Expected and Realized Returns on Volatility

Guanglian Hu — University of Sydney

Kris Jacobs* — University of Houston

Abstract

Expected returns on market volatility, which can be obtained from VIX futures prices in closed form using standard models, positively predict subsequent realized volatility returns. Volatility returns are negative on average. Following increases in volatility, expected volatility returns and subsequent realized volatility returns become more negative. Because realized volatility returns are negatively correlated with index returns, expected volatility returns also negatively predict S&P 500 index returns, but these results are less significant. The results are robust to a wide range of variations in the empirical setup and to small-sample biases.

*We would like to thank the anonymous referee, Torben Andersen, Francisco Barillas, David Bates, Tim Bollerslev, Ing-Haw Cheng, Christian Dorion, Hitesh Doshi, Jan Ericsson, Mathieu Fournier, Alex Jeanneret, Paola Pederzoli, Richard Philip, Gustavo Schwenkler, Ivan Shaliastovich, George Tauchen, Susan Thorp, Joakim Westerholm, participants in seminars at the AFA, CICF, EFA, FIRN, SoFiE, the University of Sydney, NUS, UNSW and the Virtual Derivatives Workshop, and especially Andrew Patton and Federico Bandi for helpful comments and conversations.

I. Introduction

The literature on the predictability of stock returns is voluminous, and the evidence on the predictability of index returns is an important part of this literature. Cochrane (2001) and Lettau and Ludvigson (2010) conclude that a small fraction of the variation in monthly index returns can be predicted by variables such as the price-dividend ratio, the price-earnings and dividend payout ratios, short-term interest rates, and term and default spreads. Subsequent research has suggested additional predictors, such as the variance risk premium (Bollerslev et al. (2009)), volatility of volatility (Huang et al. (2019)), and tail risk (Bollerslev and Todorov (2011); Andersen et al. (2020)).

This paper contributes to this literature on the predictability of market returns. However, rather than focus on predicting the returns on the market index, we start by investigating monthly returns on market volatility. Unlike volatility itself, which is highly persistent, we focus on realized returns on volatility based on positions in VIX futures contracts. We then investigate if these returns can be forecast using past data on VIX futures. This strategy is implementable in real time, similar to a strategy that attempts to forecast index returns using past index returns and index futures. Moreover, unlike the literature that uses variables such as price-dividend ratios to forecast index returns, we only rely on market prices measured at the same time and we avoid the use of a low-frequency variable such as dividends.

It is well-known that average realized returns on market volatility are negative, indicating that a short volatility position is on average profitable, but that the return series contains large positive outliers. We confirm this result in our sample. We then use various reduced-form models of VIX dynamics to model VIX futures and expected returns on VIX futures contracts with different maturities and holding periods. Expected volatility returns on VIX futures are also negative on average and positively correlated with the time series of subsequent realized returns.

Because of the negative correlation between index returns and volatility and/or VIX, known as the leverage effect, our results also have implications for the predictability of (S&P 500) index returns. Specifically, the positive relation between expected volatility returns and realized subsequent volatility returns suggests that expected volatility returns negatively predict subsequent index returns. We find that our results on forecasting volatility returns indeed carry over to forecasting index returns, but the results from these predictive regressions are less statistically significant.

There is considerable debate in the literature about the evidence from predictive regressions, partly because the predictable component of one-month returns is small. Several studies have advocated the use of predictive regressions using long-horizon returns to confirm the results of short-horizon regressions (Fama and French (1988); Cochrane (2001), 2011; Bollerslev et al. (2009)). Specifically, in the presence of a persistent predictor, predictability of the one-month return manifests itself in the pattern of regression slopes and R^2s at longer horizons. We therefore not only report on predicting next-month realized volatility and index returns, we also study the predictability of monthly VIX futures returns compounded over longer horizons, up to five years. We find that these predictive regressions indeed confirm our results. When we adopt the forward-backward predictive regressions of Bandi et al. (2019), we obtain hump-shaped patterns in predictive R^2s for both realized volatility and index returns, with R^2s peaking at the 24-month horizon. This finding suggests the presence of different components in expected and realized volatility returns that operate at different frequencies. We cannot reject the null hypothesis of equilibrium generated predictability (EGP), which tests if the autocorrelation in the predictor is consistent with the pattern of the loadings of the predictive regressions as a function of the forecast horizon, as would be expected in an equilibrium setup (Eraker (2025)).

Because the literature has shown that biases in predictive regressions (Stambaugh (1999)) may worsen with longer horizons (Boudoukh et al. (2022)), we show that biases in long-horizon predictive regressions for our application are very small compared to biases in forecasts based on the dividend-price ratio. These findings are consistent with Bollerslev et al. (2014). We also provide evidence on the size and power properties of predictive regressions in finite samples based on parameterizations consistent with our empirical setup.

We demonstrate that our results are robust to a wide range of variations in the empirical setup, such as the implementation of the predictive regressions, the model for the VIX, the measure of expected returns, and the contract maturity. We also show that the expected volatility return retains predictive power in the presence of other predictors, such as the slope of the VIX term structure (Johnson (2017)), the market variance risk premium and variance-of-variance (Bollerslev et al. (2009)), and the tail factor (e.g., Andersen et al. (2015); Bollerslev and Todorov (2011)). We use simple univariate sorts to demonstrate how the information in these predictors differs from the information in expected volatility returns and we discuss the relation between our results and the findings of Cheng (2018) on the VIX premium. Lastly, we show that return predictability can be further improved by using a predictor that combines expected volatility returns constructed from contracts with different maturities.

Aside from the literature on the predictability of market returns, our work is most closely related to the literature on market variance risk and the variance risk premium. Much of this literature focuses on measuring the market variance risk premium (Carr and Wu (2009); Bakshi and Kapadia (2003a); Bondarenko (2014)) as well as the variance risk premium on stocks and its relation to the market variance risk premium (Bakshi and Kapadia (2003b); Carr and Wu (2009); Driessen et al. (2009); Duarte et al. (2022)). Martin (2017) studies the relation between the market variance risk premium and the market equity risk premium, while Bollerslev et al. (2009) and Pyun (2019) study predictive regressions of the market risk premium on the variance risk premium. Like our paper, Eraker and Wu (2017) and Cheng (2018) study returns on VIX futures. Cheng (2018) focuses on the puzzle that the VIX premium does not increase with measures of risk. Eraker and Wu (2017) analyze if the distribution of VIX futures returns is consistent with dynamic equilibrium. We instead focus on expected volatility returns and ask if they can predict realized volatility returns. We find that volatility risk is priced in VIX futures markets, and that the price fluctuations correctly anticipate future market volatility returns and the corresponding S&P 500 returns. Moreover, the time series of expected volatility returns contains information that is distinct from several other risk measures suggested by theory and studies in the existing literature, such as the variance risk premium, tail risk, and the volatility of volatility.

The paper proceeds as follows. Section 2 discusses the VIX futures data. Section 3 discusses the computation of realized and expected returns on volatility. Section 4 presents the baseline empirical results. Section 5 discusses several extensions and robustness exercises, Section 6 studies statistical biases in (long-horizon) predictive regressions, and Section 7 concludes.

II. Data

We obtain daily VIX data from the Chicago Board Options Exchange (CBOE). In 1993, the CBOE developed the VXO volatility index, which was based on S&P 100 at-the-money options. In 2003, in light of structural changes in index option markets and advances in academic research, the CBOE refined the methodology and developed a new volatility index, the VIX, which is computed as a weighted average of S&P 500 option prices.[1]

We download data on VIX futures from the CBOE website, which includes daily settlement prices, open and close prices, high and low prices, volume and open interest. The sample period for the futures data is from March 26, 2004, when futures contracts on the VIX were introduced and started trading, to November 23, 2022. Our empirical analysis reports on futures contracts with different maturities. In the early part of the sample, there are typically four futures contracts listed every day. After October 2006, typically nine contracts are listed every day. We only include VIX futures contracts which expire on the Wednesday that is exactly thirty days prior to the third Friday of the following month. If that Wednesday or the Friday 30 days later is an exchange holiday, the VIX futures settle on Tuesdays instead.[2] Initially, VIX futures prices were quoted as the VIX times 10 and the contract multiplier was $100. Starting on March 26, 2007, the CBOE Futures Exchange modified the contract specification by dividing futures prices by 10 and increasing the contract multiplier to $1000, which brings the price in line with the underlying VIX index while leaving the contract size unchanged. We therefore rescale all futures prices prior to March 26, 2007 accordingly. We apply several data filters. We require futures contracts to have valid settlement prices, positive volume and positive open interest.

The sample period we use for the VIX index itself is longer, from January 2, 1990 to November 23, 2022, and we use daily data in estimation. Panel A of Figure 1 plots the VIX index for this period. The VIX index exhibits substantial variation over time and is strongly mean reverting. In our sample period, it peaks at the start of the Covid-19 pandemic, closing at a historical high of 82.69 on March 16, 2020. In the financial crisis, it

[1] See Whaley (2009) for a detailed account of the development of the VIX.

[2] Starting in July 2015, VIX futures that expire on other Wednesdays (the so-called weeklies) are also available for trading. We do not include these in the sample.

reaches a high of 80.96 on November 20, 2008. Panel A of Table 1 reports summary statistics for the VIX for two sample periods: the January 1990 to November 2022 period, as well as March 2004 to November 2022, which corresponds to the period for which VIX futures data are available. The descriptive statistics are similar in these two samples, but the VIX has a slightly higher mean, smaller standard deviation, smaller (positive) skewness, and smaller kurtosis in the longer sample.

***********Figure 1 about here*************

***********Table 1 about here*************

Panel B of Table 1 reports summary statistics for VIX futures prices.[3] VIX futures prices are higher than the VIX on average, indicating that long investors pay a premium. Furthermore, VIX futures prices increase monotonically with maturity, from 20.20 for the 1-month contract to 22.35 for the 9-month contract. This upward-sloping term structure, often referred to as the contango trap, is the reason why rolling the VIX futures contract is associated with substantial losses (e.g. Eraker and Wu (2017); Whaley (2013)). Panel B of Figure 1 plots the evolution of the VIX futures term structure over the sample period. The shape of the term structure changes considerably over time. During normal times, VIX futures prices tend to be low and have an upward-sloping term structure. In times of market turbulence, such as the financial crisis, the futures curve tends to increase and may become inverted or hump-shaped.

Panel A of Figure 2 reports the average daily trading volume and open interest per contract as a function of maturity. The majority of trading in VIX futures takes place at the short end of the term structure, with much larger volume and open interest for

[3] These summary statistics rely on linear interpolation on each day in the sample to calculate VIX futures prices with constant maturities up to nine months. However, throughout the paper our empirical analysis is exclusively based on the market prices of the VIX futures.

short-maturity contracts. Panel B shows the average daily trading volume and open interest per contract over time. The past decade has witnessed tremendous growth in the VIX futures market. The average daily volume per maturity was 33390 contracts in 2018, more than 250 times higher than in 2004 and more than 50 times higher than in 2008.

***********Figure 2 about here*************

III. Realized and Expected Returns on Volatility

We first compute ex-post realized returns from investing in VIX futures for different investment horizons. We then discuss how to compute expected returns to holding VIX futures.

A. Realized Returns on Volatility

In this section, we analyze ex-post realized returns from investing in VIX futures. Following the existing literature (e.g., Hong and Yogo (2012); Singleton (2013); Cheng (2018)), we measure the return on a futures contract as the return on a fully collateralized position.[4] Specifically, consider a VIX futures contract at time t that will expire at time T and currently trades at a price of $F_t(T)$. The simple return $R_{t,T}$ from holding this contract from time t to expiration is given by:

$$R_{t,T} = \frac{VIX_T}{F_t(T)} - 1 \tag{1}$$

[4]In our baseline implementation, we assume that the collateral does not earn a return. Results are virtually unchanged if we instead assume that the collateral earns the risk-free rate, following Gorton and Rouwenhorst (2006) or Hong and Yogo (2012), because the risk-free rate is very small compared to the changes in the futures prices. Table A1 in the Online Appendix reports these results.

where VIX_T is the settlement value of the contract at maturity.[5] A long position in VIX futures provides a hedge against volatility risk because it has a positive payoff when the VIX increases. During our sample period, 219 VIX futures contracts mature. For each of these contracts, we establish a long position one month, two months, three months, four months, and five months prior to the expiration date and compute the corresponding hold-to-maturity returns.[6]

Panel A of Table 2 reports the summary statistics for the resulting VIX futures returns.[7] To make the returns comparable, we scale all hold-to-maturity returns to monthly returns. Consistent with our observations on the term structure of VIX futures prices, the average returns to holding VIX futures are negative regardless of the horizon, which indicates that long investors on average pay a premium for buying VIX futures. Moreover, the average return decreases (in absolute value) as a function of time-to-maturity. Our finding that the volatility risk premium is negative and economically large, especially at short maturities, confirms existing evidence regarding the returns on market volatility, see for instance Andries et al. (2023), Cheng (2018), Dew-Becker et al. (2017), and Eraker and Wu (2017).

***********Table 2 about here*************

Panel A of Table 2 indicates that a long position in volatility using the 1-month contract incurs an average loss of -3.7% per month with an annualized Sharpe ratio of -0.405 over our sample period.[8] Panel A of Figure 3 plots the time series of 1-month hold-to-maturity VIX futures returns. The realized return on a long volatility position is

[5] VIX_T is determined by a Special Opening Quotation (SOQ) of the VIX, calculated from a sequence of opening prices of the underlying SPX index options.

[6] We implement these holding periods using 21, 42, 63, 84 and 105 trading days.

[7] The number of observations is less than 219 for longer horizons because the early part of the sample contains months when only futures contracts with short maturities are trading.

[8] Cheng (2018) reports a return of -3.5% per month with an annualized Sharpe ratio of -0.68 for short

negative most of the time, but occasionally it results in large positive payoffs. For example, in September 2008, during the financial crisis, a long position in a one-month VIX futures contract resulted in a positive return of 120%. In March 2020 (the beginnning of the Covid-19 pandemic), the return to a long position was 354%.

Panel B of Table 2 provides additional insight into the term structure of volatility returns. We decompose the average 5-month hold-to-maturity return into five consecutive monthly holding period returns. Consider a VIX futures contract that expires in June (month t). We decompose the return of holding this contract from January to June (t-5, t) into five monthly returns: January to February (t-5, t-4), February to March (t-4, t-3), March to April (t-3, t-2), April to May (t-2, t-1), and May to June (t-1, t). Consistent with the findings in Panel A, Panel B shows that VIX futures have on average more negative returns when they are closer to maturity.

Panel C reports summary statistics for the one-month holding period returns on VIX futures with different maturities. Each month t, we collect VIX futures that expire in month $t+1$, $t+2$, ..., $t+5$, and we report on the monthly returns to holding those contracts from month t to month $t+1$. We establish a long position in VIX futures exactly one month prior to the futures expiration date in month $t+1$.[9] For the front contract, the resulting returns are identical to the one-month holding-to-maturity returns analyzed in Panel A, with an average of -3.7%. For other maturities, these are monthly holding period returns. Consistent with our findings in Panels A and B, we find that the VIX futures holding period returns are on average negative and that they exhibit a strong negative

maturities based on the March 2004 to November 2015 sample period. The descriptive statistics in Eraker and Wu (2017) imply a return of -3.3% per month with an annualized Sharpe ratio of -0.61 for the one-month maturity based on the January 2006 to May 2013 sample period. The Sharpe ratio in our sample is somewhat lower due to the inclusion of the Covid-19 period.

[9] It is possible to extend this analysis up to $t+9$, but the longest-maturity contracts are less liquid and have more missing observations.

relationship with maturity: the average return decreases monotonically from -3.7% for the 1-month contract to -0.4% for the 5-month contract. However, despite these differences, the holding period returns on VIX futures contracts are highly correlated across maturities. For example, the correlations between the return on the 1-month futures contract and the returns on 2-, 3-, 4-, and 5-month contracts are 0.95, 0.94, 0.93, and 0.92 respectively.

