel represents the historical mean benchmark forecast. Both the option-bound forecast (Panel A) and the short interest index forecast (Panel B) exhibit significantly higher volatility and suggest implausibly negative or unrealistically large values for numerous months during the out-of-sample testing period. In contrast, the combined predictor (Panel C) generates forecasts that are relatively less volatile, and the pooling approach (Panel D) further reduces forecast variability. It appears that combining information from both markets better regulates forecast variability. We will conduct additional statistical tests shortly.

In summary, we demonstrate that combining information from both the derivative market and the stock market significantly improves market risk premium forecasts compared to using either type of information alone.

4.2 Upper bounds

In this subsection, following Campbell and Thompson (2008); Pettenuzzo, Timmermann, and Valkanov (2014), we consider imposing a simple economic upper bound on our forecasts. The Sharpe ratio reported in the previous section corresponds to the *ex-post* measure of the realized portfolio returns in the out-of-sample period. In contrast, we delve into the concept of the *ex-ante* Sharpe ratio, which reflects the perspective of an investor at time t

$$SR_{j,t} = \frac{\hat{r}_{j,t+1}}{\hat{\sigma}_{t+1}}, \tag{14}$$

where $\hat{r}_{j,t+1}$ is the forecasted excess return on the S&P 500 index based on strategy j at time t, and $\hat{\sigma}_{t+1}$ is the forecasted volatility computed from historical return using a 5-year rolling window.

Figure 2 presents the distribution of the ex-ante Sharpe ratios from the Martin's (2017) bound, the combined bound with the aggregate short interest index (*SII*), and the pooling approach. On average, the ex-ante annualized Sharpe ratio is approximately 0.2 for the Martin (2017) bound, 0.75 for pooling, and 1 for the combined predictor of the bound and *SII*. These notably high Sharpe ratios from our combined predictors and pooling forecasts suggest the possibility of an enhancement through the imposition of an upper bound on the Sharpe ratio.

Hence, based on MacKinlay (1995) and Cochrane and Saa-Requejo (2000), we use a value varying from 0.6 to 1 to constrain the above Sharpe ratio, yielding a new and economically constrained forecast. Next, we repeat our out-of-sample one-month market risk premium forecasts for the combined predictors presented in Table 1 using both 0.6 and 1 as the maximum Sharpe ratio.

Specifically, we truncate the forecast if the ratio of the forecasted return to forecasted volatility (annualized) surpasses the chosen upper bound value, thereby ensuring a Sharpe ratio within the set limit.

Panel A of Table 2 reports the out-of-sample R^2_{OS} for our combined predictors with a maximum Sharpe ratio upper bound of 0.6 across various evaluation periods. Compared to the results without the imposition of upper bounds, we observe notably improved results with larger R^2_{OS} statistics in all evaluation periods. For instance, during the evaluation period spanning from 2001 to 2020, the R^2_{OS} statistic rises from 0.809% (0.695%) in Table 1 to 2.383% (2.148%) when combining IS_{HJTZ} (SII) with option bounds. Additionally, the Pooling approach yields an out-of-sample R^2 statistic as high as 2.199%, a marked improvement from the 1.417% in Table 1 when no economic restrictions are applied. Similar trends are observed in other out-of-sample testing periods.

Panel B of Table 2 reports the results of the out-of-sample forecasts with an upper bound of 1 for the Sharpe ratio. We also observe an improvement in out-of-sample forecasts in each evaluation period. For example, the new forecast of combing SII with option bounds yields an R^2_{OS} of 1.718% relative to 0.695% in Table 1 for the evaluation periods between 2001 and 2020. Collectively, we show a significant improvement in the out-of-sample performance by imposing an economically reasonable upper bound on the maximum Sharpe ratio.

4.3 Statistical explanation

In this subsection, we offer explanations for the exceptional out-of-sample performance achieved through the combination of options prices and traditional stock market predictors from a statistical perspective.

4.3.1 Forecast stabilization

We initiate our analysis by conducting a bias-variance assessment. Given that the R^2_{OS} statistic essentially involves a comparison of mean squared prediction errors (MSPEs) between two forecasting methods as shown in Equation (13), we adopt the decomposition of MSPE proposed by Theil (1966):

$$MSPE = \left(\bar{\hat{e}}\right)^2 + Var(\hat{e}), \tag{15}$$

where $\hat{e}$ signifies the forecast error, $\left(\bar{\hat{e}}\right)^2$ is the squared forecast bias, and $Var(\hat{e})$ is the forecast variance. Consequently, the reduction in forecast variance resulting from the combination of forecasts can contribute to a decrease in MSPE, thus potentially leading to an enhancement in the out-of-sample R^2_{OS} statistics as long as the process of combining forecasts doesn't lead to a substantial increase in bias.

Figure 3 presents three scatterplots illustrating the forecast variance and the squared forecast bias for the out-of-sample forecasts based on the historical mean, option bounds, stock market predictors, and combined predictors for the period 2001–2020. In Panel A, the forecasts derived from option bounds consistently exhibit significantly higher forecast variance compared to the historical average benchmark. In Panel B, compared with the historical mean forecast, investor sentiment displays both higher forecast variance and a higher squared forecast bias. Although the use of the short interest index slightly reduces the forecast variance, the corresponding squared forecast bias increases eightfold relative to the benchmark. As a result, neither option bounds nor stock market predictors generate smaller Mean Squared Prediction Errors (MSPEs) relative to the benchmark, leading to negative R^2_{OS} values in Table 1.

Panel C presents the bias-variance decomposition for the three combined predictors and the pooled forecasts of the three individual predictors. We find that combining option bounds with either IS_{HJTZ} or IS_{SII} results in much lower forecast variance relative to the benchmark. Despite a larger squared forecast bias, the reduction in forecast variance outweighs the increase in squared forecast bias. As a result, they both yield good out-of-sample performance relative to the historical mean forecast. Interestingly, we also find that pooling the individual forecasts reduces both squared forecast bias and forecast variance, resulting in a much smaller MSPE and thereby much larger R^2_{OS} statistics. This finding is consistent with Rapach, Strauss, and Zhou (2010), which suggests that pooling can effectively regularize forecast variability, thereby consistently generating substantial forecasting gains over time. In summary, combining information on option bounds and investor sentiment reduces forecast variance, ultimately improving out-of-sample forecasts.

4.3.2 Encompassing test

We proceed to compare the information content of different forecasts using the forecast encompassing tests developed by Chong and Hendry (1986); Fair and Shiller (1990). The forecast encompassing test provides a formal way to determine whether one forecast is statistically better at explaining the variation in future returns than another. It can help us understand whether the combined predictor significantly improves forecasting performance compared to individual predictors or historical benchmarks.

Let's consider an optimal composite forecast of r_{t+1}, which is a convex combination of forecasts from two models, labeled as i and j,

$$\hat{r}^{\star}_{t+1} = (1-\lambda)\hat{r}_{i,t+1} + \lambda \hat{r}_{j,t+1}, \quad 0 \le \lambda \le 1. \tag{16}$$

If $\lambda = 0$, this suggests that model j does not carry any useful information and thus its forecast is encompassed by model i. Conversely, if $\lambda > 0$, the forecast of model i does not encompass the model j forecast because of useful information contained in model j. Thus, forecast encompassing tests indicate that it is useful to combine forecasts from models i and j compared with using solely model i if we reject the null hypothesis of encompassing.

To test the null hypothesis that model i encompasses j ($H_0 : \lambda = 0$), against the one-sided alternative hypothesis that the model i does not encompass j ($H_1 : \lambda > 0$), we follow Harvey, Leybourne, and Newbold (1998) to compute the modified HLN-statistic over the out-of-sample evaluation period of T_0. Define $d_{t+1} = (\hat{e}_{i,t+1} - \hat{e}_{j,t+1})\hat{e}_{i,t+1}$, where $\hat{e}_{i,t+1} = r_{t+1} - \hat{r}_{i,t+1}$ and $\hat{e}_{j,t+1} = r_{t+1} - \hat{r}_{j,t+1}$. Let $\bar{d} = \frac{1}{T_0}\sum_k d_k$, and we compute

$$MHLN = \frac{T_0 - 1}{T_0}\left(\left[\hat{V}(\bar{d})\right]^{-\frac{1}{2}}\bar{d}\right) \sim t_{T_0 - 1}, \tag{17}$$

where $\hat{V}(\bar{d}) = \frac{1}{T_0}\hat{\phi}_0$ and $\hat{\phi}_0 = \frac{1}{T_0}\sum_k (d_k - \bar{d})^2$.

Table 3 reports the Harvey, Leybourne, and Newbold's (1998) *MHLN* statistic p-values applied to the evaluation periods between 2001 and 2020. Each entry in the table corresponds to the null hypothesis that the forecast estimated in the row heading is encompassed by the forecast based on the column heading. We reject the null hypothesis that the option bound forecasts encompass the combined predictors' forecasts. For example, in the third column, the p-value of "Bound & IS_{HJTZ}" is 0.03 and statistically significant. We also observe highly significant results for "bound + slackness" and "Bound (OLS)" forecasts. Additionally, we can reject the null hypothesis that the traditional predictors' forecasts encompass the combined predictors' forecasts at the 5% level. By contrast, we can not reject the null hypothesis that the combined predictors' forecasts encompass either option bounds' forecasts or traditional predictors' forecasts.

In summary, the forecast encompassing tests justify the use of information from both the option market and the stock market in equity risk premium forecasts.

4.4 Economic values

Apart from the statistical accuracy, we next compare the benchmark and competing forecasts in terms of their economic values to an investor. Specifically, consider a mean-variance investor who allocates across equities and a risk-free asset (the Treasury bill) each month. At the end of month t, the investor faces the following objective function

$$\arg_{\hat{\omega}_{t+1}} \hat{\omega}_{t+1}\hat{r}_{t+1} - \frac{\gamma}{2}\hat{\omega}_{t+1}^2\hat{\sigma}_{t+1}^2, \tag{18}$$

where γ denotes the coefficient of relative risk aversion, $\{\hat{\omega}_{t+1}, 1-\hat{\omega}_{t+1}\}$ are allocation weights to the market portfolio and the risk-free asset at month $t+1$, $\hat{r}_{t+1}$ is the investor's market excess return forecast, and $\hat{\sigma}_{t+1}^2$ is the investor's forecast of the variance of the market excess return. The optimal mean-variance portfolio weight on the market can be computed as

$$\hat{\omega}^{\star}_{t+1} = \left(\frac{1}{\gamma}\right)\left(\frac{\hat{r}_{t+1}}{\hat{\sigma}^2_{t+1}}\right). \tag{19}$$

We follow Campbell and Thompson (2008) to set γ to be 3 and to constrain the portfolio weight on stocks to lie between $[0, 1.5]$ each month in Equation (19). Over the out-of-sample periods, we compute four quantities (performance measures), based on the mean $\hat{\mu}_j$ and standard deviation $\hat{\sigma}_j$ of the out-of-sample realized returns by a forecasting method j. First, we measure the *out-of-sample Sharpe ratio* (SRatio)

$$\hat{s}_j = \frac{\hat{\mu}_j}{\hat{\sigma}_j}. \tag{20}$$

To test whether the Sharpe ratios of the two strategies are statistically distinguishable, we follow DeMiguel, Garlappi, and Uppal (2009) to compute the p-value of their difference.

Second, we compute the *certainty-equivalent return* (CER) of each strategy,

$$C\hat{E}R_j = \hat{\mu}_j - \frac{\gamma}{2}\hat{\sigma}^2_j. \tag{21}$$

Relative to a benchmark, we also compute the CER difference, which is known as the utility gain in the forecasting literature (see, e.g., Rapach and Zhou, 2022).

Next, we compute DeMiguel, Garlappi, and Uppal's (2009) *return-loss value* with respect to Rapach, Strauss, and Zhou's (2010) simple pooling. Precisely, suppose $\{\hat{\mu}_b, \hat{\sigma}_b\}$ are the monthly out-of-sample mean and volatility of the excess returns from the benchmark, the return-loss from the competing forecast j is

$$\text{return-loss}_j = \left(\frac{\hat{\mu}_b}{\hat{\sigma}_b}\right) \times \hat{\sigma}_j - \hat{\mu}_j. \tag{22}$$

In other words, the return-loss is the additional return needed for strategy j to perform as well as the benchmark. Therefore, a negative return-loss value indicates that the method j outperforms the simple pooling in terms of the Sharpe ratio.

Lastly, we compute the *performance fee* suggested in Fleming, Kirby, and Ostdiek (2001). It can be interpreted as the maximum fee that a quadratic-utility investor would be willing to pay to switch from the benchmark to the alternative. To estimate this fee, we find the value of Δ that solves

$$\sum_t \left[\left(R_{j,t} - \Delta\right) - \frac{\gamma}{2(1+\gamma)} \left(R_{j,t} - \Delta\right)^2 \right] = \sum_t \left[R_{b,t} - \frac{\gamma}{2(1+\gamma)} R_{b,t}^2 \right], \tag{23}$$

where $R_{j,t}$ and $R_{b,t}$ denote the out-of-sample realized returns by the competing forecast j and the benchmark, respectively. We report the estimate of Δ as annualized fees in basis points.

Figure 4 plots log cumulative excess returns for portfolios using market excess return forecasts based on option bound, short interest index, and the combined predictor. The figure also depicts the cumulative excess return for the portfolio based on the historical average benchmark. Figure 4 reveals that the portfolio relying solely on option bounds underperforms the portfolio based on the historical mean benchmark, suffering significant losses during the financial crisis in 2008/09. Conversely, the portfolio that relies exclusively on the short interest index outperforms the benchmark. Remarkably, the portfolio that integrates information from both options and the short interest index demonstrates even better performance. Furthermore, pooling across individual forecasts from the combined predictors offers better resilience against downside risk during market stress while maintaining potential for upside gains.

Table 4 reports the above performance measures for returns over the period 2001–2020, all annualized. In Panel A, benchmark values are reported when using the historical average forecast, including average excess return, standard deviation, Sharpe ratio, and certainty equivalent return. It is observed in Panel B that forecasts based on Martin (2017) option bound and Back, Crotty, and Kazempour (2022) slackness-adjusted bound result in standard deviations twice the benchmark magnitude, leading to a significantly smaller Sharpe ratio and negative CER. Positive return-loss values and negative performance fees both indicate a strong preference of investors for the historical average benchmark over the option bound forecast. This aligns with the negative out-of-sample R_{OS}^2 statistics detailed in Table 1.

Panel C and D of Table 4 present the performance measures based on forecasts from stock market predictors and combined predictors. It is observed that utilizing stock market predictors alone yields reasonably substantial economic values relative to the benchmark, with larger Sharpe ratios, larger CERs, negative return-loss values, and positive performance fees. Additionally, combining information from both stock and option markets generates even greater forecasting improvements compared to using either type of information in isolation. For instance, relative to the Sharpe ratio of 0.212 from the historical mean benchmark, the short interest index yields a Sharpe ratio of 0.435—twice the magnitude, with the difference being marginally significant at the 10% level. Furthermore, combining the short interest index with option bounds results in a Sharpe ratio of 0.525—a 20% increase over the short interest index alone. This difference relative to the benchmark becomes more significant at the 5% level. Moreover, performance fees rise from 189 bps to 385 bps as we transition from using the short interest index alone to using the combined predictor. The substantial performance fee suggests investors are willing to pay approximately 385 bps per annum to access information from both markets. These results hold up well under a proportional transaction cost.

In conclusion, Table 4 emphasizes the value of combining forward-looking (option data) and backward-looking (sentiment variables) attributes for forecasting the market risk premium.

5 Robustness

In this section, we perform robustness checks for our out-of-sample tests. We begin by conducting separate out-of-sample forecasts based on NBER-dated business cycles. Additionally, we delve into long-horizon return forecasts. Finally, we present the results using alternative functional forms and an extended data sample.

5.1 Business cycles

To provide a visual representation of the consistency in forecast construction across various market conditions, we proceed to present time-series plots depicting the forecasting error in conjunction with the NBER-dated business cycles. In particular, we calculate the cumulative differences in squared forecast errors between the historical average benchmark and each alternative forecast, spanning the period 2001-2020 (Welch and Goyal, 2008; Rapach, Strauss, and Zhou, 2010),

$$\text{square error difference} = (r_t - \bar{r}^{HA}_{t|t-1})^2 - (r_t - \hat{r}_{t|t-1})^2, \tag{24}$$

where $\bar{r}^{HA}_{t|t-1}$ is the historical average (HA), r_t is the realized market excess return, and $\hat{r}_{|t-1}$ is the forecast based on a competing model.

The cumulative differences, as depicted in Figure 5, offer a straightforward approach to determine whether a competing forecast outperforms the benchmark within any given subsample. By comparing the curve's height at the subsample's commencement and conclusion, such determination becomes readily achievable. An elevated (reduced) curve at the conclusion indicates that the competing forecast displays a lower (higher) Mean Squared Forecast Error (MSFE) relative to the benchmark throughout the specified period. A predominantly positively sloped curve denotes consistent out-of-sample enhancements from the competing forecast, while a sharply negatively sloped segment suggests a period of considerable underperformance. Notably, forecasts solely grounded in the option bound or the short interest index in Figure 5 do not consistently exhibit gains in accuracy over time. However, when combining the option bound with the short interest index, accuracy gains improve, as evident from the curve turning positive around 2010. Moreover, pooling individual forecasts from combined predictors results in significantly higher accuracy gains, with the cumulative curve consistently staying above zero after 2009 and concluding at a higher point by the end of the testing period.