B. Computing Expected Returns on Volatility

We now discuss the computation of *expected* returns to holding VIX futures. This computation requires us to make some choices regarding implementation. First, note that the futures price $F_t(T)$ is known at time t and consider the conditional expectation of the return in equation (1):

$$ER_{t,T}^V = \frac{E_t^P(VIX_T)}{F_t(T)} - 1. \tag{2}$$

where the superscript V indicates that this is a return on a long position in volatility, and the superscript P is used to denote the expectation under the physical measure. Equation (2) indicates that we need to make an assumption regarding the dynamics of the VIX under the physical measure.

Second, note that since the VIX itself is not tradeable, the price of VIX futures cannot be pinned down by cost of carry. However, no-arbitrage implies that the fair value of a VIX futures contract is the expected value of the VIX at maturity under the risk neutral measure, $F_t(T) = E_t^Q(VIX_T)$, and therefore equation (2) can also be written as:

$$ER_{t,T}^V = \frac{E_t^P(VIX_T)}{F_t(T)} - 1 = \frac{E_t^P(VIX_T)}{E_t^Q(VIX_T)} - 1. \tag{3}$$

Equation (3) indicates that the expected futures return can be computed using the observed futures price, or alternatively using a model for the VIX dynamic under the risk-neutral measure. We report results based on expected returns computed using the observed futures price.[10]

C. The Dynamics of the VIX

Regardless of how we implement the denominator in equation (3), our empirical strategy requires the specification of a model for the VIX under the physical measure. We report results based on several models. First consider a very simple model, which assumes that the VIX follows a square-root process under the physical measure:

$$dVIX_t = \kappa(\theta - VIX_t)dt + \sigma\sqrt{VIX_t}dW \tag{4}$$

where θ is the long-run mean of the VIX and κ is the mean-reversion parameter. Given this dynamic, the P-expectation in equation (3) is straightforward to compute:

$$E_t^P(VIX_T) = \theta + (VIX_t - \theta)e^{-\kappa(T-t)}. \tag{5}$$

Given the mean-reverting property of the VIX in equation (4), the model states that the expected value of the future VIX is a weighted average of the long-run mean and the current VIX. This model is admittedly overly simplistic and we therefore consider three alternative models for the VIX: the ARMA(2, 2) model used in Cheng (2018), the heterogeneous autoregressive (HAR) model of Corsi (2009) adapted for the VIX, where the predictors are the average VIX over the past month, week and day, and finally a two-factor model (e.g., Lee and Engle (1999); Alizadeh et al. (2002); Christoffersen et al. (2009); Engle and Rangel (2008)). This model assumes that the VIX mean reverts to a stochastic mean which itself follows a square root process. Online Appendix A provides more details on this model.

[10]In unreported results, we find that the alternative expected returns, computed using the expected value of the VIX under Q, also predict subsequent volatility and index returns.

We start by comparing the performance of these four dynamics for forecasting VIX itself. We evaluate the models' performance for three forecasting horizons: 1, 2, and 3 months ahead. Our implementation follows Bekaert and Hoerova (2014). Specifically, we divide the entire VIX sample (January 1990 to November 2022) into an in-sample period used to estimate the model parameters and an out-of-sample period used for evaluating forecasting performance. We consider three different sample splits: We use either 65%, 75% and 85% of data for estimating the parameters, and the remaining 35%, 25%, and 15% of the data is used for assessing out-of-sample forecasting performance.[11] As in Bekaert and Hoerova (2014), we estimate and evaluate the models using daily data and the parameters are not updated in the out-of-sample period. We consider various performance measures, including the root mean squared error (RMSE), mean absolute error (MAE), the R^2 of the Mincer-Zarnowitz regression of realized VIX values onto forecasted values in the out-of-sample period, and the average correlation (Corr) between a given model's forecast and the forecasts generated by the winning model based on each of the above three criteria. This latter measure provides insight into the economic proximity between different model forecasts.

Table 3 reports the results, with Panels A, B, and C reporting on the 65-35, 75-25, and 85-15 sample splits respectively, and Panel D reporting on the average performance across the three sample splits and the overall rankings. The overall rank is computed as the average rank across all four performance measures, with the best performing model receiving a rank of 1 and the worse performing model receiving a rank of 4.[12] Panel D indicates that for the purpose of forecasting the one-month ahead VIX, the one-factor

[11] This corresponds to in-sample periods ending in May 2011, August 2014, and December 2017, respectively.

[12] For example, when forecasting the one-month ahead VIX, in Panel D the OF model has rank of 1 for RMSE, 2 for MAE, 2 for R^2, and 1 for Corr. Consequently the average rank is 1.5 for the OF model.

model emerges as the best model overall, achieving an overall rank of 1.5 with the lowest RMSE, second lowest MAE, second highest Mincer-Zarnowitz regression R^2, and highest average correlation. For the purpose of forecasting the VIX two and three months ahead, the HAR model is the best overall performer, followed by the one-factor model, the two-factor model, and the ARMA(2,2) model. Unsurprisingly, the forecasting performance of all models drops significantly at two- and three-month horizons.

***********Table 3 about here*************

We conclude that a clearly superior model does not emerge from this forecast comparison. Related, the differences in model forecasting performance are relatively small in several dimensions. Rather than report on (a different) best-performing model for every forecast horizon, we therefore use the same model in our baseline analysis, and report on the other models in the robustness analysis in Section D.. For simplicity we use the one-factor model as the baseline model.

IV. Forecasting Volatility Returns and S&P 500 Returns

We first report our baseline results on forecasting volatility returns, and then our baseline results on forecasting S&P 500 returns, both based on the one-month contract. Next we report on returns for contracts with longer maturities.

A. Forecasting Realized Volatility Returns

We first discuss the details of the implementation of the expected volatility return used in our baseline analysis. We estimate the VIX dynamics and expected returns recursively using an expanding window to ensure there is no look ahead bias.[13] At each

[13]We can estimate the model parameters using the entire sample period, using VIX data from January 2, 1990 to November 23, 2022, but this approach is subject to data-snooping concerns: If the P-expectation at

time t, we use VIX data from January 1990 up to time t to estimate the parameters of the model in equation (4) via maximum likelihood by exploiting the fact that the transition density of the square-root process is available in closed form (Pearson and Sun (1994)). Appendix A provides additional details on the estimation of the one-factor model and the time-series properties of the VIX. We then use those parameters and equation (5) to compute the physical one-month expectation of the VIX index. Together with the one-month futures prices, this allows us to compute the one-month expected volatility return at time t, $(ER^V_{t,t+1})$.[14]

***********Figure 3 about here*************

The solid red line in Panels B and C of Figure 3 plots the resulting time series of 1-month hold-to-maturity expected returns on VIX futures using the square root model. Compared to the realized returns in Panel A, the time series of expected returns is more persistent. Panel B also plots the time series of the (squared) VIX.[15] The expected return becomes more negative following sharp increases in volatility. The scatter plot of the VIX against the baseline expected volatility returns in Panel A of Figure 4 confirms this pattern. Note that this stylized fact is related to but distinct from the one investigated by Moreira and Muir (2017, 2019), who document that the risk-adjusted return to a long S&P 500 position is lower following spikes in volatility.

***********Figure 4 about here*************

time t is computed using model parameters that are estimated using the entire sample, the model parameters will incorporate information from future returns, which may bias the results.

[14]This recursive implementation does not lead to a shorter sample, because the VIX time series is available starting in 1990, prior to the start of the VIX futures data (and therefore the expected return series).

[15]The time-series variation in the realized variance and therefore its relation to the expected volatility return is very similar.

Our baseline analysis uses these expected returns in forecasting regressions to predict future realized returns on the one-month VIX futures contract for different return horizons. The literature on forecasting index, stock and bond returns indicates that it may be beneficial to examine forecasting regressions at both short and long horizons, in the sense that the patterns in the regression coefficients and R^2s for the long-horizon regressions can be used to confirm the results of short-horizon regressions (see for instance Fama and French (1988); Cochrane (2001), 2011; Cochrane and Piazzesi (2005); Bollerslev et al. (2009), 2014; Kostakis et al. (2015)).[16] We therefore not only consider predicting next-month realized returns, we also study the predictability of monthly VIX futures returns compounded over longer horizons, up to five years (60 months), scaled by the forecast horizon. Specifically, we construct time series of h-month futures returns, and we report on $h = 1, 3, 6, 12, ..., 60$. These predictive regressions are thus given by:

$$\frac{1}{h}\sum_{i=1}^{h} r_{t,t+i}^{VIX} = \alpha_{t+h} + \beta_{t+h} ER_{t,t+1}^{V} + \epsilon_{t+h}, h = 1, 3, 6, 12, ..., 60 \tag{6}$$

where h is the index for the time horizon (in months), $\sum_{i=1}^{h} r_{t,t+i}^{VIX}$ is the cumulative future log return from holding VIX futures over horizon h calculated by summing up the monthly realized log returns from Panel A in Figure 3, and $ER_{t,t+1}^{V}$ is the expected (log) return from holding 1-month VIX futures to maturity. We estimate a single forecasting regression for each horizon (i.e., for a given forecast horizon h, the regression in equation (6) is estimated once). An alternative approach is to estimate a different forecasting regression at each time t and report on the average slopes and R^2s of these regressions. We report on this alternative implementation in Section C..

Panel A of Table 4 reports the results. The literature has emphasized potential statistical problems with overlapping returns in long-horizon regressions. We discuss these

[16] A related literature considers predictive regressions with long-horizon returns when predictors have multiple components that operate at different frequencies (Bandi et al. (2019); Bandi and Tamoni (2022)).

issues in Section VI.. Table 4 reports t-statistics according to Hansen and Hodrick (1980) and Hodrick (1992), because our findings in Section VI. confirm existing results that these t-statistics have better small-sample properties in the presence of overlapping returns compared to OLS or Newey-West t-statistics.[17]

***********Table 4 about here*************

The slope coefficients are positive at all horizons. The t-statistics are U-shaped as a function of the forecast horizon. The pattern across horizons differs somewhat between the Hodrick and Hansen-Hodrick t-statistics, but both indicate statistical significance at multiple horizons. The R^2 are much higher for long forecast horizons. As discussed by Cochrane (2001) for example, R^2s mechanically increase as a function of the forecast horizon in forecasting regressions with persistent regressors, provided that the loading on the predictor is nonzero to start with. The point estimate of the loadings decrease as a function of the horizon, consistent with a predictor that is mean-reverting. The pattern at long horizons is consistent with the first order sample autocorrelation of 0.85 for an AR(1) predictor.[18]

To provide additional intuition for the results in Table 4 and the high R^2s at long horizons, Panel B of Figure 4 scatter plots the baseline expected volatility returns against subsequent 5-year realized returns on VIX futures. We focus on long-horizon returns in

[17] Hodrick (1992) proposes several alternative computations of the standard error when forecasting long horizon returns. We use the standard errors usually referred to as Hodrick 1B. This implementation exploits covariance stationarity to remove the overlapping structure of the error terms, is guaranteed to be positive, and is valid under the null hypothesis of no predictability. The first standard error proposed in Hodrick (1992), labeled 1A, is the conditionally heteroskedastic counterpart to the standard errors of Hansen and Hodrick (1980). The resulting estimate of the covariance matrix is not guaranteed to be positive definite.

[18] At shorter horizons, the implied autocorrelation for an AR(1) predictor is lower, suggesting the presence of MA coefficients.

these plots because they are less variable, which leads to a higher R^2 and therefore a scatter plot that makes the relation with expected returns easier to gauge. More negative one-month expected returns anticipate multi-year lower (more negative) subsequent returns on VIX futures. Panel C of Figure 3 plots the corresponding time series of the one-month expected returns $x_{t,t+1}$ along with the subsequent 5-year realized returns on VIX futures.[19]

A potential concern is that our sample is relatively short. The financial crisis and the subsequent shock to returns represent a major event in this short sample. The same is true for the Covid-19 crisis, but presumably its effect on subsequent (long-horizon) returns is only partially realized as yet in our sample. We therefore need to be especially careful, because the variation in long-horizon returns and the resulting predictability might be largely due to this single event. Panel C of Figure 4 shows that this is not the case. We scatter plot five-year realized returns against expected returns and distinguish between returns that overlap with the financial crisis (in blue) and those that do not (in red). Clearly, the positive relation between expected and subsequently realized returns is present in both subsamples.

In summary, we document that expected volatility returns are statistically and economically significant predictors of realized returns on VIX futures. Next we discuss the relation between these findings and the existing literature on the predictability of S&P 500 index returns.

B. Forecasting S&P 500 Index Returns

The leverage effect refers to the negative correlation between (index) returns and innovations to (market) variance. The economic mechanism behind this negative correlation is strongly debated, but there is consensus about the stylized fact. Usually this

[19]Note that the time series of realized returns is shorter because it uses information up to month $t + 60$ at time t.

correlation is measured using daily, weekly, or monthly returns. We report the correlation between returns to holding the S&P 500 and the one-month VIX futures contract as a function of the time horizon. Panel C of Table 4 shows that this stylized fact also obtains for longer return horizons and that the magnitude of the correlation is remarkably constant across horizons. The correlations between the two returns are highly negative and range from -0.76 to -0.85.

We therefore conjecture that the finding in Section A. that expected returns can predict realized VIX futures returns may have implications for predicting index returns. To test this hypothesis, we use the baseline expected returns to predict subsequent index returns over different horizons:

$$\frac{1}{h}\sum_{i=1}^{h} r_{t,t+i}^{SP} = \alpha_{t+h} + \beta_{t+h} ER_{t,t+1}^{V} + \epsilon_{t+h}, h = 1, 3, 6, 12, ..., 60 \tag{7}$$

where $\sum_{i=1}^{h} r_{t,t+i}^{SP}$ denotes the log index return over horizon h, and $ER_{t,t+1}^{V}$ again is the baseline one-month log expected volatility return.

Panel D of Table 4 summarizes the results using the baseline implementation. Index returns are from CRSP. Consistent with our conjecture, expected returns on volatility indeed contain predictive information about future index returns. Specifically, the expected volatility returns negatively predict index returns and the predictive patterns mirror those in the volatility return predictive regressions in Panel A. However, on average across horizons the statistical significance and the R^2s are lower compared to the volatility returns in Panel A.[20] Another difference is that for these predictive regressions, the Hodrick t-statistics are generally a bit higher than the HH t-statistics. Panel C of Figure 3 plots the subsequent 5-year S&P 500 index returns (the broken red line) together with the expected

[20] Heston and Todorov (2024) report the related finding that the predictive power of the option-implied variance is higher for the volatility return than for the futures return in a cross-section of twenty futures contracts.

volatility return and the subsequent 5-year VIX futures return. The plot visually confirms the high correlations between the three time series.

C. Forecasting the Returns on Longer-Maturity VIX Futures Contracts

Section A. reports on predictive regressions using the returns on the 1-month VIX futures contract. Table 5 instead reports on predictive regressions based on monthly returns on the 2-, 3-, 4-, and 5-month VIX futures contracts compounded to longer horizons. Panel C of Table 2 reports descriptive statistics for these returns.

For completeness, we repeat the results for the 1-month contract from Table 4. The results for the 1-month contract generalize to other maturities. For example, at the three-year forecast horizon, the expected volatility return significantly predicts the returns to 2-, 3-, 4-, and 5-month contracts, with R^2s of 39.9%, 29.63%, 26.39%, and 26.03% respectively. This is not surprising because of the high correlations between VIX futures returns on contracts with different maturities discussed in Section A.. Also unsurprisingly, the loadings of the forecasting regressions are more stable across forecast horizons for contracts with longer maturities. Overall, the R^2s of the predictive regressions are somewhat higher for shorter-maturity contracts.

***********Table 5 about here*************

V. Alternative Predictors and Robustness Analysis

In this section, we present robustness results and explore how our results are related to the evidence in the existing literature, which almost exclusively considers predictive regressions for returns on the S&P 500. First we present results from univariate sorts. We then consider the predictive power of expected volatility returns in the presence of predictors used in the existing literature. We also investigate the robustness of our results with respect to the computation of expected returns and the implementation of the predictive regressions. We first report on results where we recursively estimate the predictive regressions. Next we use alternative models for the dynamics of the VIX to compute expected returns and we use a predictor that combines the information in futures contracts with different maturities. Finally we further explore the term structure of return predictability using the forward-backward aggregation approach of Bandi et al. (2019) and we test whether the term structure of return predictability is consistent with the persistence of the expected volatility returns (Eraker (2025)).

A. Univariate Sorts on Expected Returns

Simple (univariate) sorts can be useful to illustrate nonlinear relationships in the data. We sort the 219 monthly expected volatility returns used in Panel A of Table 4 and depicted in Panel B of Figure 3 into three groups/portfolios. Panel B of Table 4 reports the average subsequent VIX futures returns for tercile portfolios sorted by the expected volatility return. We use the baseline implementation with the one-month VIX futures contract and the expected volatility return from the one-factor model. Once again, we study not only realized returns over the following month but also longer horizon returns. Consistent with the returns in the predictive regressions, we divide the h-month log returns by h. The long-short returns are statistically significant at all horizons. They are economically large, especially at short horizons. The signs of the long-short returns are consistent with the sign of the loadings in the predictive regressions.

B. Predicting Variance Returns with Alternative Predictors

Table 6 presents the results of univariate and bivariate predictive regressions of the S&P 500 and volatility returns on the expected volatility return and other predictors. Once again these results are based on the baseline implementation with the one-month VIX futures contract and the expected volatility return from the one-factor model. We do not think of these results as a horse race. The expected volatility return in equation (2) is suggested by theory as a natural predictor for future market volatility returns and by extension for index returns. This connection to theory is less direct for several other predictors used in the literature. A high correlation between any such predictor and the expected volatility returns therefore does not reduce the relevance of the expected volatility return as a novel predictor; instead, it merely illustrates its connection to the existing empirical literature. Panel C of Table 1 reports the correlations between the expected volatility return and these predictors of S&P 500 returns that have been studied in the existing literature. The absolute value of this correlation is highest for the VIX term structure slope.

***********Table 6 about here*************

Tables A2 and A3 in the Online Appendix report univariate results for multiple forecasting horizons. Results are of course somewhat dependent on the forecasting horizon, because the predictive power and statistical significance of different predictors peak at different horizons. To give the alternative predictors their best chance, Table 6 reports on an intermediate (three-year) return horizon. The overall conclusion is that the predictive power of the expected volatility returns remains in the presence of other predictors. We now discuss these results in more detail.

The Variance Risk Premium We remarked earlier that expected volatility returns become more negative when volatility increases. This may be related to the findings of Bollerslev et al. (2009) that the market variance risk premium has predictive power for index returns. We therefore document the predictive power of the variance risk premium for subsequent S&P 500 index returns and VIX futures returns in our sample. We adopt the definition of the variance risk premium in Bollerslev et al. (2009), $VIX^2_{t,t+1} - RV_{t-1,t}$, where $RV_{t-1,t}$ denotes the realized variance between $t-1$ and t, and we express the variance risk premium in monthly percentage squared terms. The sample for realized variance and hence the variance risk premium ends in December 2020.

Table A3 confirms the robustness of the finding in Bollerslev et al. (2009) that the variance risk premium positively predicts index returns, and Table A2 confirms our prior that this predictive power also obtains for volatility returns because of the leverage effect.[21] Table 6 indicates that for the purpose of predicting index returns and volatility returns, the information in the expected volatility return differs from that in the variance risk premium.

Decomposing the Variance Risk Premium Bandi and Perron (2008) and Sizova (2013) report that lagged variance predicts index returns. Bandi et al. (2019) report a hump-shaped pattern in forward-backward regressions, with maximum R^2 at the 16-year horizon. The lagged variance is one of the two components of the implementation of the variance risk premium in Bollerslev et al. (2009). We therefore report on predictive regressions using both components of the variance risk premium.[22] Following the construction of the variance risk premium, the realized variance and the squared VIX are also expressed in monthly percentage squared terms. The sample for the VIX ends in November 2022. The sample for realized variance ends in December 2020.

Tables A2 and A3 show that the risk-neutral and physical variances have some (long-horizon) predictive power by themselves. In Table A2 (A3), the slope estimates are

[21] There are some qualitative differences between the results in Table A3 and the results in Bollerslev et al. (2009), where the R^2 of the predictive regression peaks at the three-month horizon. We verified that these differences are due to the sample period. When we use the 1990-2007 sample in Bollerslev et al. (2009), our results are very similar to theirs. Moreover, Table A4 in the Online Appendix reports that for the 1990-2020 sample, the maximum R^2 also obtains for a three-month horizon.

[22] Kilic and Shaliastovich (2019) instead decompose the variance risk premium into good and bad variance premia.

negative (positive) for both variance measures for all return horizons. For a given maturity, the loadings are also very similar for both variances, while the R^2s are slightly higher for the risk-neutral variance. Table 6 shows that the predictive ability of the expected volatility return remains in the presence of the realized variance or the squared VIX.

The Variance Risk Premium and the Volatility Return Next we rewrite the volatility return in equation (1) as:

$$R_{t,T} = \frac{VIX_T - F_t(T)}{F_t(T)}. \tag{8}$$

Comparing this return with the variance risk premium, we see that the expected volatility return differs from the variance risk premium in (at least) three ways: 1) It uses volatility rather than variance units; 2) The second term in the numerator is the futures price rather than the realized variance; and 3) The volatility return in equation (8) is scaled by the futures price. To further investigate the relation to the variance risk premium, Tables A2 and A3 therefore document the predictive power of the non-scaled predictor $VIX_T - F_t(T)$. While the slope of course differs from the baseline results in Table 4, the patterns in the t-statistics and R^2s are similar. Row 5, labeled 'Scaled VRP', reports on the variance risk premium divided by the futures price as a predictor. These results are more similar to those obtained for the VRP. We conclude that the expected volatility return contains information that is distinct from the variance risk premium.

Volatility of Volatility The higher moments of the return distribution capture risk that is different from the variance. We therefore also study if volatility of volatility can predict future returns. Following Huang et al. (2019), to measure volatility of volatility we use the VVIX index which is calculated from VIX options in an analogous way to the VIX index. We obtain VVIX data for August 2006 to November 2022 from the CBOE website. Tables A2 and A3 indicate that the VVIX has some predictive power, mainly for intermediate horizons. The pattern of the R^2s and t-statistics as a function of the forecast horizon for VVIX differs from the pattern for expected volatility returns in our sample, but also from the patterns for the VIX and the VRP. Table 6 once again indicates that including this variable does not affect the predictive power of the expected volatility return.

The Tail Factor Recent studies have highlighted the importance of tail risk.[23] We obtain data on the Left Tail Volatility index (LTV) from tailindex.com. LTV is a measure of return volatility generated by the left tail of the one-week risk-neutral return distribution, inferred from short-maturity out-of-the-money put options. Tables A2 and A3 indicate that LTV is informative about future returns, especially at longer horizons. In both tables, the sign at the one-month horizon differs from the sign for all other horizons, but it is not statistically significant. The bivariate regression in Table 6 shows that the predictive power of the expected volatility return remains.

The VIX Slope Johnson (2017) finds that the second principal component of the VIX term structure (the slope factor) contains predictive information about future returns on a range of volatility assets including variance swaps, VIX futures, and straddles at the daily and monthly horizons. For example, he finds that the slope predicts VIX futures returns at the monthly horizon with an R^2 of 11.44%. We confirm these results for our sample. We obtain the slope factor data from Travis Johnson's website. The sample for the slope factor ends in December 2017. Although our returns are measured differently from Johnson (2017), Tables A2 and A3 confirm that the slope factor predicts VIX futures returns and show that this also translates into index return predictability. The predictive power for futures returns peaks at the 3-year horizon with an R^2 of 23%. The bivariate regression in Table 6 shows that the t-statistics for both regressors are smaller compared to

[23] See for instance Bollerslev and Todorov (2011), Bollerslev et al. (2015), Andersen et al. (2015), Andersen et al. (2020), and Andersen et al. (2022).

the univariate regressions, which is not surprising given that the correlation between the two predictors is -0.63. However, the predictive power of the expected volatility return remains when including the VIX slope in the predictive regression.

Univariate Sorts Univariate sorts are also useful to illustrate the similarities and differences between the informational content in expected volatility returns and other forecasting variables. Panels A-C of Table A5 in the Online Appendix report on tercile portfolios sorted by the variance risk premium, the VVIX, and LTV. The signs of the long-short returns are mostly consistent with the predictive regressions, but for the VRP, the return at the one month horizon is almost zero and not significant. We verified that the sign is often negative in subsamples, while the sign for longer horizons is positive. This finding is related to Cheng (2018), who refers to it as the VIX premium puzzle. Similar to the expected volatility returns, the return difference for the variance risk premium is highly statistically significant at longer horizons.

Table A5 also confirms that the information in the VVIX differs from that in expected returns and the VRP because the economically largest and statistically most significant return spreads occur at intermediary horizons. For LTV, the sign at the one-month horizon also differs from that at longer horizons, and the long-short returns are largest at intermediate horizons.

Summary In summary, Table A3 confirms that predictors suggested by the existing literature have some predictive power for S&P 500 index returns, and Table A2 shows that this translates into predictive power for volatility returns. However, Table 6 suggests that the expected volatility return seems to contain different or additional information, and therefore retains predictive power for realized volatility returns and S&P 500 index returns when these other predictors are included in the predictive regression.

C. Out-of-Sample Recursive Predictive Regressions

Our baseline implementation does not suffer from data-snooping because expected returns are computed recursively, i.e., we construct the expected return predictor by re-estimating the VIX model parameters each month. However, for simplicity we present results based on a single forecasting regression. We now report on predictive regressions using the same expected returns, but re-estimating the predictive regression each period. Panel B of Tables 7 and 8 reports the average slope estimates, t-statistics and R^2s from these predictive regressions as a function of the forecast horizon. The literature often refers to these recursive regressions as out-of-sample predictive regressions, while our baseline results, which are repeated in Panel A of Tables 7 and 8 for convenience, are referred to as in-sample predictive regressions.

***********Table 7 about here*************

Tables 7 and 8 indicate that the impact on the results is limited. Figure 5 provides more details by plotting the resulting time series of the regression slopes, R^2s, and t-statistics. Panels A, C, and E report on the one-month horizon and Panels B, D, and F on the five-year horizon. By construction the time-series for the five-year returns are shorter. For a given forecast horizon, the slopes and R^2s over time are of a similar order of magnitude, and do not seem to be affected by outliers. However, the slopes and especially the R^2s decrease towards the end of the sample period, perhaps due to the impact of the Covid-19 crisis on volatility returns. Panels E and F present the time series of t-statistics in the recursive regressions. Note the strong pattern in the Hodrick t-statistics in Panel F, in contrast with the pattern in the HH t-statistics, suggesting potential biases in the Hodrick t-statistics when using long horizons in small samples. In Section VI. below, we present simulation evidence that is consistent with these patterns.

*************Table 8 about here***************

*************Figure 5 about here***************

D. VIX Dynamics and Predictive Regressions

In Section C., we compare the predictive performance of four different VIX dynamics for the future VIX: a simple square root specification, a two-factor generalization of this simple model, an ARMA(2,2) specification, and the heterogeneous autoregressive (HAR) model of Corsi (2009).[24] We find that no model clearly dominates, and that the models' predictive performance displays more similarities than differences. We have therefore reported results based on the simplest model, the square root specification for the VIX. We now present the results of predictive regressions for volatility returns on the one-month contract for the three other models.

We focus on an ARMA(2,2) dynamic because Cheng (2018) uses ARMA dynamics for the VIX and mainly reports results based on an ARMA(2,2) dynamic.[25] The main focus in Cheng (2018) is the puzzle that the ex ante estimate of the VIX premium embedded in VIX futures falls or stays flat when risk rises. This ex-ante risk premium is a measure of the expected return on volatility. Despite this VIX premium puzzle, Cheng (2018) finds that estimated risk premiums reliably forecast future one-month premiums, similar to our finding that expected returns on volatility forecast realized returns. Cheng (2018, 2020) analyzes realized one-month volatility returns, while we also study long-horizon returns.

[24]For alternative reduced form models for volatility modeling and forecasting, see for example Corsi (2009), Bekaert and Hoerova (2014), and Bekaert et al. (2013).

[25]We investigated other autoregressive specifications of the VIX and found that the predictive regressions are not greatly impacted. We use monthly hold-to-maturity returns, but results are also similar when using monthly holding period returns as in Cheng (2018). The timing of returns in Cheng (2018) is also different. He considers the monthly return of holding a futures contract expiring in month $t+1$ from the end of month $t-1$ to the end of month t while we look at monthly hold-to-maturity returns.

Panels C-E of Table 7 report on predictive regressions for forecasting VIX futures returns. As in the baseline analysis, all expected return measures are based on the recursive implementation in which we estimate volatility forecasting models recursively to generate the numerator. The results are qualitatively similar to the baseline results in Panel A, but the R^2s are generally lower, especially at long horizons. Panels C-E of Table 8 present the corresponding results for forecasting S&P 500 returns. Again they are qualitatively similar to the baseline results.

E. Simple Returns

Cochrane (2001) shows that when predicting index returns using the price-dividend ratio, results for log returns and simple returns are different at long horizons. Panel F of Table 7 reports results for predicting simple rather than log VIX futures returns, and Panel F of Table 8 reports on simple S&P 500 returns. In both cases, the results are similar to the baseline results in Panel A.[26]

F. Combining Expected Returns Across Maturities

Cochrane and Piazzesi (2005) show that a simple combination of forward rates predicts excess bond returns. Motivated by their results, we investigate whether combining expected volatility returns across maturities can improve forecasting performance. Similar to computing expected 1-month holding-to-maturity returns ($ER^V_{t,t+1}$), each month we can also compute the expected returns from holding a two-month ($ER^V_{t,t+2}$), three-month ($ER^V_{t,t+3}$), four-month ($ER^V_{t,t+4}$), and five-month VIX futures contract ($ER^V_{t,t+5}$) to maturity by using market prices of two-, three-, four-, and five-month contracts and the baseline VIX forecasting model to compute the corresponding physical expectations. We then extract the first principal component of these expected returns across different maturities

[26] It is not possible to compute Hodrick t-statistics when using simple returns.

and use it to predict the returns on (1-month) VIX futures and the S&P 500 index.[27]

Panel G of Table 7 (8) reports the results for forecasting VIX futures (S&P 500 index) returns. We find that combining expected returns across maturities in general leads to a notable increase in the R^2 of the predictive regressions, especially for VIX futures. For example, when using the first common component to predict returns, for the 24-month forecast horizon the R^2 is 30.14% and 11.59% for VIX futures and the S&P 500 index respectively, as compared to 17.07% and 5.13% for the baseline expected return reported in Panel A.[28]

G. Forward-Backward Regressions

This section further explores the term structure of volatility return predictability. We once again use our baseline predictive regression setup with the one-month futures contract and the simple square-root one-factor model to construct the expected volatility return. We implement the forward-backward regressions of Bandi et al. (2019), who regress long horizon returns (which are aggregated forward) onto long-run past expected volatility returns (which are aggregated backward), as follows:

$$r_{t+1,t+h} = \alpha_h + \beta_h ER^V_{t-h+1,t} + \epsilon \tag{9}$$

where $r_{t+1,t+h}$ is the long run VIX futures or S&P 500 index (log) return over different future horizons indexed by h, and $ER^V_{t-h+1,t}$ is the sum of the expected volatility returns over the past h months. Long-horizon forward-backward regressions capture information

[27]The expected volatility returns exhibit a strong factor structure, with the first principal component capturing 98.3% of the total variation and second component explaining 1.4%.

[28]Results are similar if we simply use the average expected return across different maturities. This is not entirely surprising as the first principal component is highly correlated with the level of the expected returns. The second principal component does not have predictive power.

that cannot be captured by the classical long-horizon regressions in Tables 4 and 5, where the loadings and R^2s mechanically increase with horizon in the presence of persistent regressors. This approach can also be used to identify components of the predictor that operate at different frequencies, thereby overcoming low signal-to-noise problems in the predictive regression.

***********Table 9 about here*************

Panel A of Table 9 reports the results for forecasting VIX futures returns and Panel B for S&P 500 index returns. Consistent with the findings reported in Bandi et al. (2019) for predictive regressions of market returns on variance, we observe hump-shaped patterns in predictability in both panels. Estimates are statistically significant at horizons less than or equal to 24 months.[29] Panel A of Figure 6 plots the R^2s of the backward-forward regressions as a function of the horizon. The R^2s peak at the 24-month horizon for both predictive regressions. These findings suggest the presence of different components in expected and realized volatility returns that operate at different frequencies. Note that these results are not inconsistent with the patterns we document in Tables 4 and Table 5, because those results do not rely on backward aggregation.