Notably, the failure of the option bound proposed by Martin (2017) is primarily attributed to the NBER-dated recessions surrounding the 2008/09 global financial crisis. During this period, the option bound significantly underperforms the historical average benchmark. A plausible explanation is that the effectiveness of the Martin (2017) bound is contingent on the NCC assumption. However, we observe frequent violations of the NCC during the evaluation period spanning from 2001 to 2020, as depicted in Figure 6. These violations are concentrated either in periods of high investment sentiment or during market recessions, such as the crisis in 2008/09. This can be attributed to an increased demand for options as hedging instruments during severe market conditions, leading to substantial increases in option prices. Since the option bound is essentially a weighted average of market prices of index options, its usage as a ”meaningful expected return,” as in Martin (2017), can result in unrealistic positive forecasts during periods of stock market stress when option prices skyrocket.

To formally assess out-of-sample forecasting across different market periods, we calculate separate R^2_{OS} statistics for NBER-dated expansions and recessions within the out-of-sample testing intervals. In total, three recessions occurred during the out-of-sample period spanning from January 2001 to December 2020, corresponding to business-cycle peaks in 2001:03, 2007:12, and 2020:02, and troughs in 2001:11, 2009:06, and 2020:04. Given the limited sample size during recessions (approximately 30 observations), we solely evaluate the statistical significance of positive R^2_{OS} statistics within expansions.

Consistent with the preceding argument, we observe that the R^2_{OS} statistics using the option bound are more negative in recessions than in expansions in Panel A, Table 5. This discrepancy suggests considerably poorer predictability of the option bound during recessionary periods, which also accounts for the negative R^2_{OS} statistics calculated for the entire sample period. Moving to Panel B, it becomes evident that sentiment-based predictors perform inadequately, yielding notably negative R^2_{OS} statistics during recessions, while generally exhibiting positive and significant R^2_{OS} during expansions. Consequently, the R^2_{OS} statistics for the entire sample period turn negative. For instance, IS_{HJTZ} generates a slightly significant R^2_{OS} statistic of 0.307% during expansions, but a negative R^2_{OS} of -1.332% during recessions. In Panel C, it is noticeable that combined predictors and pooling both yield positive R^2_{OS} statistics during both expansion and recession periods (except for an exception for *SSI*). As an illustration, upon combining the option bound with IS_{HJTZ}, the R^2_{OS} statistics amount to 0.814% and 0.800% for expansions and recessions, ultimately culminating in a significantly positive R^2_{OS} of 0.809% for the entire out-of-sample period.

To summarize, the combination of option bounds with stock market predictors is more likely to restore market risk premium predictability during both expansion and recession periods, ultimately surpassing the historical average benchmark across the board.

5.2 Long-horizon forecasts

In this subsection, we repeat the out-of-sample tests using various long horizons. As previously discussed, we diverge from the common monthly, quarterly, or annual regressions found in the literature due to the horizon-matching characteristic inherent in Martin (2017)'s theory. Given the feasibility of computing option bounds for various horizons via option maturities alignment, we examine the out-of-sample performances over 3-month, 6-month, and 12-month horizons.

Table 6 reports the out-of-sample R^2_{OS} statistics for the evaluation period spanning from 2001 to 2020. For the 3-month horizon forecast, option bounds fail to surpass the historical average forecast, resulting in R^2_{OS} statistics that are either negative or positive yet insignificant. In contrast, stock market predictors begin to outperform the historical mean benchmark forecast, displaying notably positive and significant R^2_{OS} values, ranging from 1.181% to 3.471%. Additionally, the combined predictors of IS_HJTZ, SSI, and pooling generate even more robust out-of-sample performance relative to the benchmark forecast. For the 3-month horizon forecasts, this combined approach yields notably positive R^2_{OS} statistics, ranging from 2.435% to 6.065%.

In general, the performance of out-of-sample forecasting improves as the forecast horizon lengthens. For semi-annual and annual forecasting, both option bounds and traditional predictors yield positive and significant R^2_{OS} statistics. This outcome is expected given the overall upward trend observed in the market over the long term. Similar results are obtained by Campbell and Thompson (2008) when they forecast annual market return using macro variables (Panel B, Table 2, in their paper). Despite this enhanced performance, the combination of the two information sets continues to yield superior outcomes, particularly evident with the application of the pooling method. To illustrate, pooling results in notably high R^2_{OS} statistics, reaching 14.105% and 26.975% for the semi-annual and annual horizons, respectively. In summary, our combined predictors consistently demonstrate superior performance compared to using either information set alone across various forecast horizons.

5.3 Alternative function forms

So far, we have demonstrated that the generalization of Martin (2017) bound by integrating sentiment-based variables generates superior out-of-sample forecasting gains. In Equation (8), we have employed an exponential form for the function $k(.)$. In this subsection, we consider two alternative function forms $k(x_t)$. Subsequently, we replicate the out-of-sample forecasts across 1-, 3-, 6- and 12-month horizons, using these new combined predictors.

Case 1: Exponential form with individual terms. Equation (8) only scales and operates through interaction. We modify $k(x_t)$ to incorporate both interaction and individual terms[4]

$$b_t\left[k(x_t)\right] = \left[\underbrace{\exp(ax_t + b) + c}_{k(x_t)}\right] \times \left[(T-t)R_{f,t}\text{SVIX}^2_{t\to T} + R_{f,t}\right] - R_{f,t}. \tag{25}$$

Case 2: Linear form. We consider a linear function form for $k(x_t) = ax_t + b$ so that

$$b_t\left[k(x_t)\right] = \left[\underbrace{a + bx_t}_{k(x_t)}\right] \times \left[(T-t)R_{f,t}\text{SVIX}^2_{t\to T} + R_{f,t}\right] - R_{f,t}. \tag{26}$$

The results are presented in Table 7. We find consistent results as the results observed in Tables 1 and 6. These alternative combined predictors consistently exhibit substantial outperformance in comparison to the historical mean forecast across diverse horizons. In most instances, the out-of-sample R^2 statistics are statistically significant at the 5% level or better. For example, after incorporating the short interest index into option bounds, we obtain R^2_{OS} statistics of 0.448%, 3.575%, 8.764%, and 14.387% for the 1-, 3-, 6-, and 12-month horizons, respectively. Moreover, the pooling forecasts consistently outperform the prevailing mean at various horizons with R^2_{OS} statistics, ranging from 1.325% to 26.456%.

[4] We thank Tyler Beason for this great suggestion.

5.4 Extended sample period

Our sample begins in 1996 due to the availability of option data. Both the option bounds and our combined predictors hinge on constructing SVIX from market prices of index options. Given the acknowledged strong correlation between SVIX and the publicly traded VIX index (with a correlation coefficient as high as 0.99), we collect the VIX data from CBOE starting from 1990. Subsequently, we reconstruct the combined predictor as follows,

$$b_t\left[k(x_t)\right] = \begin{cases} k(x_t)\left[(T-t)R_{f,t}\mathrm{SVIX}^2_{t\to T} + R_{f,t}\right] - R_{f,t}, & \text{if} \quad t \geq 1996, \\ k(x_t)\left[(T-t)R_{f,t}\mathrm{VIX}^2_{t\to T} + R_{f,t}\right] - R_{f,t}, & \text{if} \quad t < 1996 \end{cases} \tag{27}$$

Recall that,

$$\mathrm{SVIX}^2_{t\to T} = \frac{2}{(T-t)R_{f,t}S_t^2}\left[\int_0^{F_{t,T}} \mathrm{put}_{t,T}(K)dK + \int_{F_{t,T}}^{\infty} \mathrm{call}_{t,T}(K)dK\right]. \tag{28}$$

Similarly, we can formulate the equation for VIX based on index option prices,

$$\mathrm{VIX}^2_{t\to T} = \frac{2R_{f,t}}{(T-t)}\left[\int_0^{F_{t,T}} \frac{1}{K^2}\mathrm{put}_{t,T}(K)dK + \int_{F_{t,T}}^{\infty} \frac{1}{K^2}\mathrm{call}_{t,T}(K)dK\right]. \tag{29}$$

VIX and SVIX both capture important aspects of market return, and and their disparity can be minimal under certain conditions, such as log-normality (see Martin, 2017). Figure 7 depicts the 30-day SVIX index and the CBOE VIX index. Both indices showcase nearly identical patterns. While SVIX gauges risk-neutral volatility, VIX measures risk-neutral entropy, thus portraying a slight variation in their definitions.

$$VIX^2_{t\to t+T} = \frac{2}{T}L_t^Q\left(\frac{R_{t\to t+T}}{R_{f,t\to t+T}}\right), \tag{30}$$

where $L_t^Q(X) \equiv log\mathbb{E}_t^Q X - \mathbb{E}_t^Q logX$.