***********Figure 6 about here*************

H. Equilibrium Generated Predictability

Eraker (2025) proposes a test of equilibrium generated predictability (EGP). The intuition behind the test is that in an equilibrium setup, the autocorrelation in the

[29] We use the rescaled t-statistics of Valkanov (2003) ($t/\sqrt{T}$, case 2) to evaluate statistical significance. Valkanov (2003) shows that standard t-statistic does not converge to a well defined distribution, but the scaled t-statistic does. While the limiting distribution of $t/\sqrt{T}$ is non-normal, it can be simulated following Valkanov (2003) to obtain critical values.

predictor ought to be related in the pattern of the loadings of the predictive regression as a function of the forecast horizon. We follow Eraker (2025) and report on the following Q statistic:

$$Q = (\hat{b} - b^*)'\hat{\Omega}(\hat{b} - b^*) \tag{10}$$

where $\hat{b}$ is the vector of estimated slopes from OLS regressions of future one-month VIX futures returns on the expected volatility return, b^* is the vector of corresponding theory-implied slopes, and $\hat{\Omega}$ is the estimated variance-covariance matrix. Specifically, the $\hat{b}_h$ are estimated using the following regressions:

$$r_{t+h-1,t+h} = \alpha_h + b_h ER^V_t + \epsilon \quad h = 1, 2, ...60 \tag{11}$$

where $r_{t+h-1,t+h}$ is the one period return on VIX futures h months from now.[30] The theory-implied slopes are computed from the contemporaneous regression of VIX futures returns on the changes in the expected volatility return, under the assumption that the expected return follows an AR(1) process, that is:

$$r_t = \alpha_0 + b_0 \Delta ER^V_t + \epsilon \tag{12}$$

$$ER^V_t = \rho ER^V_{t-1} + u_t. \tag{13}$$

We find that b_0 is negative, consistent with the notion that a positive shock to expected returns is associated with a price decrease and hence a lower contemporaneous realized return. A negative b_0 implies that the implied slopes b^* are positive, since $b^*_h = b_0(\rho^h - \rho^{h-1})$.

[30] By construction the averages of these one-period slopes over the relevant horizons collapse to the slopes reported in Table 4.

Panel B of Figure 6 plots the realized one-period slopes together with theory-implied slopes. The one-period estimated slopes are mostly positive, suggesting that expected volatility returns are positively related to future VIX futures returns. However, they turn negative around the 4-year horizon, which explains the decrease in the R^2s at the 4-year horizon observed in Table 4. As mentioned above, the theory-implied slopes are always positive and decrease with the horizon.

Panel C of Table 9 reports the Q statistics for different horizons h as well as the associated p values, based on the statistic's χ^2 distribution with $h-1$ degrees of freedom. For example, when $h = 12$, we use the first 12 one-period slopes to compute the Q statistic. The results suggest that the EGP null hypothesis cannot be rejected across horizons. For instance, when we use the entire term structure (i.e. $h = 60$), the Q statistic is 51.868 with a p value of 0.734. Panel D reports the same test for forecasting S&P 500 index returns and finds that the EGP null cannot be rejected either.[31]

VI. Inference in Long-Horizon Volatility Return Regressions

It is well-known that predictive regressions can induce statistical biases, and the literature has recognized that these biases may be further compounded in long horizon predictive regressions with overlapping returns. In this section we investigate these biases in the context of our data and sample, and we study size and power based on the parameterization of the predictive regressions.

[31] However, consistent with the findings of Eraker (2025), we find that the EGP null is strongly rejected for the variance risk premium. The results for the variance risk premium and expected volatility differ because their time-series properties (i.e., persistence) are very different, resulting in different patterns in the predictive regressions.

A. Bias in OLS Slopes in Predictive Regressions

It is well-known that forecasting with a persistent predictor may induce biases, even when returns are not overlapping (Stambaugh (1986), 1999). Furthermore, there is an extensive literature on potential problems with long-horizon forecasting regressions with overlapping returns. We now discuss how these issues impact our empirical results. For clarity we repeat our baseline forecasting regression:

$$\frac{1}{h}\sum_{i=1}^{h} r_{t,t+i} = \alpha_{t+h} + \beta_{t+h} ER^{V}_{t,t+1} + \epsilon_{t+h}, h = 1, 3, 6, 12, ..., 60 \tag{14}$$

where $r_{t,t+i}$ is the log return from holding VIX futures or the S&P 500 index for different horizons indexed by i, and $ER^{V}_{t,t+1}$ is the expected log volatility return.

The seminal work by Stambaugh (1986, 1999) discusses biases in the slope estimate in the single-period predictive regression when the predictor is persistent and its innovations are correlated with returns, and provides a plug-in adjustment for the biases based on the Kendall (1954) approximation. The literature has long recognized that these biases may also be present and may be further compounded in long horizon predictive regressions with overlapping returns. Recently, Boudoukh et al. (2022) (henceforth BIR) show how to compute the equivalent of the Stambaugh bias in the slope estimates for the case of long horizon regressions. We apply the BIR correction to our predictive regressions with expected volatility returns. Note that the BIR correction reduces to the Stambaugh correction in the special case of the one-period (one-month) horizon.

Panel A of Table 10 compares the bias-adjusted slopes of the predictive regressions according to BIR with the unadjusted slopes based on the baseline measure of expected returns in Table 4. The OLS slope estimates in Panel A of Table 4 are slightly biased upwards when predicting VIX futures returns, because the innovations to expected returns are slightly negatively correlated with the innovations to realized volatility returns. For the index return predictability regressions, Panel B of Table 10 indicates that the OLS estimate in Panel D of Table 4 is slightly biased downward due to the positive correlation between the innovations to expected returns and index returns.

***********Table 10 about here*************

These biases in the OLS slope coefficients are very small in our setting, mainly because the expected volatility return is much less persistent than the dividend price-ratio, which is used in BIR to quantify biases.[32] Bollerslev et al. (2014) reach a similar conclusion regarding the bias when using the VRP as a predictor. For expected volatility returns, another factor is that the correlation between the innovations to the predictor and the forecasting regression is small in magnitude.[33]

B. Comparing Standard Errors

Another potential problem with long-horizon regressions is the computation of the standard errors. In our baseline analysis, we report Hansen and Hodrick (1980) and Hodrick (1992) standard errors to account for the overlapping nature of the data because we found that these approaches have the best statistical properties. Panel C of Table 10 compares these standard errors with OLS standard errors. We also report on Newey and West (1987) standard errors, with lag length equal to the number of overlapping observations.[34] These standard errors are often used in the literature on predictive

[32] The expected volatility return in Table 10 has an AR(1) coefficient of 0.85.

[33] Lewellen (2004) points out that the bias may be even smaller when taking all relevant information into account.

[34] This means, for example, that we include 35 lags for monthly predictive regressions with 3-year returns. Our results are also robust to using the rule of thumb lag selection suggested by Newey and West (1994), which is based on sample length.

regressions. The OLS and Newey-West t-statistics are clearly much higher than the Hansen-Hodrick and Hodrick t-statistics. In Panel D, we report on the results of the weighted least squares (WLS) approach in Johnson (2018) applied to the volatility forecasting regressions. We follow Johnson (2018) and use the squared return as a proxy for conditional variance. Panel D shows that the slope and t-statistics change substantially at short and medium horizons, but at 3, 4, and 5-year horizons, the results are very similar to the OLS results. This is intuitive because at short horizons the conditional return variance is likely to fluctuate substantially, whereas the conditional variance does not move that much at longer horizons, and the results for WLS and OLS are similar.

C. Size and Power in Predictive Regressions

A large literature documents that the finite sample distribution of predictive regression statistics can be very different from asymptotic distribution and the evidence for stock return predictability becomes much weaker or statistically insignificant when taking into account the small sample size of long horizon regressions and the use of overlapping returns (see, among others, Nelson and Kim (1993); Goetzmann and Jorion (1993); Kirby (1997); Boudoukh et al. (2008)). In this section, we use simulations to investigate if our findings can be attributed to small sample bias. We simulate 219 monthly returns on VIX futures under the assumption of no predictability and an AR(1) process for our baseline measure of expected volatility return x_t:

$$r_{t+1}^{VIX} = \bar{r}^{VIX} + v_{t+1} \tag{15}$$

$$x_{t+1} = c + \rho x_t + u_{t+1} \tag{16}$$

where $\bar{r}^{VIX}$ is the mean return of one-month VIX futures, v_{t+1} and u_{t+1} are shocks.[35] We then use the simulated data to estimate the regression of $\sum_{i=1}^{h} r_{t+i}^{VIX}$ on x_t over different horizons as in our actual empirical analysis, and record the slope coefficient, various t-statistics, and R^2. We repeat this 50000 times to construct a finite sample distribution of slope coefficients, t-statistics, and R^2s. This is a standard framework for testing against the null of no predictability in the predictability literature.[36]

Panel A of Figure 7 compares the empirical slopes (data) with the mean and 95th percentile of the finite sample distribution of the slope, based on the parameterization of the predictive regression with the baseline expected volatility return. The mean of the simulated slope estimates exceeds zero across all horizons, suggesting slope coefficients are biased upwards in our case.[37] However, the mean slope in the simulation is very close to the truth (zero), thereby confirming the very small BIR bias slope corrections in Table 10. These results also confirm that the BIR correction formula works well. Most importantly, comparing the empirical slopes with the 95th percentile shows that the empirical slope coefficients are statistically significant at the 5% level at most horizons.

*************Figure 7 about here***************

Panel B of Figure 7 compares the R^2 found in the data with the mean and 95th percentile of the simulated R^2 across 50000 simulated samples. Consistent with existing studies, we find that in predictive regressions with overlapping returns and a persistent

[35] The parameters are as follows: $\bar{r}^{VIX} = -0.0676$, $c = -0.0035$, $\rho = 0.8503$, $\sigma_v = 0.2141$, $\sigma_u = 0.0504$, and $corr(\sigma_v, \sigma_u) = -0.0538$. Note that these estimates are for log returns, while Table 2 reports on simple returns.

[36] We do not assume a normal distribution for v_{t+1} and u_{t+1}; instead we use the wild bootstrap approach of Neely et al. (2014) and draw from actual residuals. See Online Appendix B for the details of this procedure.

[37] This follows once again from the fact that the innovations to expected returns are negatively correlated with the innovations to realized volatility returns ($corr(v_{t+1}, u_{t+1}) = -0.0538$) (Stambaugh (1999)).

predictor, R^2s increase with the horizon even if there is no predictability (see also Bollerslev et al. (2014)). However, the high R^2s in our predictive regressions are significant at most horizons and cannot be entirely explained by small sample bias at longer horizons. For example, the mean and 95th percentile of R^2 under the null of no predictability are 4.49% and 17.93% respectively for the three-year horizon, whereas the corresponding R^2 in the data is 33.57%.

We now turn to the size properties of the t-statistics documented in Table 10. Panel C of Figure 7 plots the 95th percentile of various t-statistics across 50000 simulated samples, indicating the size properties of different standard errors in finite samples. Consistent with existing studies, we find that the t-statistics are biased upward in finite samples. At the 1-month horizon, the bias in the t-statistics does not appear to be all that severe: the 95% critical values of the sampling distribution of all t-statistics are close to 1.645 (1.664 for OLS, 1.671 for HH, 1.693 for Hodrick, and 1.693 for NW). It becomes progressively worse, however, as we move to longer horizons. For example, at the 4-year horizon, OLS (NW) t-statistics above 5.637 (3.176) are observed in 5% of all simulated samples, indicating that there is a serious size distortion and OLS and NW t-statistics tend to reject the null hypothesis far too often in finite samples. HH and Hodrick t-statistics perform better, although they also have incorrect size. At the 4-year horizon, the corresponding 95% critical values of HH and Hodrick t-statistics are 2.110 and 2.0670, respectively. Comparing the 95th percentile values to the corresponding t-statistics we found in the data (Table 10), we find empirical t-statistics are significant across the majority of horizons and therefore the statistical evidence against the null of no predictability cannot be entirely explained by small sample bias. Finally, the Hodrick t-statistics are closest to the asymptotic values for most horizons, except for very long horizons, where HH t-statistics are most reliable.

We also conduct an analysis of the power of different t-statistics to examine their performance when the null hypothesis of no predictability is false. Appendix C reports the results. The performance ranking of the various t-statistics is less clear for power than for size, and depends on the parameterization.

In summary, we confirm the presence of statistical biases in predictive regressions with long-horizon returns. However, for our application the Stambaugh bias (and the BIR extension to long horizons) are very small for all reasonable parameterizations. We confirm results in the existing literature that HH and Hodrick standard errors have better finite sample size properties for our empirical exercise. The relative ranking of various standard errors based on power depends on the parameterization of the forecasting regression.

VII. Conclusion

We use the information in VIX futures prices to study realized and expected returns on market volatility. Expected volatility returns on VIX futures are available in closed form. We compute realized monthly returns on volatility based on fully collateralized positions in VIX futures contracts and investigate if these returns can be forecast using the ex-ante expected returns.

Expected volatility returns positively predict subsequent realized volatility returns. Volatility returns are negative on average. Following increases in volatility, expected volatility returns and subsequent realized volatility returns become more negative. We confirm our results using predictive regressions with long-horizon returns. Our results also have implications for the predictability of index returns. Because of the negative correlation between index returns and volatility returns, expected volatility returns negatively predict subsequent index returns, but these results are statistically less significant.

Our predictability results are robust to a wide range of variations in the empirical setup and a number of important statistical concerns. The predictive power of expected returns remains when controlling for other predictors, such as the slope of the VIX term structure in Johnson (2017), a tail factor estimated from index options (Andersen et al. (2015)), the market variance (Bandi and Perron (2008); Bandi et al. (2019)), and the market variance risk premium (Bollerslev et al. (2009)).

References

Alizadeh, S.; M. W. Brandt; and F. X. Diebold. "Range-Based Estimation of Stochastic Volatility Models." *Journal of Finance*, 57 (2002), 1047–1091.

Andersen, T. G.; N. Fusari; and V. Todorov. "The Risk Premia Embedded in Index Options." *Journal of Financial Economics*, 117 (2015), 558–584.

Andersen, T. G.; N. Fusari; and V. Todorov. "The Pricing of Tail Risk and the Equity Premium: Evidence from International Option Markets." *Journal of Business & Economic Statistics*, 38 (2020), 662–678.

Andersen, T. G.; V. Todorov; and M. Ubukata. "Tail Risk and Return Predictability for the Japanese Equity Market." *Journal of Econometrics*, 222 (2022), 344–363.

Andries, M.; T. Eisenbach; J. Kahn; and M. Schmalz, "The Term Structure of the Price of Variance Risk." (2023), working Paper, Federal Reserve Bank of New York.

Bakshi, G., and N. Kapadia. "Delta-Hedged Gains and the Negative Market Volatility Risk Premium." *Review of Financial Studies*, 16 (2003a), 527–566.

Bakshi, G., and N. Kapadia. "Volatility Risk Premiums Embedded in Individual Equity Options: Some New Insights." *Journal of Derivatives*, 11 (2003b), 45–54.

Bandi, F. M., and B. Perron. "Long-Run Risk-Return Trade-Offs." *Journal of Econometrics*, 143 (2008), 349–374.

Bandi, F. M.; B. Perron; A. Tamoni; and C. Tebaldi. "The Scale of Predictability." *Journal of Econometrics*, 208 (2019), 120–140.

Bandi, F. M., and A. Tamoni. "Spectral Financial Econometrics." *Econometric Theory*, 38 (2022), 1175–1220.

Bekaert, G., and M. Hoerova. "The VIX, the Variance Premium and Stock Market Volatility." *Journal of Econometrics*, 183 (2014), 181–192.

Bekaert, G.; M. Hoerova; and M. L. Duca. "Risk, Uncertainty and Monetary Policy." *Journal of Monetary Economics*, 60 (2013), 771–788.

Bollerslev, T.; J. Marrone; L. Xu; and H. Zhou. "Stock Return Predictability and Variance Risk Premia: Statistical Inference and International Evidence." *Journal of Financial and Quantitative Analysis*, 49 (2014), 633–661.

Bollerslev, T.; G. Tauchen; and H. Zhou. "Expected Stock Returns and Variance Risk Premia." *Review of Financial Studies*, 22 (2009), 4463–4492.

Bollerslev, T., and V. Todorov. "Tails, Fears, and Risk Premia." *Journal of Finance*, 66 (2011), 2165–2211.

Bollerslev, T.; V. Todorov; and L. Xu. "Tail Risk Premia and Return Predictability." *Journal of Financial Economics*, 118 (2015), 113–134.

Bondarenko, O. "Variance Trading and Market Price of Variance Risk." *Journal of Econometrics*, 180 (2014), 81–97.

Boudoukh, J.; R. Israel; and M. P. Richardson. "Biases in Long-Horizon Predictive Regressions." *Journal of Financial Economics*, 145 (2022), 937–969.

Boudoukh, J.; M. Richardson; and R. F. Whitelaw. "The Myth of Long-Horizon Predictability." *Review of Financial Studies*, 21 (2008), 1577–1605.

Carr, P., and L. Wu. "Variance Risk Premiums." *Review of Financial Studies*, 22 (2009), 1311–1341.

Cheng, I.-H. "The VIX Premium." *Review of Financial Studies*, 32 (2018), 180–227.

Cheng, I.-H. "Volatility Markets Underreacted to the Early Stages of the COVID-19 Pandemic." *Review of Asset Pricing Studies*, 10 (2020), 635–668.

Christoffersen, P.; S. Heston; and K. Jacobs. "The Shape and Term Structure of the Index Option Smirk: Why Multifactor Stochastic Volatility Models Work So Well." *Management Science*, 55 (2009), 1914–1932.

Cochrane, J. H. *Asset Pricing*, Princeton University Press (2001).

Cochrane, J. H. "Presidential Address: Discount Rates." *Journal of Finance*, 66 (2011), 1047–1108.

Cochrane, J. H., and M. Piazzesi. "Bond Risk Premia." *American Economic Review*, 95 (2005), 138–160.

Corsi, F. "A Simple Approximate Long-Memory Model of Realized Volatility." *Journal of Financial Econometrics*, 7 (2009), 174–196.

Dew-Becker, I.; S. Giglio; A. Le; and M. Rodriguez. "The Price of Variance Risk." *Journal of Financial Economics*, 123 (2017), 225–250.

Driessen, J.; P. J. Maenhout; and G. Vilkov. "The Price of Correlation Risk: Evidence from Equity Options." *The Journal of Finance*, 64 (2009), 1377–1406.

Duarte, J.; C. S. Jones; and J. L. Wang. "Very Noisy Option Prices and Inference Regarding the Volatility Risk Premium." *Journal of Finance, Forthcoming*, (2022).

Engle, R. F., and J. G. Rangel. "The Spline-GARCH Model for Low-Frequency Volatility and its Global Macroeconomic Causes." *Review of Financial Studies*, 21 (2008), 1187–1222.

Eraker, B. "Predictability Puzzles." *Journal of Financial and Quantitative Analysis*, 60 (2025), 482–523.

Eraker, B., and Y. Wu. "Explaining the Negative Returns to VIX Futures and ETNs: An Equilibrium Approach." *Journal of Financial Economics*, 125 (2017), 72–98.

Fama, E. F., and K. R. French. "Dividend Yields and Expected Stock Returns." *Journal of Financial Economics*, 22 (1988), 3–25.

Goetzmann, W. N., and P. Jorion. "Testing the Predictive Power of Dividend Yields." *Journal of Finance*, 48 (1993), 663–679.

Gorton, G., and K. G. Rouwenhorst. "Facts and Fantasies About Commodity Futures." *Financial Analysts Journal*, 62 (2006), 47–68.

Hansen, L. P., and R. J. Hodrick. "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis." *Journal of Political Economy*, 88 (1980), 829–853.

Heston, S. L., and K. Todorov, "Exploring the Variance Risk Premium Across Assets." (2024), working Paper, University of Maryland.

Hodrick, R. J. "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement." *Review of Financial Studies*, 5 (1992), 357–386.

Hong, H., and M. Yogo. "What Does Futures Market Interest Tell Us About the Macroeconomy and Asset Prices?" *Journal of Financial Economics*, 105 (2012), 473–490.

Huang, D.; C. Schlag; I. Shaliastovich; and J. Thimme. "Volatility-of-Volatility Risk." *Journal of Financial and Quantitative Analysis*, 54 (2019), 2423–2452.

Johnson, T. L. "Risk Premia and the VIX Term Structure." *Journal of Financial and Quantitative Analysis*, 52 (2017), 2461–2490.

Johnson, T. L. "A Fresh Look at Return Predictability Using a More Efficient Estimator." *Review of Asset Pricing Studies*, 9 (2018), 1–46.

Kendall, M. G. "Note on Bias in the Estimation of Autocorrelation." *Biometrika*, 41 (1954), 403–404.

Kilic, M., and I. Shaliastovich. "Good and Bad Variance Premia and Expected Returns." *Management Science*, 65 (2019), 2522–2544.

Kirby, C. "Measuring the Predictable Variation in Stock and Bond Returns." *Review of Financial Studies*, 10 (1997), 579–630.

Kostakis, A.; T. Magdalinos; and M. P. Stamatogiannis. "Robust Econometric Inference for Stock Return Predictability." *Review of Financial Studies*, 28 (2015), 1506–1553.

Lee, G. G. J., and R. F. Engle. "A Permanent and Transitory Component Model of Stock Return Volatility." (1999). Working Paper, University of California, San Diego.

Lettau, M., and S. C. Ludvigson, "Measuring and Modeling Variation in the Risk-Return Trade-Off."Y. Ait-Sahalia and L. P. Hansen, eds., "Handbook of Financial Econometrics: Tools and Techniques," , vol. 1 of *Handbooks in Finance*San Diego: North-Holland (2010), 617–690.

Lewellen, J. "Predicting Returns with Financial Ratios." *Journal of Financial Economics*, 74 (2004), 209–235.

Martin, I. "What is the Expected Return on the Market?" *Quarterly Journal of Economics*, 132 (2017), 367–433.

Mincer, J. A., and V. Zarnowitz, "The Evaluation of Economic Forecasts." J. Mincer, ed., "Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance," Boston, MA: NBER (1969), 3–46.

Moreira, A., and T. Muir. "Volatility-Managed Portfolios." *Journal of Finance*, 72 (2017), 1611–1643.

Moreira, A., and T. Muir. "Should Long-Term Investors Time Volatility?" *The Journal of Financial Economics*, 131 (2019), 507–527.

Neely, C. J.; D. E. Rapach; J. Tu; and G. Zhou. "Forecasting the equity Risk Premium: The role of technical indicators." *Management Science*, 60 (2014), 1772–1791.

Nelson, C. R., and M. J. Kim. "Predictable Stock Returns: The Role of Small Sample Bias." *Journal of Finance*, 48 (1993), 641–661.

Newey, W., and K. West. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, 55 (1987), 703–708.

Newey, W. K., and K. D. West. "Automatic Lag Selection in Covariance Matrix Estimation." *Review of Economic Studies*, 61 (1994), 631–653.

Pearson, N. D., and T. Sun. "Exploiting the Conditional Density in Estimating the Term Structure: An Application to the Cox, Ingersoll, and Ross Model." *Journal of Finance*, 49 (1994), 1279–1304.

Pyun, S. "Variance Risk in Aggregate Stock Returns and Time-Varying Return Predictability." *Journal of Financial Economics*, 132 (2019), 150–174.

Singleton, K. J. "Investor Flows and the 2008 Boom/Bust in Oil Prices." *Management Science*, 60 (2013), 300–318.

Sizova, N. “Long-Horizon Return Regressions with Historical Volatility and Other Long-Memory Variables.” *Journal of Business & Economic Statistics*, 31 (2013), 546–559.

Stambaugh, R. F. “Bias in Regressions with Lagged Stochastic Regressors.” (1986). Working Paper, University of Chicago.

Stambaugh, R. F. “Predictive Regressions.” *Journal of Financial Economics*, 54 (1999), 375–421.

Valkanov, R. “Long-Horizon Regressions: Theoretical Results and Applications.” *Journal of Financial Economics*, 68 (2003), 201–232.

Whaley, R. E. “Understanding the VIX.” *Journal of Portfolio Management*, 35 (2009), 98–105.

Whaley, R. E. “Trading Volatility: At What Cost?” *Journal of Portfolio Management*, 40 (2013), 95–108.

Table 1: VIX and VIX Futures Prices: Summary Statistics

Panel A reports the mean, standard deviation (std), skewness (skew), and kurtosis (kurt) of the VIX index for the 1990-2022 and 2004-2022 periods. The latter sample corresponds to the period when VIX futures are available. Panel B reports summary statistics for VIX futures prices, based on the March 26, 2004 to November 23, 2022 sample period. We use linear interpolation to generate daily prices of VIX futures with constant maturities of 1 month, 2 months, 3 months, 4 months, 5 months, 6 months, 7 months, 8 months, and 9 months. Panel C reports correlations between our baseline measure of expected volatility returns and other predictors.

Panel A: VIX				
	Mean	Std.	Skew.	Kurt.
Jan 2, 1990 - Nov 23, 2022	19.66	8.01	2.12	8.06
Mar 26, 2004 - Nov 23, 2022	19.29	9.01	2.45	8.77

Panel B: VIX Futures Prices				
Maturity (in months)	Mean	Std.	Skew.	Kurt.
1	20.20	7.70	1.87	4.93
2	20.85	7.08	1.53	3.24
3	21.23	6.60	1.27	1.96
4	21.50	6.25	1.08	1.09
5	21.73	5.99	0.95	0.60
6	21.93	5.79	0.88	0.37
7	22.10	5.62	0.84	0.28
8	22.23	5.46	0.81	0.24
9	22.35	5.31	0.79	0.18

Panel C: Correlations Between ER^V and Other Predictors							
	ER^V	RV	VIX	VRP	VVIX	Slope	LTV
ER^V	1.00						
RV	-0.25	1.00					
VIX	-0.30	0.94	1.00				
VRP	-0.26	0.26	0.57	1.00			
VVIX	-0.21	0.45	0.42	0.27	1.00		
Slope	-0.63	-0.31	-0.31	-0.13	-0.06	1.00	
LTV	-0.26	0.76	0.87	0.59	0.43	-0.18	1.00

Table 2: VIX Futures Returns

Panel A reports summary statistics for the returns of holding VIX futures to maturity over various horizons. Over the sample period (2004-2022), a total of 219 VIX futures contracts have matured. For each contract, we establish a long position 1 month, 2 months, 3 months, 4 months, and 5 months prior to the maturity date, compute the corresponding hold-to-maturity return, and convert it to a monthly horizon. Panel B decomposes the 5-month hold-to-maturity return into five consecutive monthly holding period returns. For example, suppose a VIX futures contract expires in June (month t). We decompose the return of holding this contract from January to June (t-5, t) into five monthly returns: January to February (t-5, t-4), February to March (t-4, t-3), March to April (t-3, t-2), April to May (t-2, t-1), and May to June (t-1, t). Panel C reports summary statistics for the returns of the cross-section of VIX futures with maturities ranging from one month to five months. Each month t, we establish a long position in VIX futures that expire in month $t+1$, $t+2$, $\ldots$ $t+5$, and we compute the returns on holding those contracts from month t to month $t+1$. Means and standard deviations (Std) are monthly, while the Sharpe ratio (SR) is annualized. Skewness (Skew) and kurtosis (Kurt) are measured over the relevant horizons.

Panel A: Average HTM Returns					
Time to maturity	1m	2m	3m	4m	5m
Mean	-0.037	-0.031	-0.022	-0.018	-0.015
Std.	0.319	0.275	0.236	0.210	0.192
SR.	-0.405	-0.388	-0.330	-0.288	-0.270
Skew.	7.080	5.063	4.000	3.339	3.020
Kurt.	73.306	38.002	23.151	16.102	12.858
# of obs	219	214	202	201	198
Panel B: Decomposing 5m HTM Returns					
	(t-5, t-4)	(t-4, t-3)	(t-3, t-2)	(t-2, t-1)	(t-1, t)
Mean	-0.004	-0.003	-0.009	-0.025	-0.032
Std.	0.100	0.174	0.175	0.253	0.332
SR.	-0.156	-0.067	-0.173	-0.337	-0.335
Skew.	2.667	7.910	5.074	7.680	6.884
Kurt.	13.197	86.442	41.166	80.715	68.295
# of obs	198	198	198	198	198
Panel C: One-Month Holding Period Returns on VIX Futures					
Contract Maturity	1m	2m	3m	4m	5m
Mean	-0.037	-0.022	-0.008	-0.006	-0.004
Std	0.319	0.270	0.218	0.174	0.141
SR	-0.405	-0.279	-0.129	-0.112	-0.104
Skew	7.080	9.438	9.177	8.479	7.635
Kurt	73.306	113.238	108.313	96.737	82.567
# of obs	219	208	201	203	200

Table 3: Forecasting the VIX

This table compares the out-of-sample forecasting performance for various forecasting horizons for four VIX models: The One-Factor Model (OF), the Two-Factor Model (TF), the HAR model, and the ARMA(2, 2) (ARMA) model. We use 65% (Panel A), 75% (Panel B), and 85% (Panel C) of the data for estimating parameters and the remaining 35%, 25%, and 15% of the data for assessing out-of-sample performance. We consider various performance measures, including the root-mean-squared error (RMSE), the mean absolute error (MAE), the R^2 of the Mincer-Zarnowitz (Mincer and Zarnowitz (1969)) regression of realized values onto the VIX forecast in the out-of-sample period, and the average correlation (Corr) between the model forecast and the forecasts generated by the winning model based on each of the above three criteria. Panel D reports the average performance across the three sample splits. For each performance measure, the best (worst) performing model receives a rank of 1 (4). The overall rank is computed based on the average rank across all performance measures.

Forecasting 1-Month Ahead VIX					**Forecasting 2-Month Ahead VIX**					**Forecasting 3-Month Ahead VIX**				
Panel A: 65%					Panel A: 65%					Panel A: 65%				
	OF	TF	HAR	ARMA		OF	TF	HAR	ARMA		OF	TF	HAR	ARMA
RMSE	5.990	6.112	6.132	6.564	RMSE	6.833	7.128	6.920	7.310	RMSE	7.008	7.350	7.072	7.314
MAE	3.895	3.669	3.804	3.982	MAE	4.669	4.321	4.400	4.561	MAE	5.095	4.676	4.722	4.875
R^2	0.380	0.383	0.381	0.345	R^2	0.199	0.200	0.208	0.181	R^2	0.142	0.136	0.148	0.127
Corr	0.964	0.982	0.964	0.924	Corr	0.948	0.919	0.963	0.905	Corr	0.922	0.884	0.946	0.876
Panel B: 75%					Panel B: 75%					Panel B: 75%				
	OF	TF	HAR	ARMA		OF	TF	HAR	ARMA		OF	TF	HAR	ARMA
RMSE	6.481	6.584	6.696	7.3640	RMSE	7.234	7.450	7.405	8.012	RMSE	7.446	7.747	7.563	7.979
MAE	4.141	3.872	4.104	4.376	MAE	4.819	4.395	4.587	4.869	MAE	5.203	4.866	4.918	5.148
R^2	0.330	0.341	0.330	0.286	R^2	0.165	0.190	0.175	0.140	R^2	0.123	0.136	0.132	0.103
Corr	0.950	0.975	0.953	0.885	Corr	0.876	0.938	0.903	0.807	Corr	0.830	0.915	0.869	0.759
Panel C: 85%					Panel C: 85%					Panel C: 85%				
	OF	TF	HAR	ARMA		OF	TF	HAR	ARMA		OF	TF	HAR	ARMA
RMSE	7.659	8.001	7.988	9.007	RMSE	8.392	9.082	8.792	9.655	RMSE	8.578	9.449	9.048	9.580
MAE	4.610	4.731	4.829	5.355	MAE	4.963	5.356	5.065	5.623	MAE	5.317	5.935	5.517	5.886
R^2	0.233	0.211	0.233	0.192	R^2	0.083	0.072	0.088	0.068	R^2	0.050	0.038	0.050	0.040
Corr	1.000	0.886	0.994	0.921	Corr	0.994	0.763	0.988	0.904	Corr	1.000	0.662	0.968	0.920
Panel D: Average					Panel D: Average					Panel D: Average				
	OF	TF	HAR	ARMA		OF	TF	HAR	ARMA		OF	TF	HAR	ARMA
RMSE	6.710	6.899	6.939	7.645	RMSE	7.486	7.887	7.706	8.326	RMSE	7.677	8.182	7.894	8.291
MAE	4.216	4.091	4.246	4.571	MAE	4.817	4.690	4.684	5.017	MAE	5.205	5.159	5.052	5.303
R^2	0.314	0.312	0.315	0.274	R^2	0.149	0.154	0.157	0.129	R^2	0.105	0.103	0.110	0.090
Corr	0.971	0.948	0.970	0.910	Corr	0.939	0.874	0.951	0.872	Corr	0.917	0.820	0.928	0.852
Overall Rank	1.50	2.25	2.25	4.00	Rank	2.33	2.50	1.25	4.00	Rank	2.00	3.00	1.25	3.75

Table 4: Forecasting VIX Futures Returns and S&P 500 Returns

Panel A reports the results of predictive regressions of VIX futures returns on the one-month maturity contract on one-month expected volatility returns, for various forecast horizons. Panel B reports the average VIX futures returns for tercile portfolios sorted on the expected volatility returns. Panel C reports the correlations between VIX futures returns and S&P 500 index returns over different horizons. Panel D reports results for predictive regressions for S&P 500 index returns. Slopes are scaled by the forecast horizon. T-statistics are computed according to Hansen and Hodrick (1980) and Hodrick (1992). The sample period is from April 2004 to November 2022.