We extend the full sample to encompass the period from January 1990 to December 2020 and subsequently re-run our out-of-sample forecasts using the same evaluation periods as outlined in Table 1. The resulting R^2_{OS} statistics are presented in Table 8. We find even stronger results than those in Table 1. For example, during the evaluation period between 2001 and 2020, all combined predictors, except for the Baker and Wurgler's (2006) sentiment index, outperform both option bounds and standalone time-series predictors. The R^2_{OS} statistics are as follows: 1.625% for the combined predictor involving Huang et al. (2015)'s investor sentiment index, 1.498% for Rapach, Ringgenberg, and Zhou (2016)'s short interest index, and 1.686% for the pooling approach. Additionally, these R^2_{OS} statistics are all statistically significant at the 5% level. Collectively, we demonstrate the robustness of the out-of-sample forecasts when employing the extended sample.

6 Conclusion

Predicting the market risk premium, or expected market excess return, is one of the fundamental challenges in finance because the market risk premium is a pivotal determinant of the required rate of returns for investors to hold assets in asset pricing models. Despite the numerous studies on time-series predictability in the literature, the out-of-sample predictability continues to exhibit limited effectiveness. Recent developments by Martin (2017) and others shed light on the expected market excess return from options prices, nevertheless, empirical performance remains unsatisfactory.

In this paper, we provide the first study on combining two lines of literature on market risk premium. We theoretically derive a new bound on the market risk premium by combining risk-neutral variance with sentiment-based state variables. We further show that the new combined predictor performs well empirically and the improvement in out-of-sample forecasting is economically substantial. Collectively, our paper provides new insights into the market risk premium by drawing perspectives from both the time-series return predictability literature and the option literature.

References

Asriyan, V., Fuchs, W., Green, B., 2019. Liquidity sentiments. American Economic Review 109, 3813–3848.

Back, K., Crotty, K., Kazempour, S. M., 2022. Validity, tightness, and forecasting power of risk premium bounds. Journal of Financial Economics 144, 732–760.

Baker, M., Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. The Journal of Finance 61, 1645–1680.

Bakshi, G., Crosby, J., Gao, X., Zhou, W., 2021. Negative correlation condition and the dark matter property of asset pricing models. Working paper.

Bakshi, G., Gao, X., Xue, J., 2022. Recovery with applications to forecasting equity disaster probability and testing the spanning hypothesis in the treasury market. Journal of Financial and Quantitative Analysis 58, 1808–1842.

Black, F., 1993. Estimating expected return. Financial Analysts Journal 49, 36–38.

Breeden, D. T., Litzenberger, R. H., 1978. Prices of state-contingent claims implicit in option prices. Journal of Business 51, 621–651.

Campbell, J. Y., Shiller, R. J., 1988a. The dividend-price ratio and expectations of future dividends and discount factors. The Review of Financial Studies 1, 195–228.

Campbell, J. Y., Shiller, R. J., 1988b. Stock prices, earnings, and expected dividends. The Journal of Finance 43, 661–676.

Campbell, J. Y., Shiller, R. J., 1998. Valuation ratios and the long-run stock market outlook. The Journal of Portfolio Management 24, 11–26.

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

Chabi-Yo, F., Dim, C., Vilkov, G., 2022. Generalized bounds on the conditional expected excess return on individual stocks. Management Science .

Chabi-Yo, F., Loudis, J., 2020. The conditional expected market return. Journal of Financial Economics 137, 752–786.

Chong, Y. Y., Hendry, D. F., 1986. Econometric evaluation of linear macro-economic models. The Review of Economic Studies 53, 671–690.

Clark, T. E., West, K. D., 2007. Approximately normal tests for equal predictive accuracy in nested models. Journal of Econometrics 138, 291–311.

Cochrane, J. H., Saa-Requejo, J., 2000. Beyond arbitrage: Good-deal asset price bounds in incomplete markets. Journal of Political Economy 108, 79–119.

De Long, J. B., Shleifer, A., Summers, L. H., Waldmann, R. J., 1990. Noise trader risk in financial markets. Journal of Political Economy 98, 703–738.

DeMiguel, V., Garlappi, L., Uppal, R., 2009. Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? The Review of Financial studies 22, 1915–1953.

Dong, X., Li, Y., Rapach, D. E., Zhou, G., 2022. Anomalies and the expected market return. The Journal of Finance 77, 639–681.

Dow, C. H., 1920. Scientific stock speculation. Magazine of Wall Street.

Elton, E. J., 1999. Presidential address: Expected return, realized return, and asset pricing tests. The Journal of Finance 54, 1199–1220.

Engelberg, J., McLean, R. D., Pontiff, J., Ringgenberg, M. C., 2023. Do cross-sectional predictors contain systematic information? Journal of Financial and Quantitative Analysis 58, 1172–1201.

Fair, R. C., Shiller, R. J., 1990. Comparing information in forecasts from econometric models. The American Economic Review 80, 375–389.

Fama, E. F., French, K. R., 1988. Dividend yields and expected stock returns. Journal of Financial Economics 22, 3–25.

Fama, E. F., French, K. R., 1989. Business conditions and expected returns on stocks and bonds. Journal of Financial Economics 25, 23–49.

Fleming, J., Kirby, C., Ostdiek, B., 2001. The economic value of volatility timing. The Journal of Finance 56, 329–352.

Goyal, A., Welch, I., Zafirov, A., 2021. A comprehensive look at the empirical performance of equity premium prediction ii. Working paper.

Harvey, D. I., Leybourne, S. J., Newbold, P., 1998. Tests for forecast encompassing. Journal of Business & Economic Statistics 16, 254–259.

Heston, S. L., 2021. Recovering the variance premium. Working paper.

Hore, S., 2015. Equilibrium predictability, term structure of equity premia, and other return characteristics. Review of Finance 19, 423–466.

Huang, D., Jiang, F., Tu, J., Zhou, G., 2015. Investor sentiment aligned: A powerful predictor of stock returns. The Review of Financial Studies 28, 791–837.

Kadan, O., Tang, X., 2020. A bound on expected stock returns. The Review of Financial Studies 33, 1565–1617.

Kremens, L., Martin, I., 2019. The quanto theory of exchange rates. American Economic Review 109, 810–43.

Liu, H., Tang, X., Zhou, G., 2022. Recovering the fomc risk premium. Journal of Financial Economics 145, 45–68.

MacKinlay, A. C., 1995. Multifactor models do not explain deviations from the capm. Journal of Financial Economics 38, 3–28.

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

Martin, I. W., Nagel, S., 2022. Market efficiency in the age of big data. Journal of Financial Economics 145, 154–177.

Martin, I. W., Wagner, C., 2019. What is the expected return on a stock? The Journal of Finance 74, 1887–1929.

Merton, R. C., 1980. On estimating the expected return on the market: An exploratory investigation. Journal of Financial Economics 8, 323–361.

Neely, C. J., Rapach, D. E., Tu, J., Zhou, G., 2014. Forecasting the equity risk premium: the role of technical indicators. Management Science 60, 1772–1791.

Pettenuzzo, D., Timmermann, A., Valkanov, R., 2014. Forecasting stock returns under economic constraints. Journal of Financial Economics 114, 517–553.

Rapach, D., Zhou, G., 2022. Asset pricing: Time-series predictability. Oxford Research Encyclopedia of Economics and Finance pp. 1–34.

Rapach, D. E., Ringgenberg, M. C., Zhou, G., 2016. Short interest and aggregate stock returns. Journal of Financial Economics 121, 46–65.

Rapach, D. E., Strauss, J. K., Zhou, G., 2010. Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. The Review of Financial Studies 23, 821–862.

Spiegel, M., 2008. Forecasting the equity premium: Where we stand today. The Review of Financial Studies 21, 1453–1454.

Theil, H., 1966. Applied economic forecasting. Amsterdam: North-Holland.

Welch, I., Goyal, A., 2008. A comprehensive look at the empirical performance of equity premium prediction. The Review of Financial Studies 21, 1455–1508.

Table 1: R^2_{OS} statistics (in percent) for 1-month market risk premium forecast

This table reports out-of-sample R^2 statistics (R^2_{OS}) in percent for 1-month market excess return forecasts based on option bounds, time-series (stock market) predictors, and combined predictors. The formulation of the combined predictor is as follows:

$$b_t\left[k(x_t)\right] = k(x_t)\left[(T-t)R_{f,t}\mathrm{SVIX}^2_{t\to T} + R_{f,t}\right] - R_{f,t},$$

where $R_{f,t}$ is the gross return on an risk-free asset, $\mathrm{SVIX}^2_{t\to T}$ is computed from market prices of index options as in Martin (2017), x_t is one of time-series variables, and $k(x_t) = exp(a+bx_t)$. In the last row, we pool the individual forecasts from the combined predictors by taking an arithmetic mean.