Panel A: Forecasting VIX Futures Returns

Horizon(months)	1	3	6	12	24	36	48	60
Slope	0.46	0.32	0.24	0.18	0.16	0.17	0.12	0.11
Hodrick t	(3.44)	(2.77)	(2.06)	(1.51)	(1.56)	(1.74)	(1.49)	(1.94)
HH t	(3.06)	(2.21)	(1.81)	(1.68)	(2.29)	(3.18)	(2.80)	(4.69)
Adj. R^2	3.65%	5.23%	5.64%	7.49%	17.07%	33.57%	28.56%	50.59%

Panel B: Sort on Expected Volatility Return

Horizon (months)	1	3	6	12	24	36	48	60
L	-0.105	-0.098	-0.087	-0.082	-0.081	-0.082	-0.077	-0.078
2	-0.103	-0.069	-0.070	-0.067	-0.066	-0.062	-0.058	-0.063
H	0.005	-0.034	-0.042	-0.047	-0.045	-0.046	-0.054	-0.054
H-L	0.110	0.064	0.045	0.036	0.036	0.036	0.023	0.024
t-stat	(3.04)	(3.28)	(2.95)	(3.30)	(6.44)	(8.70)	(6.39)	(11.58)

Panel C: Correlations between S&P 500 Returns and VIX Futures Returns

Horizon (months)	1	3	6	12	24	36	48	60
Corr	-0.80	-0.78	-0.76	-0.76	-0.84	-0.85	-0.85	-0.76

Panel D: Forecasting S&P 500 Returns

Horizon (months)	1	3	6	12	24	36	48	60
Slope	-0.063	-0.027	-0.011	-0.008	-0.025	-0.032	-0.023	-0.024
Hodrick t	(-2.01)	(-1.07)	(-0.48)	(-0.34)	(-1.31)	(-1.75)	(-1.49)	(-2.27)
HH t	(-1.89)	(-0.77)	(-0.31)	(-0.25)	(-1.09)	(-1.59)	(-1.35)	(-2.04)
Adj. R^2	1.15%	0.31%	-0.25%	-0.27%	5.13%	14.38%	11.16%	21.96%

Table 5: Forecasting Longer-Maturity VIX Futures Returns

This table reports the results of predictive regressions of VIX futures returns on contracts with different maturities on the one-month expected volatility returns, for various forecast horizons. Panel A repeats the results for 1-month contracts reported in Table 4, Panel B reports on 2-month contracts, Panel C on 3-month contracts, Panel D on 4-month contracts, and Panel E on 5-month contracts.

Panel A: 1-Month VIX futures								
Horizon(months)	1	3	6	12	24	36	48	60
Slope	0.46	0.32	0.24	0.18	0.16	0.17	0.12	0.11
Hodrick t	(3.44)	(2.77)	(2.06)	(1.51)	(1.56)	(1.74)	(1.49)	(1.94)
HH t	(3.06)	(2.21)	(1.81)	(1.68)	(2.29)	(3.18)	(2.80)	(4.69)
Adj. R^2	3.65%	5.23%	5.64%	7.49%	17.07%	33.57%	28.56%	50.59%
Panel B: 2-Month VIX futures								
Horizon(months)	1	3	6	12	24	36	48	60
Slope	0.35	0.25	0.18	0.16	0.20	0.23	0.16	0.16
Hodrick t	(3.21)	(2.98)	(2.28)	(1.91)	(2.61)	(3.20)	(2.82)	(3.65)
HH t	(2.99)	(2.06)	(1.50)	(1.39)	(2.25)	(3.00)	(2.92)	(4.13)
Adj. R^2	3.66%	5.04%	4.79%	6.95%	22.58%	39.90%	35.38%	51.33%
Panel C: 3-Month VIX futures								
Horizon(months)	1	3	6	12	24	36	48	60
Slope	0.24	0.15	0.10	0.10	0.14	0.18	0.14	0.14
Hodrick t	(2.49)	(2.09)	(1.45)	(1.36)	(2.17)	(3.10)	(3.02)	(3.79)
HH t	(2.35)	(1.47)	(0.94)	(0.93)	(1.53)	(2.30)	(2.51)	(3.18)
Adj. R^2	2.18%	2.47%	1.78%	3.42%	13.34%	29.63%	32.23%	45.45%
Panel D: 4-Month VIX futures								
Horizon(months)	1	3	6	12	24	36	48	60
Slope	0.15	0.10	0.06	0.06	0.12	0.16	0.13	0.12
Hodrick t	(1.68)	(1.47)	(1.01)	(1.02)	(2.04)	(3.15)	(3.14)	(3.81)
HH t	(1.66)	(1.02)	(0.64)	(0.66)	(1.35)	(2.11)	(2.41)	(2.94)
Adj. R^2	0.84%	0.96%	0.57%	1.51%	10.61%	26.39%	30.59%	41.82%
Panel E: 5-Month VIX futures								
Horizon(months)	1	3	6	12	24	36	48	60
Slope	0.14	0.08	0.05	0.06	0.10	0.15	0.13	0.12
Hodrick t	(1.80)	(1.41)	(0.93)	(0.96)	(1.96)	(3.14)	(3.29)	(3.88)
HH t	(1.73)	(0.97)	(0.58)	(0.61)	(1.23)	(2.05)	(2.47)	(2.93)
Adj. R^2	0.98%	0.84%	0.41%	1.30%	9.30%	26.03%	32.32%	42.61%

Table 6: Multivariate Forecasting Regressions

Panel A reports the results of predictive regressions of 3-year VIX futures returns on one-month expected volatility returns and other predictors. Panel B reports on predictive regressions for index returns. The other predictors are realized variance (RV), the VIX index, the variance risk premium (VRP), the VVIX index, the tail risk index (LTV), and the VIX term structure (Slope). The slope estimates for the predictors other than the expected volatility return are scaled by 1,000. T-statistics are computed according to Hansen and Hodrick (1980) and Hodrick (1992).

Panel A: Forecasting VIX Futures Returns													
ER^V	0.17							0.16	0.16	0.16	0.15	0.14	0.17
Hodrick t	(1.74)							(1.79)	(1.80)	(1.82)	(1.42)	(1.64)	(1.33)
HH t	(3.18)							(3.15)	(3.13)	(3.09)	(3.28)	(3.18)	(2.69)
VRP		-0.39						-0.19					
Hodrick t		(-1.38)						(-1.13)					
HH t		(-2.27)						(-1.96)					
RV			-0.17						-0.09				
Hodrick t			(-1.18)						(-0.83)				
HH t			(-1.68)						(-1.28)				
VIX				-0.18						-0.09			
Hodrick t				(-1.26)						(-0.93)			
HH t				(-1.93)						(-1.51)			
VVIX					-0.50						-0.51		
Hodrick t					(-1.61)						(-1.70)		
HH t					(-1.82)						(-2.32)		
LTV						-3.35						-2.50	
Hodrick t						(-1.95)						(-1.96)	
HH t						(-3.62)						(-3.93)	
Slope							-11.21						-3.01
Hodrick t							(-2.04)						(-0.41)
HH t							(-3.22)						(-0.83)
Adj. R^2	33.57%	7.31%	7.01%	10.98%	6.27%	24.67%	22.93%	35.04%	35.06%	36.08%	36.86%	46.34%	41.50%
Panel B: Forecasting S&P 500 Returns													
ER^V	-0.032							-0.03	-0.03	-0.03	-0.02	-0.03	-0.02
Hodrick t	(-1.75)							(-1.74)	(-1.73)	(-1.70)	(-1.12)	(-1.47)	(-1.07)
HH t	(-1.59)							(-1.53)	(-1.50)	(-1.46)	(-1.32)	(-1.35)	(-1.17)
VRP		0.09						0.06					
Hodrick t		(1.68)						(1.62)					
HH t		(1.66)						(1.69)					
RV			0.04						0.03				
Hodrick t			(1.49)						(1.24)				
HH t			(1.38)						(1.22)				
VIX				0.04						0.03			
Hodrick t				(1.58)						(1.39)			
HH t				(1.50)						(1.38)			
VVIX					0.17						0.17		
Hodrick t					(3.19)						(3.32)		
HH t					(2.34)						(2.53)		
LTV						0.87						0.73	
Hodrick t						(2.27)						(2.30)	
HH t						(3.00)						(3.55)	
Slope							2.90						1.66
Hodrick t							(2.06)						(0.92)
HH t							(2.70)						(1.32)
Adj. R^2	14.38%	4.81%	5.08%	7.80%	7.34%	23.84%	10.53%	15.80%	16.08%	17.15%	18.52%	27.91%	24.64%

Table 7: Robustness: Predicting VIX Futures Returns

This table reports robustness results for predictive regressions with VIX futures returns. Panel A repeats the baseline results from Table 4. Panel B reports average slopes, R^2s and t-statistics from predictive regressions estimated recursively (out-of-sample). Panels C to E compute expected returns based on an ARMA(2,2) model, a HAR model, and a two-factor model for the VIX. Panel E uses simple returns instead of log returns. Panel G uses the first principal component of expected volatility returns across different maturities.

Panel A: ER^V								
Horizon(months)	1	3	6	12	24	36	48	60
Slope	0.46	0.32	0.24	0.18	0.16	0.17	0.12	0.11
Hodrick t	(3.44)	(2.77)	(2.06)	(1.51)	(1.56)	(1.74)	(1.49)	(1.94)
HH t	(3.06)	(2.21)	(1.81)	(1.68)	(2.29)	(3.18)	(2.80)	(4.69)
Adj. R^2	3.65%	5.23%	5.64%	7.49%	17.07%	33.57%	28.56%	50.59%
Panel B: Out of Sample								
Horizon(months)	1	3	6	12	24	36	48	60
Slope	0.44	0.32	0.23	0.14	0.19	0.21	0.16	0.13
Hodrick t	(2.93)	(2.47)	(1.71)	(1.05)	(1.62)	(2.04)	(1.96)	(3.86)
HH t	(2.54)	(1.76)	(1.37)	(1.05)	(2.00)	(3.54)	(3.54)	(5.09)
Adj. R^2	3.55%	4.97%	4.82%	4.40%	18.57%	44.46%	40.68%	56.48%
Panel C: ARMA Model								
Horizon(months)	1	3	6	12	24	36	48	60
Slope	0.52	0.38	0.23	0.14	0.15	0.17	0.10	0.09
Hodrick t	(2.68)	(2.74)	(1.70)	(1.14)	(1.32)	(1.77)	(1.29)	(1.70)
HH t	(2.54)	(2.18)	(1.47)	(1.21)	(1.85)	(2.86)	(1.76)	(2.32)
Adj. R^2	2.42%	4.13%	2.68%	2.57%	7.96%	19.49%	10.94%	18.71%
Panel D: HAR Model								
Horizon(months)	1	3	6	12	24	36	48	60
Slope	0.51	0.35	0.19	0.12	0.12	0.15	0.09	0.08
Hodrick t	(2.51)	(2.41)	(1.35)	(0.93)	(1.12)	(1.67)	(1.23)	(1.60)
HH t	(2.31)	(1.93)	(1.19)	(0.96)	(1.46)	(2.52)	(1.59)	(2.05)
Adj. R^2	1.94%	2.94%	1.44%	1.25%	4.44%	13.86%	7.91%	13.59%
Panel E: Two-Factor Model								
Horizon(months)	1	3	6	12	24	36	48	60
Slope	0.73	0.44	0.31	0.20	0.19	0.22	0.13	0.14
Hodrick t	(3.09)	(2.71)	(2.09)	(1.41)	(1.42)	(1.97)	(1.49)	(2.22)
HH t	(2.61)	(1.85)	(1.55)	(1.30)	(1.75)	(2.69)	(1.87)	(3.09)
Adj. R^2	2.56%	2.73%	2.79%	2.72%	6.63%	15.59%	9.68%	22.88%
Panel F: Simple Returns								
Horizon(months)	1	3	6	12	24	36	48	60
Slope	0.57	0.30	0.21	0.16	0.16	0.16	0.11	0.11
HH t	(2.47)	(1.98)	(1.61)	(1.56)	(2.24)	(3.15)	(2.74)	(4.65)
Adj. R^2	2.27%	4.12%	4.46%	6.54%	16.67%	33.02%	27.56%	49.92%
Panel G: Combining Expected Returns								
Horizon(months)	1	3	6	12	24	36	48	60
Slope	0.12	0.10	0.09	0.07	0.06	0.04	0.03	0.03
Hodrick t	(3.09)	(2.72)	(2.47)	(1.88)	(1.64)	(1.38)	(1.25)	(1.85)
HH t	(2.75)	(2.42)	(2.37)	(2.25)	(2.78)	(3.23)	(3.36)	(4.93)
Adj. R^2	3.28%	7.59%	12.02%	16.40%	30.14%	38.32%	35.63%	60.76%

Table 8: Robustness: Predicting S&P 500 Returns

This table reports robustness results for predictive regressions with S&P 500 index returns. Panel A repeats the baseline results from Table 4. Panel B reports average slopes, R^2s and t-statistics from predictive regressions estimated recursively (out-of-sample). Panels C to E compute expected returns based on an ARMA(2,2) model, a HAR model, and a two-factor model for the VIX. Panel E uses simple returns instead of log returns. Panel G uses the first principal component of expected volatility returns across different maturities.