Panel A: Option bounds are represented when $k(x_t) = 1$. The bound is employed directly as a measure of market excess return. Additionally, we use the bound as a predictor in predictive regression, termed as "Bound (OLS)". Furthermore, the past mean slackness, which is the realized market excess return minus the bound, is utilized as a correction to the bound, as suggested by Back, Crotty, and Kazempour (2022).

Panel B: Traditional time-series predictors are depicted when $b_t\left[k(x_t)\right] = x_t$. Specifically, we utilize sentiment indices including Baker and Wurgler (2006)'s sentiment index (IS_{BW}), Huang et al. (2015)'s sentiment index (IS_{HJTZ}), and Rapach, Ringgenberg, and Zhou (2016)'s short interest index (SSI).

The out-of-sample periods are 2001:01–2020:12, 2006:01–2020:12, and 2011:01–2020:12, as indicated in the column headings. Using the Clark and West (2007) test, asterisks (*) and double asterisks (**) denote significance at the 10% and 5% levels for positive R^2_{OS}, respectively.

Out-of-sample periods	2001−2020	2006−2020	2011−2020
Panel A: Option bounds			
Bound	−1.450	−2.214	0.741
Bound + Slackness	−2.533	−2.439	2.227
Bound (OLS)	−8.445	−9.519	−3.206
Panel B: Traditional predictors			
$x_t = IS_{BW}$	−1.443	−0.401	1.869*
$x_t = IS_{HJTZ}$	−0.268	0.085	0.860
$x_t = SSI$	−0.776	1.520**	−0.553
Panel C: Combined predictors			
$x_t = IS_{BW}$	−0.788	−0.481	1.717*
$x_t = IS_{HJTZ}$	0.809**	−0.632	1.504*
$x_t = SSI$	0.695**	1.822**	0.186*
Pooling	1.417**	0.649*	1.649*

Table 2: R^2_{OS} statistics (in percent) with upper bounds

This table reports out-of-sample R^2 statistics (R^2_{OS}) in percent for 1-month market excess return based on the combined predictor such that

$$b_t\left[k(x_t)\right] = k(x_t)\left[(T-t)R_{f,t}\mathrm{SVIX}^2_{t\to T} + R_{f,t}\right] - R_{f,t},$$

where $\mathrm{SVIX}^2_{t\to T}$ is computed from market prices of index options as in Martin (2017), x_t is one of three sentiment variables, and $k(x_t) = exp(a+bx_t)$. We consider two Sharpe ratio values as upper bounds, as indicated in panel headings. In the last row of each panel, we pool the individual forecasts by taking an arithmetic mean. Based on the Clark and West's (2007) test, *, ** and *** indicate significance at the 10%, 5% and 1% levels for the positive R^2_{OS}, respectively.

Out-of-sample periods	2001–2020	2006–2020	2011–2020
Panel A: Maximum Sharpe ratio of 0.6			
$x_t = IS_{BW}$	0.231	0.36	1.905**
$x_t = IS_{HJTZ}$	2.383**	0.532	1.776*
$x_t = SSI$	2.148**	2.413**	1.497*
Pooling	2.199**	1.198**	1.795*
Panel B: Maximum Sharpe ratio of 1.0			
$x_t = IS_{BW}$	−0.627	−0.325	1.705*
$x_t = IS_{HJTZ}$	1.282**	−0.121	1.576*
$x_t = SSI$	1.718**	2.347**	1.009*
Pooling	1.774**	0.955*	1.802*

Table 3: Forecast encompassing test results, *MHLN* statistic *p*-values

This table reports *p*-values for the Harvey, Leybourne, and Newbold's (1998) *MHLN* statistic for the out-of-sample forecasting. The forecasts are derived from option bounds, traditional predictors, and combined predictors. The statistic corresponds to an upper-tail test of the null hypothesis that the forecast given in the column heading encompasses the forecast given in the row heading against the alternative hypothesis that the forecast given in the column heading does not encompass the forecast given in the row heading. The out-of-sample period is 2001:01–2020:12.

(1)	(2) HA	(3) Bound	(4) Bound + Slackness	(5) Bound (OLS)	(6) IS_{BW}	(7) IS_{HJTZ}	(8) *SSI*	(9) Bound + IS_{BW}	(10) Bound + IS_{HJTZ}	(11) Bound + *SSI*	(12) Pooling
HA		0.154	0.059	0.020	0.031	0.055	0.021	0.075	0.177	0.098	0.323
Bound	0.743		0.038	0.024	0.104	0.178	0.070	0.150	0.310	0.218	0.490
Bound + Slackness	0.797	0.686		0.036	0.325	0.463	0.324	0.288	0.588	0.568	0.794
Bound (OLS)	0.791	0.651	0.608		0.536	0.614	0.568	0.479	0.685	0.667	0.760
IS_{BW}	0.292	0.092	0.077	0.011		0.692	0.120	0.502	0.823	0.270	0.960
IS_{HJTZ}	0.088	0.038	0.043	0.009	0.090		0.045	0.182	0.759	0.125	0.746
SSI	0.077	0.024	0.029	0.018	0.051	0.078		0.071	0.174	0.888	0.496
Bound + IS_{BW}	0.161	0.069	0.062	0.009	0.176	0.352	0.067		0.644	0.142	0.804
Bound + IS_{HJTZ}	0.050	0.030	0.034	0.007	0.046	0.124	0.030	0.072		0.076	0.517
Bound + *SSI*	0.029	0.016	0.023	0.015	0.021	0.028	0.067	0.035	0.073		0.256
Pooling	0.028	0.017	0.021	0.008	0.007	0.033	0.030	0.030	0.127	0.100	

Table 4: Economic values

This table provides various economic metrics for a mean-variance investor characterized by a relative risk aversion coefficient of three. The investor reallocates between equities and risk-free bills on a monthly basis during the out-of-sample period from 2001:01 to 2020:12. The allocation weights depend on the return forecasts as indicated by panel headings. The computed performance measures encompass out-of-sample average excess return, standard deviation, Sharpe ratio (SRatio), certainty equivalent return (CER), DeMiguel, Garlappi, and Uppal (2009) return-loss value, and Fleming, Kirby, and Ostdiek (2001) performance fee (Fee). For both SRatio and CER, we additionally calculate the disparity between various competing models and the historical mean benchmark. All results are annualized. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

(1)	(2) Avg. Ret (%)	(3) S.D. (%)	(4) SRatio	(5) CER (%)	(6) SRatio diff	(7) CER diff	(8) Ret-loss (%)	(9) Fee (bps)
				Panel A: Prevailing mean benchmark				
HA	1.994	9.390	0.212	0.672				
				Panel B: Option bounds				
Bound	0.007	15.790	0.000	−3.733	−0.212	−4.408	3.346	−442.558
Bound + Avg slackness	1.310	20.796	0.063	−5.177	−0.149	−5.841	3.106	−589.686
Bound (OLS)	2.701	18.569	0.145	−2.471	−0.067	−3.132	1.243	−316.664
				Panel C: Traditional predictors				
$x_t = IS_{BW}$	5.651	20.548	0.275	−0.683	0.063	−1.348	−1.287	−140.101
$x_t = IS_{HJTZ}$	7.509	18.459	0.407	2.398	0.194*	1.732	−3.589	169.751
$x_t = SSI$	8.961	20.608	0.435	2.590	0.222*	1.946	−4.584	189.232
				Panel D: Combined predictors				
$x_t = IS_{BW}$	7.503	19.281	0.389	1.926	0.177	1.260	−3.408	121.886
$x_t = IS_{HJTZ}$	8.403	19.470	0.432	2.717	0.219*	2.049	−4.268	200.599
$x_t = SSI$	10.154	19.342	0.525	4.543	0.313**	3.896	−6.047	385.479
Pooling	9.121	18.737	0.487	3.855	0.274**	3.194	−5.142	315.806

Table 5: R^2_{OS} statistics (in percent) for NBER-dated business cycles

This table presents out-of-sample R^2 statistics (R^2_{OS}) as percentages for 1-month market excess return forecasts derived from option bounds, time-series predictors, and combined predictors. Additionally, we provide distinct R^2_{OS} values for NBER-dated expansions and recessions. The out-of-sample period considered spans from 2001:01 to 2020:12. In the last row of Panel C, we pool the individual forecasts from combined predictors by taking an arithmetic mean. Based on the Clark and West's (2007) test, * and ** indicate significance at the 10% and 5% levels for the positive R^2_{OS}, respectively.

Sub-periods	Overall	Expansions	Recessions
Panel A: Option bounds			
Bound	−1.450	0.768	−5.560
Bound + Avg slackness	−2.533	0.162	−7.527
Bound (OLS)	−8.445	−5.219	−14.425
Panel B: Traditional predictors			
$x_t = IS_{BW}$	−1.443	2.000**	−7.825
$x_t = IS_{HJTZ}$	−0.268	0.307*	−1.332
$x_t = SSI$	−0.776	−0.146	−1.944
Panel C: Combined predictors			
$x_t = IS_{BW}$	−0.788	2.321**	−6.549
$x_t = IS_{HJTZ}$	0.809**	0.814*	0.800
$x_t = SSI$	0.695**	−0.084	2.139
Pooling	1.417**	2.135**	0.088

Table 6: R^2_{OS} statistics (in percent) for longer-horizon forecasts

This table displays out-of-sample R^2 statistics (R^2_{OS}) as percentages for 3-, 6-, and 12-month market excess return forecasts utilizing option bounds, stock market predictors, and combined predictors. In the last row of Panel C, we pool three individual forecasts from three combined predictors by taking an arithmetic mean. The out-of-sample period is 2001:01–2020:12. Based on the Clark and West's (2007) test, ** and *** indicate significance at the 5% and 1% levels for the positive R^2_{OS}, respectively.

Horizons	3-month	6-month	12-month
Panel A: Option bounds			
Bound	0.799	5.582***	7.329***
Bound + Slackness	−2.693	−2.107	−13.936
Bound (OLS)	−7.572	8.250***	2.229**
Panel B: Traditional predictors			
$x_t = IS_{BW}$	1.181***	6.227***	13.252***
$x_t = IS_{HJTZ}$	2.300***	5.861***	12.594***
$x_t = SSI$	3.471***	9.445***	12.671***
Panel C: Combined predictors			
$x_t = IS_{BW}$	−1.136	2.341***	8.478***
$x_t = IS_{HJTZ}$	2.435***	6.421***	11.907***
$x_t = SSI$	4.055***	9.430***	15.711***
Pooling	6.065***	14.105***	26.975***

Table 7: R^2_{OS} statistics (in percent) for alternative function forms of $k(x_t)$

This table reports out-of-sample R^2 statistics (R^2_{OS}) in percent for 1-, 3-, 6-, and 12-month market excess return forecasts based on the combined predictors such that

$$b_t\,[k(x_t)] = k(x_t)\left[(T-t)R_{f,t}\mathrm{SVIX}^2_{t\to T} + R_{f,t}\right] - R_{f,t},$$

where $R_{f,t}$ is the gross return on an risk-free asset, $\mathrm{SVIX}^2_{t\to T}$ is computed from market prices of index options as in Martin (2017), and x_t is one of the sentiment variables. We consider two alternative forms for $k(x_t)$ as indicated by the panel headings. In the last row of each panel, we pool the individual forecasts from combined predictors by taking an arithmetic mean. The out-of-sample period is 2001:01–2020:12. Based on the Clark and West's (2007) test, *, ** and *** indicate significance at the 10%, 5% and 1% levels for the positive R^2_{OS}, respectively.

Horizons	1-month	3-month	6-month	12-month
	Panel A: Exponential form $k(x_t) = \exp(a+bx_t)+c$			
$x_t = IS_{BW}$	−0.879	−1.282	2.043***	8.041***
$x_t = IS_{HJTZ}$	0.883**	2.348***	5.885***	11.297***
$x_t = SSI$	0.768**	3.360***	8.528***	15.465***
Pooling	1.443**	5.783***	13.489***	26.456***
	Panel B: Linear form $k(x_t) = (a+bx_t)$			
$x_t = IS_{BW}$	−0.285	1.585***	6.087***	14.390***
$x_t = IS_{HJTZ}$	0.875**	2.690***	5.880***	13.904***
$x_t = SSI$	0.448**	3.575***	8.764***	14.387***
Pooling	1.325**	6.245***	13.746***	26.065***

Table 8: R^2_{OS} statistics (in percent) by using both SVIX and VIX

This table reports out-of-sample R^2 statistics (R^2_{OS}) in percent for 1-month market excess return based on the combined predictor. The combined predictor takes the form such that,

$$b_t\left[k(x_t)\right] = \begin{cases} k(x_t)\left[(T-t)R_{f,t}\mathrm{SVIX}^2_{t\to T} + R_{f,t}\right] - R_{f,t}, & \text{if} \quad t \geq 1996 \\ k(x_t)\left[(T-t)R_{f,t}\mathrm{VIX}^2_{t\to T} + R_{f,t}\right] - R_{f,t}, & \text{if} \quad t < 1996 \end{cases},$$

where $R_{f,t}$ is the gross return on an risk-free asset, $\mathrm{SVIX}_{t\to T}$ is computed from market prices of index options as in Martin (2017), $\mathrm{VIX}_{t\to T}$ is CBOE VIX index, x_t is one of the three sentiment variables including Baker and Wurgler's (2006) sentiment index (IS_{BW}), Huang et al.'s (2015) sentiment index (IS_{HJTZ}) and Rapach, Ringgenberg, and Zhou's (2016) short interest index (SSI); and $k(x_t) = exp(a + bx_t)$. In the last row of Panel C, we pool the individual forecasts from combined predictors by taking an arithmetic mean. The full sample spans from 1990 to 2020, and we consider four out-of-sample testing periods as indicated by the column headings below. Based on the Clark and West's (2007) test, *, ** and *** indicate significance at the 10%, 5% and 1% levels for the positive R^2_{OS}, respectively.

Out-of-sample periods	2001−2020	2006−2020	2011−2020
Panel A: Option bounds by SVIX and VIX			
Bound	−1.450	−2.214	0.741
Bound + Slackness	−2.024	−2.268	2.436
Bound (OLS)	−6.761	−8.018	−2.277
Panel B: Traditional predictors			
$x_t = IS_{BW}$	−0.513	−0.128	1.979*
$x_t = IS_{HJTZ}$	0.730*	0.525	0.87
$x_t = SSI$	0.303**	1.798**	0.749*
Panel C: Combined predictors			
$x_t = IS_{BW}$	−0.061	−0.176	1.822*
$x_t = IS_{HJTZ}$	1.625**	−0.103	1.528*
$x_t = SSI$	1.498**	2.027**	1.279*
Pooling	1.686**	0.825*	1.872*