Panel A: ER^V								
Horizon(months)	1	3	6	12	24	36	48	60
Slope	-0.063	-0.027	-0.011	-0.008	-0.025	-0.032	-0.023	-0.024
Hodrick t	(-2.01)	(-1.07)	(-0.48)	(-0.34)	(-1.31)	(-1.75)	(-1.49)	(-2.27)
HH t	(-1.89)	(-0.77)	(-0.31)	(-0.25)	(-1.09)	(-1.59)	(-1.35)	(-2.04)
R2	1.15%	0.31%	-0.25%	-0.27%	5.13%	14.38%	11.16%	21.96%
Panel B: Out of Sample								
Horizon(months)	1	3	6	12	24	36	48	60
Slope	-0.053	-0.017	-0.001	0.010	-0.034	-0.048	-0.041	-0.034
Hodrick t	(-1.43)	(-0.65)	(-0.07)	(0.29)	(-1.50)	(-2.13)	(-2.03)	(-4.91)
HH t	(-1.35)	(-0.41)	(-0.04)	(0.20)	(-1.09)	(-2.22)	(-2.53)	(-2.89)
Adj. R^2	0.54%	-0.32%	-0.63%	-0.32%	7.29%	28.03%	30.91%	38.10%
Panel C: ARMA Model								
Horizon(months)	1	3	6	12	24	36	48	60
Slope	-0.070	-0.032	-0.005	-0.004	-0.030	-0.038	-0.025	-0.025
Hodrick t	(-1.33)	(-1.05)	(-0.20)	(-0.20)	(-1.38)	(-1.83)	(-1.44)	(-2.61)
HH t	(-1.56)	(-0.76)	(-0.12)	(-0.12)	(-1.23)	(-1.86)	(-1.36)	(-1.87)
Adj. R^2	0.64%	0.18%	-0.44%	-0.45%	4.09%	11.55%	7.39%	13.70%
Panel D: HAR Model								
Horizon(months)	1	3	6	12	24	36	48	60
Slope	-0.066	-0.026	0.001	0.001	-0.025	-0.033	-0.023	-0.023
Hodrick t	(-1.22)	(-0.81)	(0.05)	(0.04)	(-1.20)	(-1.75)	(-1.41)	(-2.49)
HH t	(-1.37)	(-0.60)	(0.03)	(0.03)	(-1.00)	(-1.67)	(-1.28)	(-1.68)
Adj. R^2	0.39%	-0.10%	-0.47%	-0.48%	2.19%	7.89%	5.48%	10.13%
Panel E: Two Factor Model								
Horizon(months)	1	3	6	12	24	36	48	60
Slope	-0.097	-0.031	-0.020	-0.013	-0.028	-0.040	-0.032	-0.031
Hodrick t	(-1.58)	(-0.82)	(-0.67)	(-0.46)	(-1.11)	(-1.85)	(-1.63)	(-2.63)
HH t	(-1.58)	(-0.55)	(-0.37)	(-0.29)	(-0.90)	(-1.52)	(-1.42)	(-1.82)
Adj. R^2	0.68%	-0.16%	-0.25%	-0.30%	1.69%	6.13%	5.80%	9.95%
Panel F: Simple Returns								
Horizon(months)	1	3	6	12	24	36	48	60
Slope	-0.063	-0.025	-0.009	-0.006	-0.026	-0.033	-0.023	-0.025
HH t	(-1.85)	(-0.69)	(-0.24)	(-0.20)	(-1.09)	(-1.61)	(-1.33)	(-2.04)
Adj. R^2	1.09%	0.17%	-0.34%	-0.35%	5.09%	14.45%	10.70%	21.70%
Panel G: Combining Expected Returns								
Horizon (months)	1	3	6	12	24	36	48	60
Slope	-0.017	-0.012	-0.009	-0.008	-0.010	-0.007	-0.005	-0.006
Hodrick t	(-1.96)	(-1.45)	(-1.16)	(-1.04)	(-1.54)	(-1.28)	(-1.11)	(-2.00)
HH t	(-1.68)	(-1.09)	(-0.76)	(-0.78)	(-1.45)	(-1.47)	(-1.39)	(-2.28)
Adj. R^2	0.93%	1.40%	1.18%	2.10%	11.59%	14.41%	11.52%	25.67%

Table 9: Exploring The Term Structure of Return Predictability

Panels A and B of this table report on the forward-backward regressions of Bandi et al. (2019). The $t/\sqrt{T}$ statistics of Valkanov (2003) are reported in brackets. Statistical significance of the $t/\sqrt{T}$ statistics at the 10%, 5% and 1% levels, computed using the simulated critical values in Valkanov (2003), is denoted by *, **, and *** respectively. Panels C and D report on the equilibrium generated predictability (EGP) test of Eraker (2025).

Panel A: Forecasting VIX Futures Returns								
Horizon (months)	1	3	6	12	24	36	48	60
Slope	0.464	0.269	0.180	0.187	0.277	0.117	-0.056	-0.055
Scaled t	(0.30***)	(0.28***)	(0.25*)	(0.36*)	(0.74**)	(0.4)	(-0.19)	(-0.19)
R^2	3.65%	3.16%	2.42%	5.77%	23.79%	9.43%	2.02%	2.68%
Panel B: Forecasting S&P 500 Index Returns								
Horizon (months)	1	3	6	12	24	36	48	60
Slope	-0.063	-0.016	0.005	-0.024	-0.074	-0.039	-0.002	0.008
Scaled t	(-0.19***)	(-0.07)	(0.03)	(-0.17)	(-0.69**)	(-0.59*)	(-0.04)	(0.14)
R^2	1.15%	-0.22%	-0.44%	0.94%	21.64%	18.79%	-0.68%	1.06%
Panel C: EGP test for VIX Futures Returns								
N	6	12	24	36	48	60		
Q_N	8.692	10.926	16.926	31.539	44.142	51.868		
p value	0.122	0.449	0.813	0.636	0.592	0.734		
Panel D: EGP test for S&P 500 Index Returns								
N	6	12	24	36	48	60		
Q_N	9.369	11.140	18.059	32.304	43.862	50.601		
p value	0.153	0.517	0.8	0.645	0.643	0.801		

Table 10: Statistical Biases in Predictive Regressions

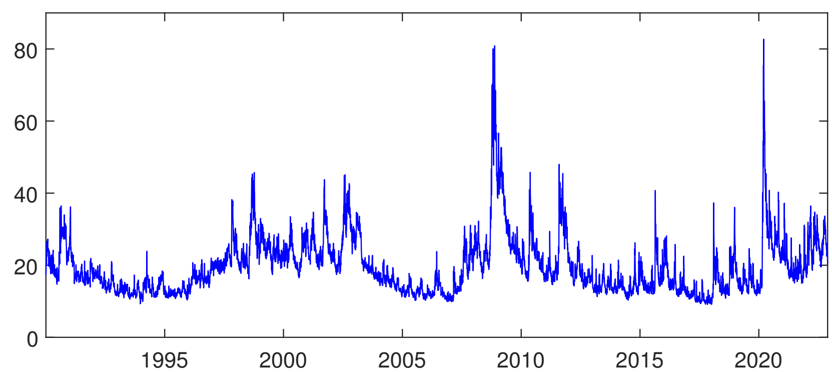
Panels A and B compare the slope coefficients in the VIX futures and S&P 500 index return predictive regressions with their bias-corrected counterparts. We use the formula in Boudoukh et al. (2022) to compute bias-adjusted slopes. Panel C compares various t-statistics assessing the statistical significance of predictive regressions with volatility returns. Panel D reports the results of the weighted least squares (WLS) approach in Johnson (2017).

Panel A: The Stambaugh Bias in the Volatility Return Forecasting Regression								
Horizon (months)	1	3	6	12	24	36	48	60
OLS slope	0.464	0.318	0.236	0.177	0.163	0.172	0.120	0.113
Bias-corrected slope	0.460	0.315	0.233	0.175	0.160	0.170	0.118	0.111
Panel B: The Stambaugh Bias in the S&P 500 Return Forecasting Regression								
Horizon (months)	1	3	6	12	24	36	48	60
OLS slope	-0.0631	-0.0266	-0.0111	-0.0077	-0.0254	-0.0324	-0.0232	-0.0244
Bias-corrected slope	-0.0626	-0.0262	-0.0107	-0.0074	-0.0251	-0.0321	-0.0230	-0.0241
Panel C: Alternative t-statistics								
Horizon (months)	1	3	6	12	24	36	48	60
OLS slope	0.464	0.318	0.236	0.177	0.163	0.172	0.120	0.113
OLS t	(3.04)	(3.60)	(3.71)	(4.21)	(6.41)	(9.67)	(8.33)	(12.80)
Hodrick t	(3.44)	(2.77)	(2.06)	(1.51)	(1.56)	(1.74)	(1.49)	(1.94)
HH t	(3.06)	(2.21)	(1.81)	(1.68)	(2.29)	(3.18)	(2.80)	(4.69)
NW t	(3.44)	(3.36)	(2.38)	(1.89)	(3.76)	(4.91)	(3.57)	(7.70)
Panel D: WLS								
Horizon (months)	1	3	6	12	24	36	48	60
Slope	0.006	0.017	0.015	0.094	0.005	0.186	0.104	0.114
OLS t-statistic	(0.29)	(1.32)	(1.90)	(4.93)	(0.19)	(9.07)	(6.92)	(12.36)

Figure 1: The Time Series of the VIX and VIX Futures Prices

Panel A plots the VIX. The sample period is from January 1990 to November 2022. Panel B plots daily prices of VIX futures with constant maturities of 1 month, 2 months, 3 months, 4 months, 5 months, 6 months, 7 months, 8 months, and 9 months. For each day in the sample, we use linear interpolation to generate these prices. The sample period is from March 26, 2004 to November 23, 2022.

Panel A: The VIX



Panel B: VIX Futures Prices

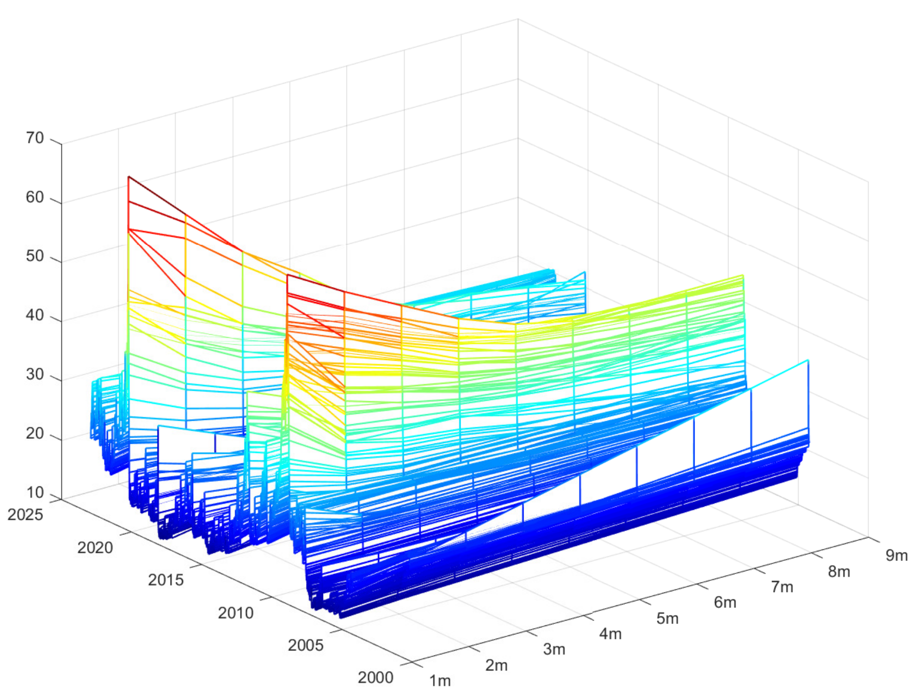
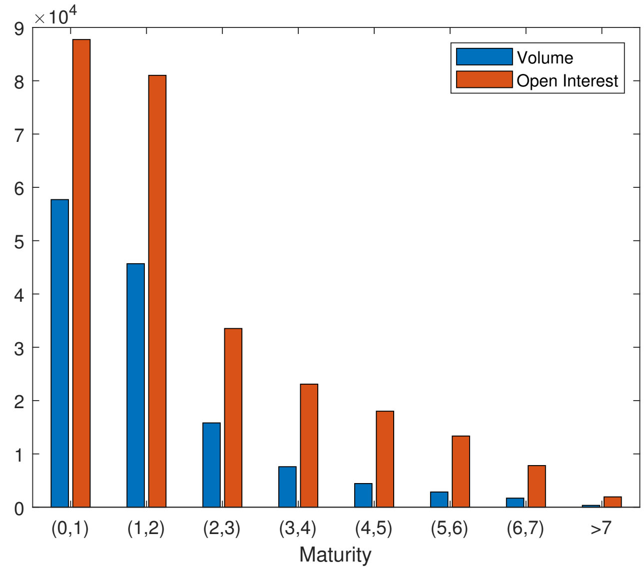


Figure 2: VIX Futures: Average Trading Volume and Open Interest

Panel A plots the average daily trading volume and open interest for VIX futures by time-to-maturity. Panel B plots the daily average trading volume and open interest per contract by year. The sample period is from March 26, 2004 to November 23, 2022.

Panel A: By Maturity



Panel B: By Year

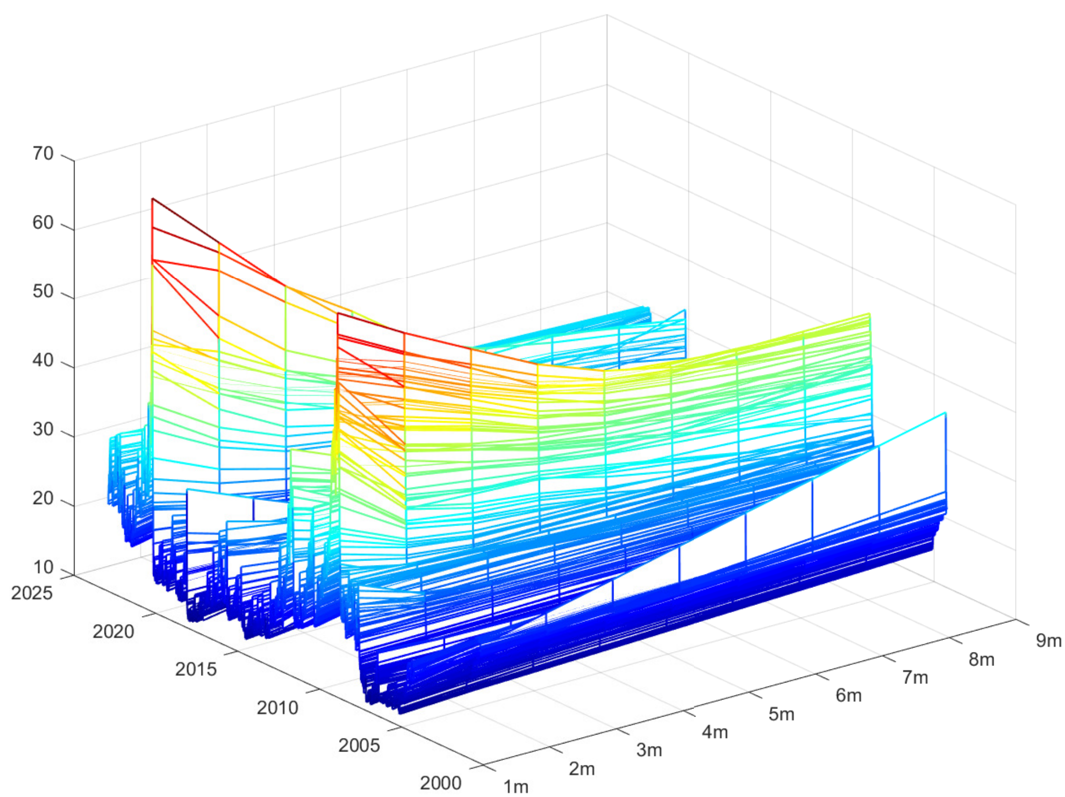
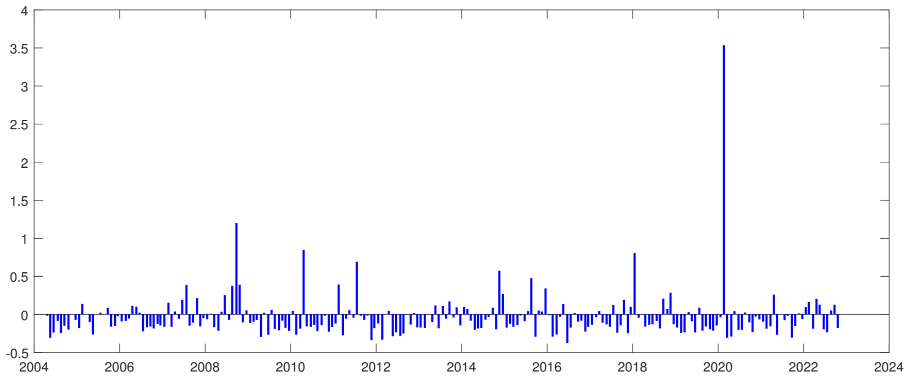


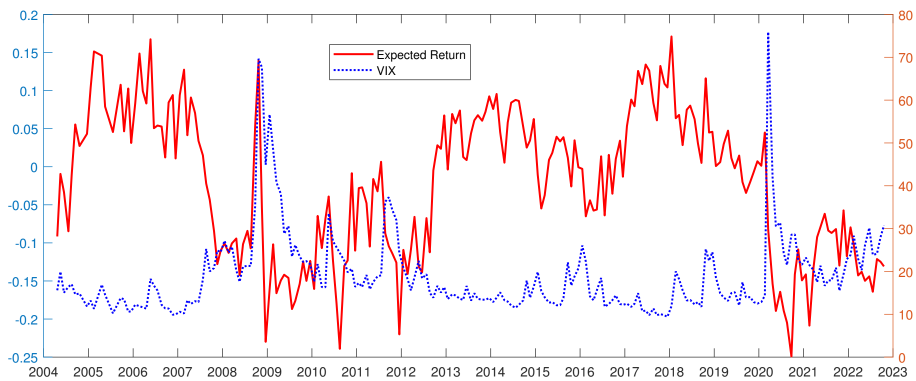
Figure 3: Monthly VIX Futures Returns and Expected Volatility Returns

Panel A plots monthly realized returns based on holding the 1-month VIX futures contract to maturity. Panel B plots expected 1-month hold-to-maturity VIX futures returns against the VIX. Panel C plots the expected returns from Panel B, together with subsequent 5-year VIX futures and stock returns.