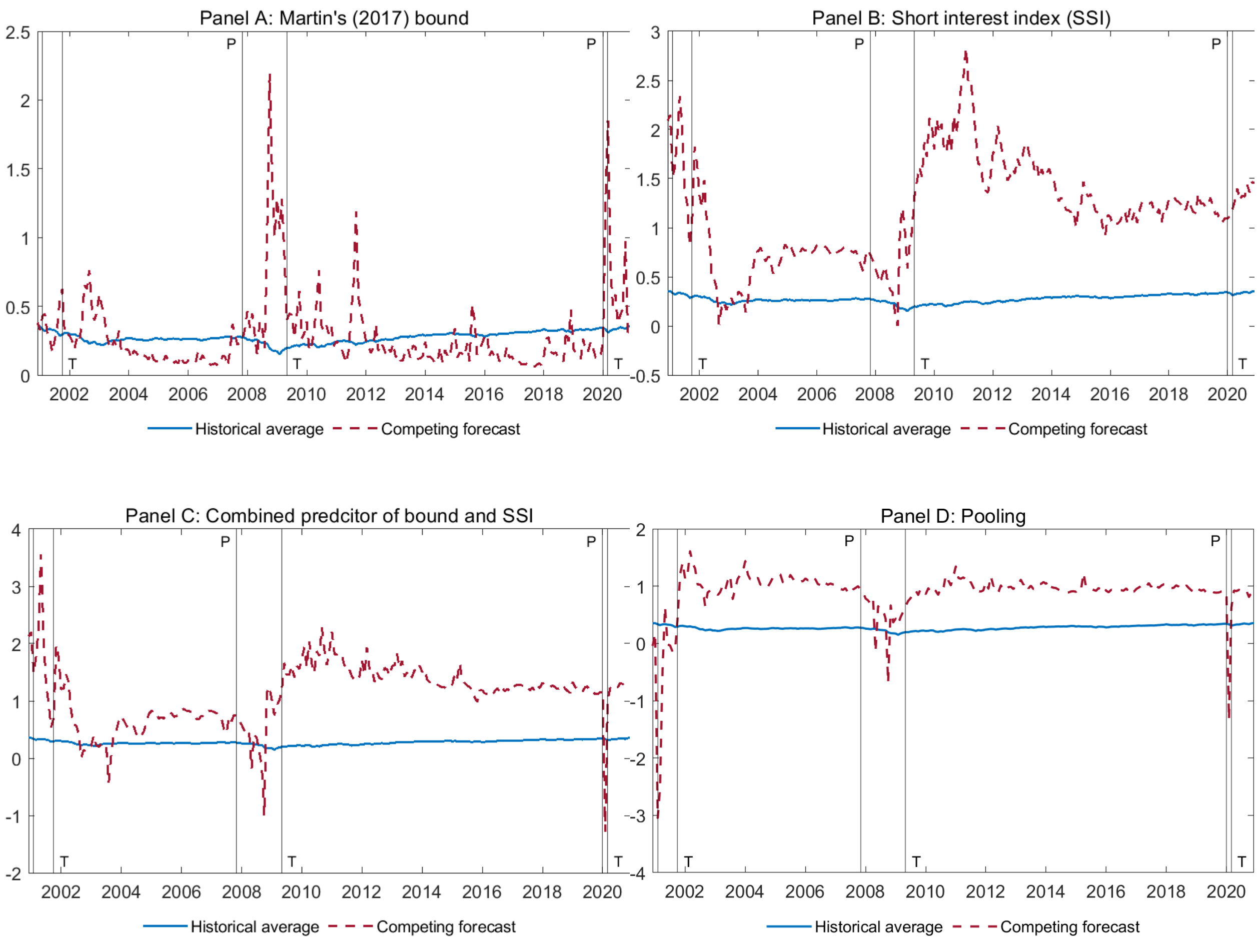

Figure 1: Out-of-sample forecasts of market excess return

This figure illustrates out-of-sample market excess return forecasts (expressed as percentages) for the period 2001:01 to 2020:12. The forecasts are generated using the predictor (or method) specified in the respective panel heading, along with the historical average benchmark. The term "Pooling" refers to the aggregation of three individual forecasts from three combined predictors using an arithmetic mean. Vertical lines denote NBER-dated business-cycle peaks (P) and troughs (T).

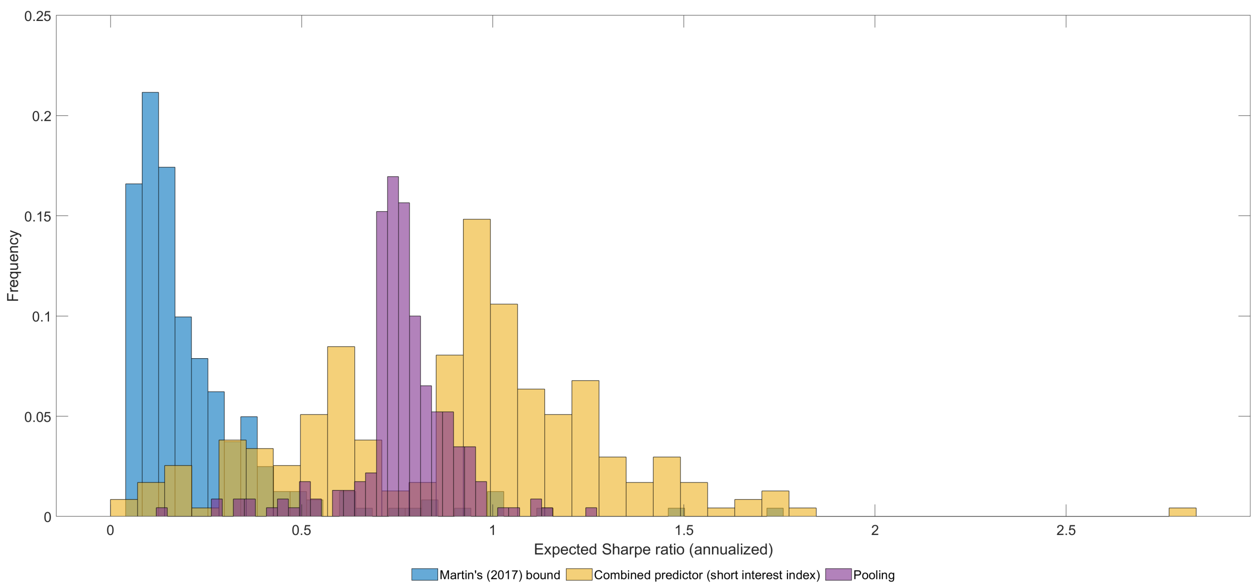

Figure 2: Distribution of ex-ante Sharpe ratios

This figure plots the *ex-ante* Sharpe ratio perceived by the investor at time t when forecasting the return for the subsequent period within the span of 2001:01 to 2020:12 . The formula for calculating the Sharpe ratio is as follows:

$$SR_{j,t} = \frac{\hat{r}_{j,t+1}}{\hat{\sigma}_{t+1}},$$

where $\hat{r}_{j,t+1}$ is the forecasted excess return on the S&P 500 index at time t based on either Martin (2017), combining bound and short interest index, or by pooling three individual forecasts from combined predictors. $\hat{\sigma}_t$ is the forecasted volatility, computed from historical return using a 5-year rolling window.

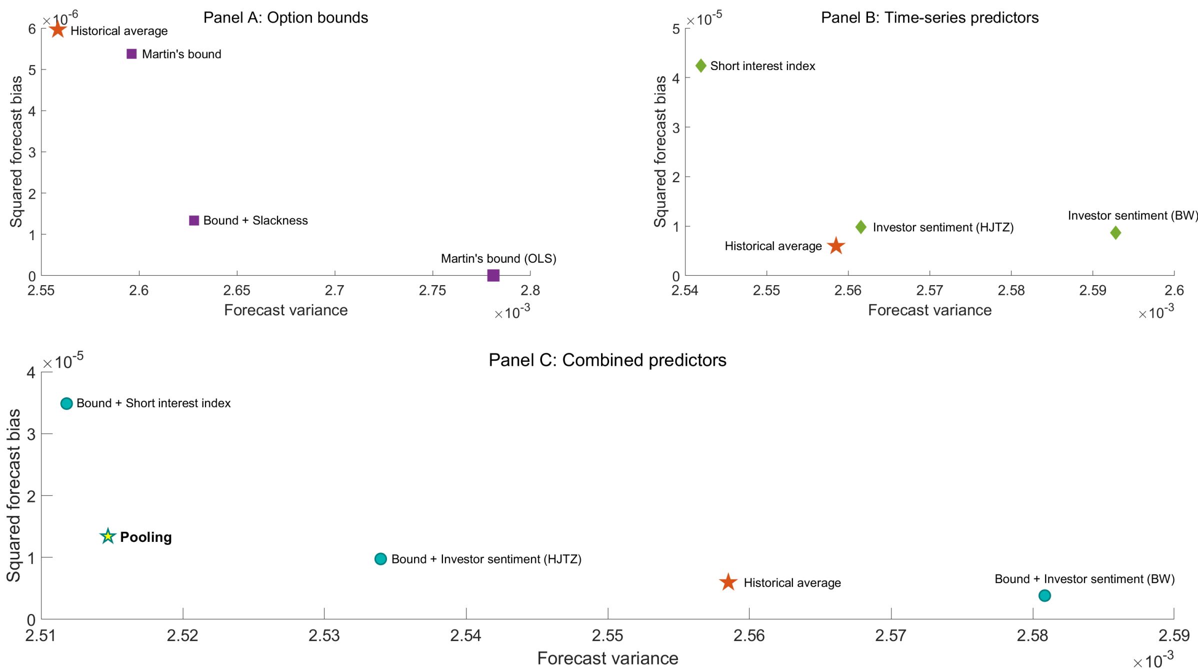

Figure 3: Bias-Variance decomposition

This figure presents the bias-variance decomposition of the Mean Squared Prediction Error (MSPE) using forecasts derived from option bounds, stock market predictors, and combined predictors during the out-of-sample period from 2001:01 to 2020:12.

Figure 4: Log cumulative excess returns for portfolios out-of-sample

Each panel depicts the log cumulative excess return for a portfolio constructed using the market excess return forecast in the panel heading and the historical mean benchmark forecast for the period 2001–2020. Vertical lines indicate NBER-dated business-cycle peaks (P) and troughs (T).

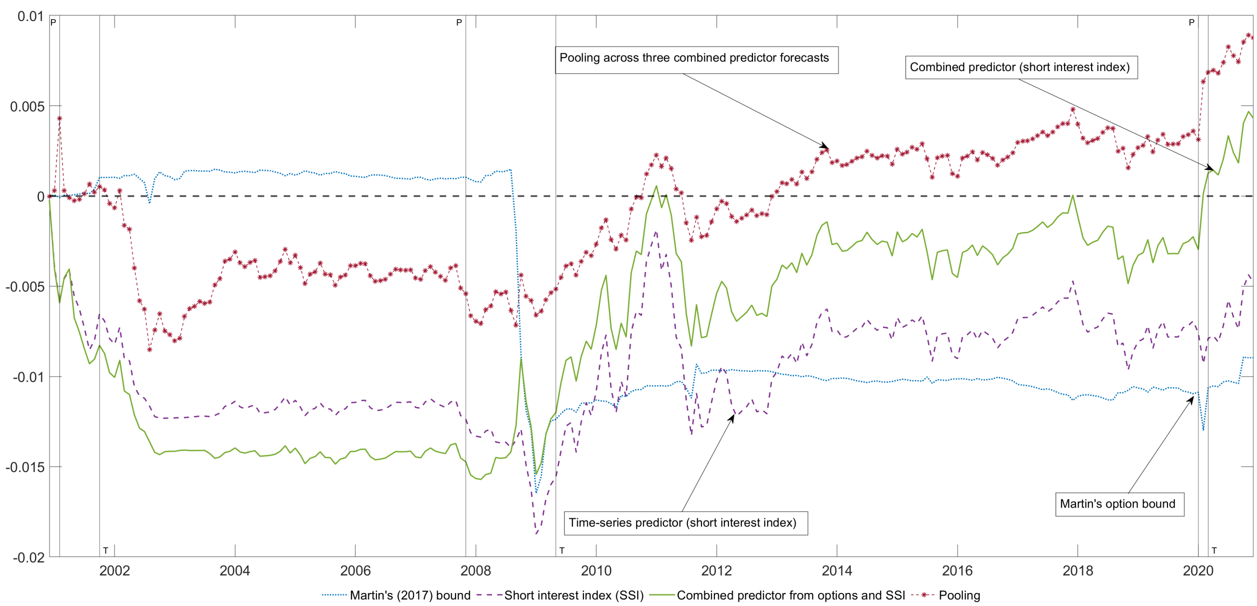

Figure 5: Cumulative error difference relative to historical average

This figure plots the cumulative square prediction error for 2001:01–2020:12 such that

$$\text{square error difference} = (r_t - \bar{r}^{HA}_{t|t-1})^2 - (r_t - \hat{r}_{t|t-1})^2,$$

where $\bar{r}^{HA}_{t|t-1}$ is the historical average (HA), r_t is the realized market excess return, and $\hat{r}_{|t-1}$ denotes the forecast derived from either an option bound, a time-series (stock market) predictor, or a combined predictor. Pooling denotes the aggregation of the three individual forecasts from the three combined predictors using an arithmetic mean. Vertical lines indicate NBER-dated business-cycle peaks (P) and troughs (T).

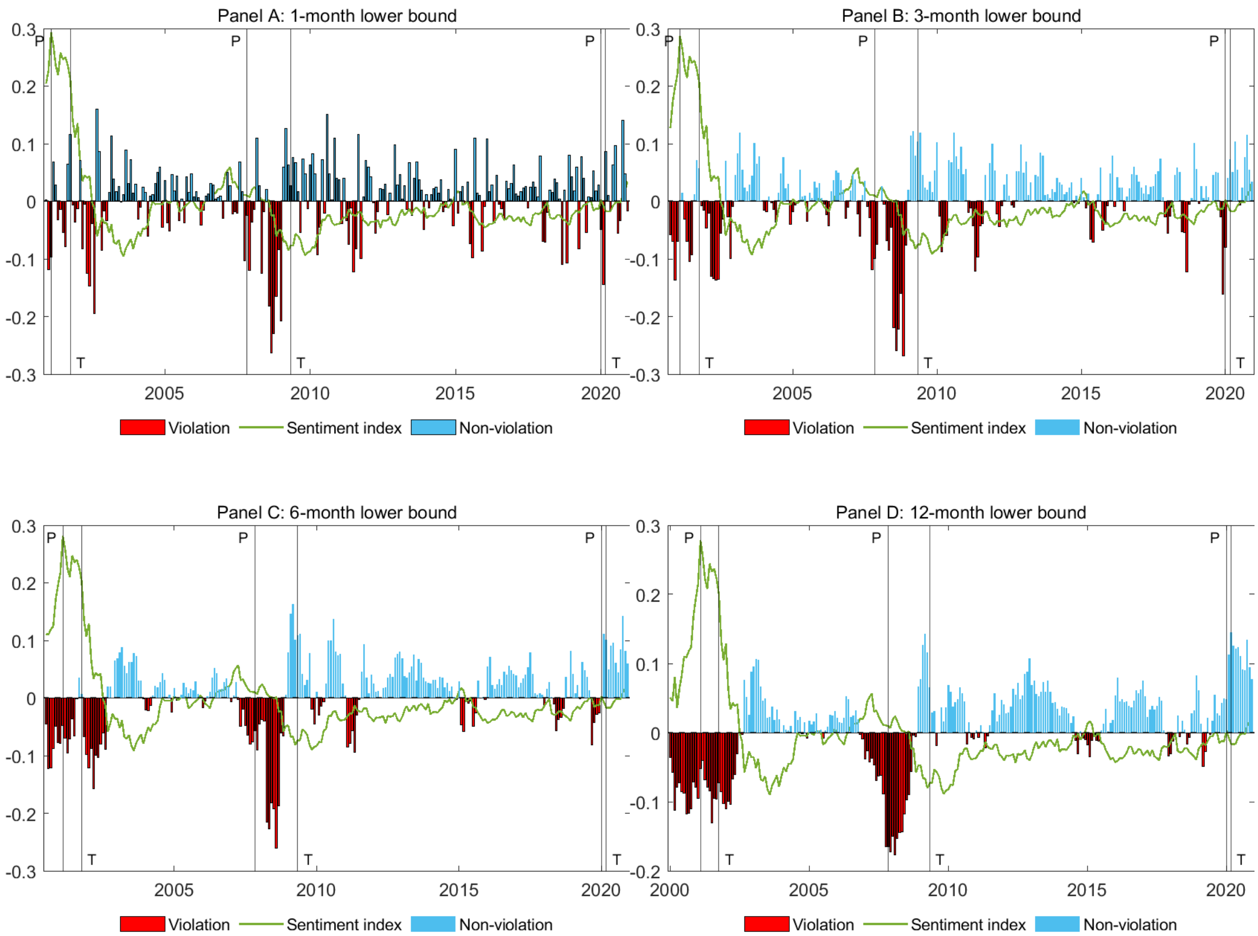

Figure 6: Violation of the lower bound by Martin (2017)

This figure plots the violation of the inequality in Martin (2017) for 2001:01–2020:12 such that

$$\mathbb{E}_t(R_T) - R_{f,t} \geq \frac{1}{R_{f,t}} Var_t^{\mathbb{Q}}(R_T).$$

The red bar signifies the violation of the above inequality, while the blue bar indicates its non-violation. The green line represents investor sentiment by Baker and Wurgler (2006), which has been normalized to 1 unit. Vertical lines denote NBER-dated business-cycle peaks (P) and troughs (T).

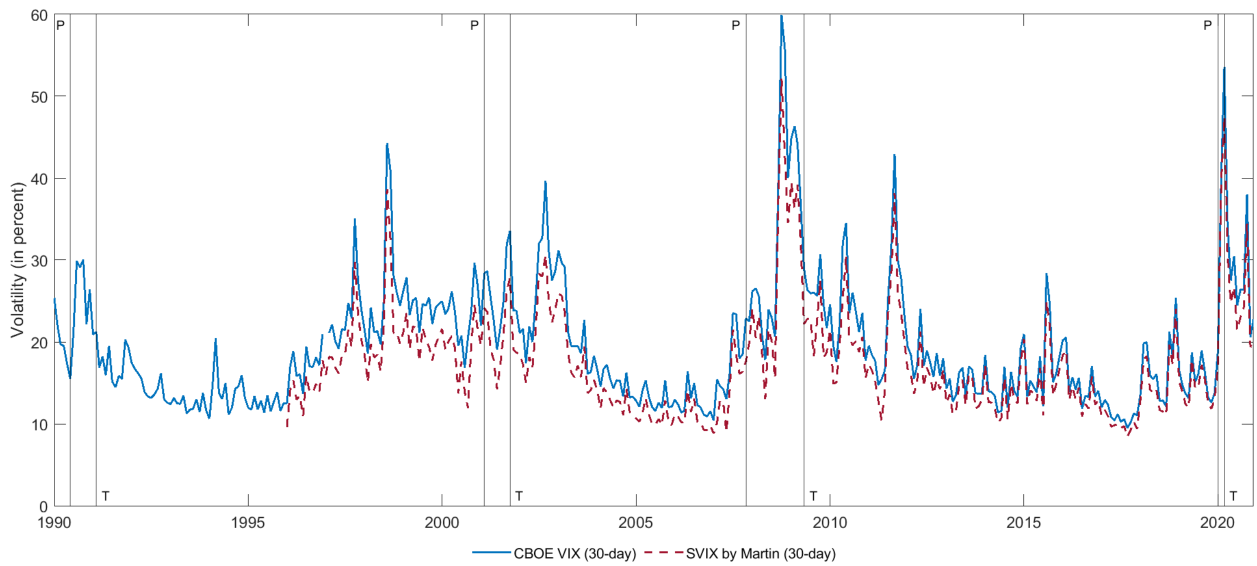

Figure 7: CBOE VIX and SVIX (in percent)

This figure illustrates the CBOE VIX index and Martin (2017) SVIX index (both expressed as percentages) from January 1990 to December 2020. Vertical lines are used to indicate NBER-dated business-cycle peaks (P) and troughs (T).