Panel A: Realized VIX Futures Returns



Panel B: Expected VIX Futures Returns and the VIX



Panel C: Expected Returns and Subsequent 5-Year VIX Futures and S&P 500 Returns

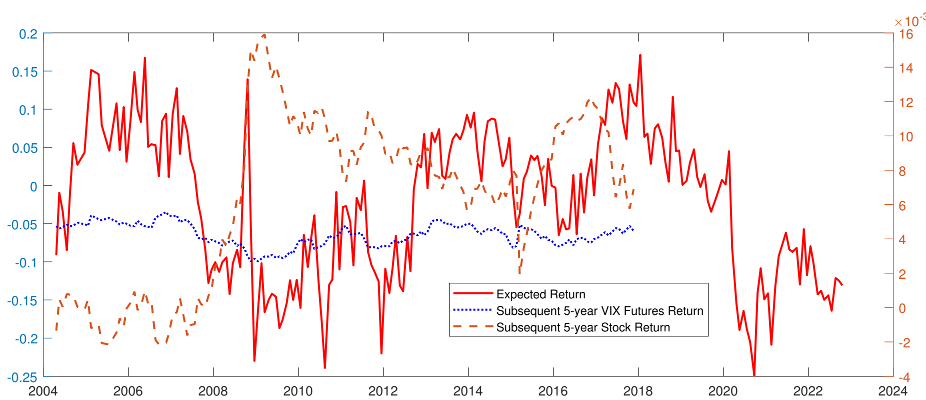


Figure 4: Expected Returns, Realized Returns and the VIX

This figure plots the VIX and subsequent 5-year VIX futures returns against expected volatility returns. Panel A scatter-plots the VIX against expected volatility returns. Panels B and C scatter plot expected returns against realized returns. Panel C highlights the role of the financial crisis.

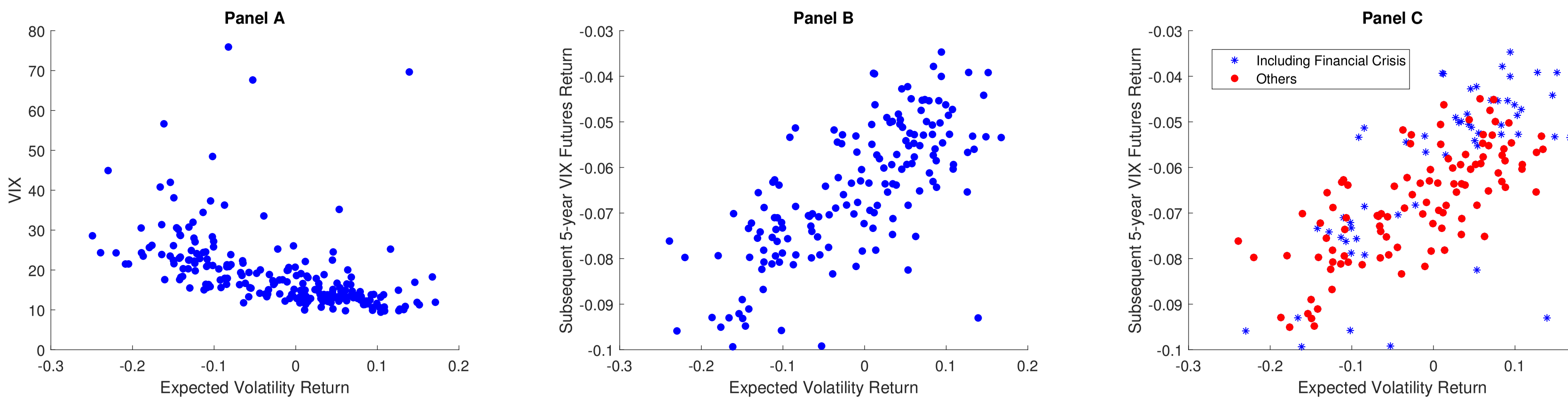


Figure 5: Time Series of Slopes, R^2s and t-statistics in Out-of-Sample Regressions

The left panels report on the one-month horizon and the right panels on the 60-month horizon. Panels A-B report on the slope, Panels C-D on the R^2, and Panels E-F on the t-statistics for out-of-sample (recursive) predictive regressions with one-month VIX futures returns.

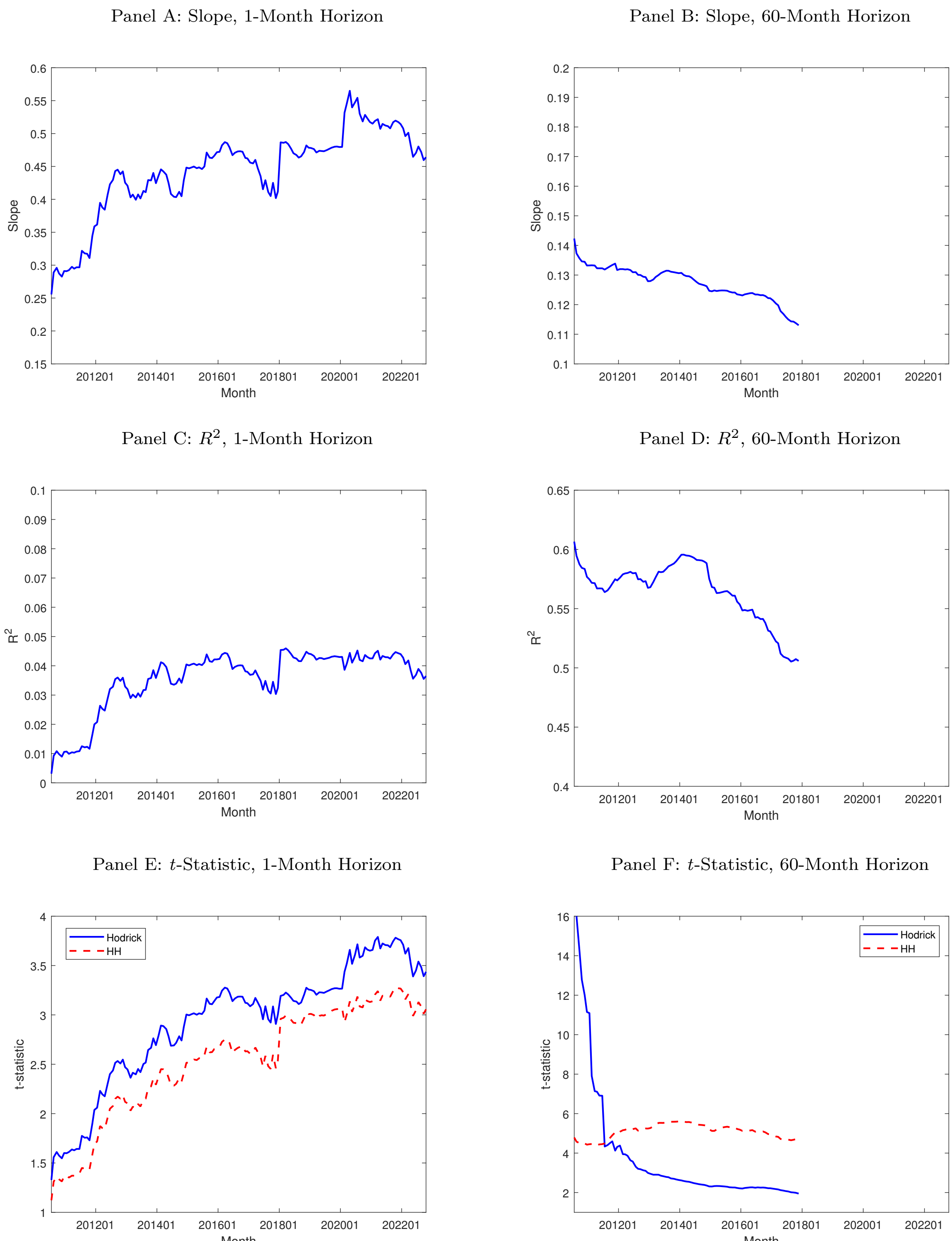
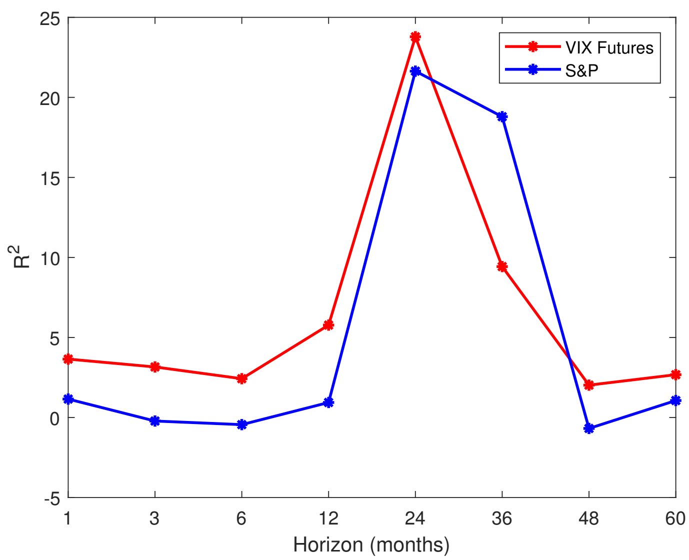


Figure 6: Exploring The Term Structure of Predictability

Panel A plots the R^2s of the forward-backward regressions in Bandi et al. (2019) as a function of the aggregation horizon. Panel B plots the equilibrium and realized slopes used in the EGP test of Eraker (2025) as a function of the horizon. Results are based on the baseline one-factor model and the VIX futures contract with one-month maturity.

Panel A: R^2s for Forward-Backward Regressions



Panel B: EGP Test: Equilibrium and Realized Slopes

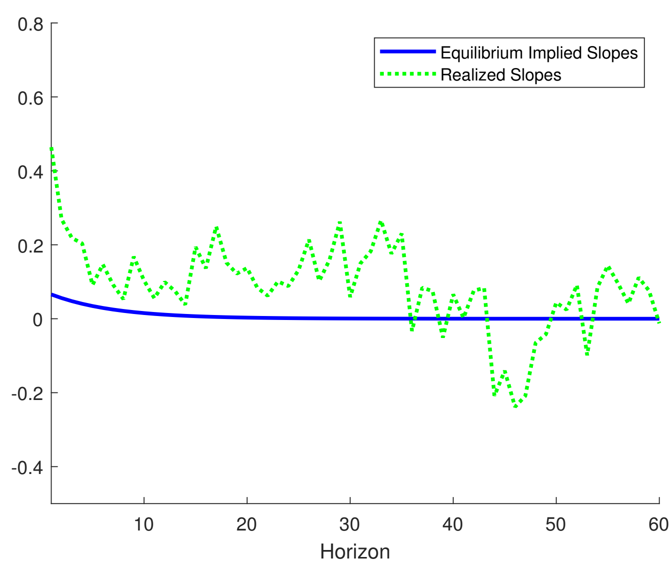
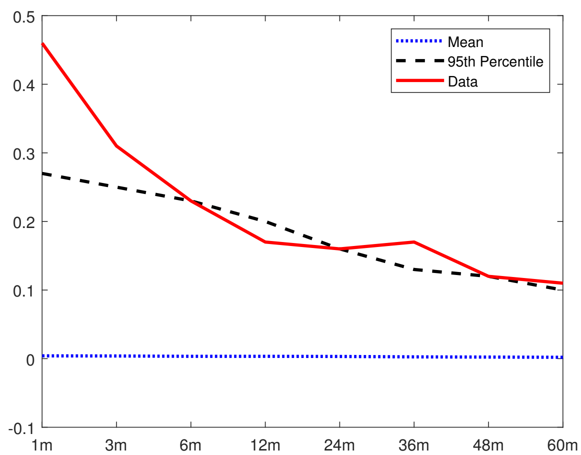


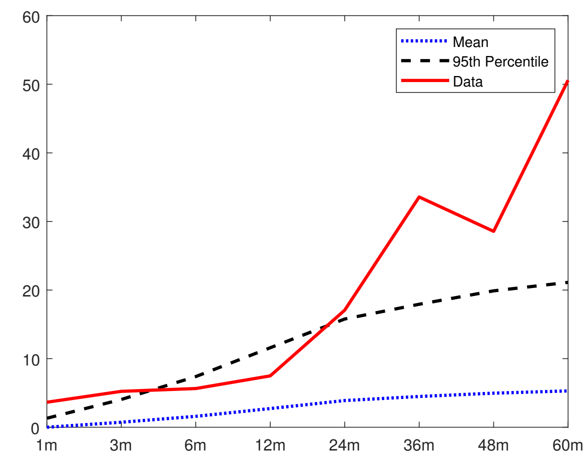
Figure 7: Assessing Finite Sample Bias in Predictive Regressions

Panel A plots the slope coefficients in predictive regressions for volatility returns. We compare the slope coefficients from the data with the mean and 95th percentile of their finite sample distribution under the null of no predictability. Panel B plots the R^2 from the data with the mean and 95th percentile of the finite sample distribution under the null of no predictability. Panel C plots the 95th percentile of the finite sample distribution of various t-statistics under the null of no predictability. The null hypothesis is parameterized using the baseline expected volatility returns.

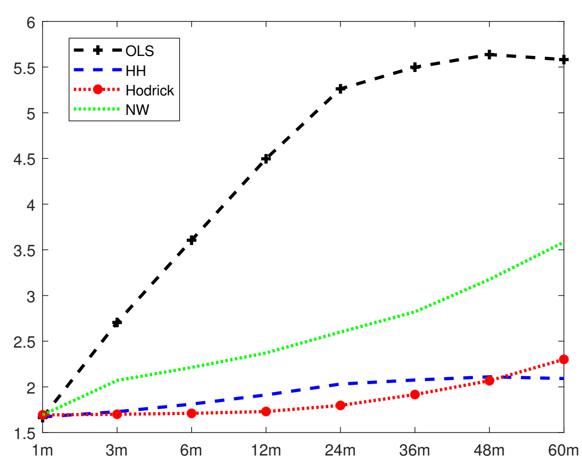
Panel A: Slope



Panel B: R^2



Panel C: t-statistic



Expected and Realized Returns on Volatility

Online Appendix

Guanglian Hu
University of Sydney

Kris Jacobs
University of Houston

A Estimation of the Square-Root Process

Since the transition density of the square-root process is known in closed-form (Pearson and Sun, 1994), we can estimate the physical parameters of the one-factor model via maximum likelihood. The transition density of the VIX is given by:

$$P(V_{t+\Delta}|V_t) = ce^{-u-v}(\frac{v}{u})^{\frac{q}{2}} I_q(2\sqrt{uv})$$

where $c = \frac{2\kappa}{\sigma^2(1-e^{-\kappa\Delta})}$, $u = cV_t e^{-\kappa\Delta}$, $v = cV_{t+\Delta}$, $q = \frac{2\kappa\theta}{\sigma^2} - 1$, and $I_q(\cdot)$ is the modified Bessel function of the first kind of order q. We estimate the physical parameters by maximizing the log-likelihood function:

$$(\kappa, \theta, \sigma) = \arg\max \sum_{t=1}^{T} \ln(P(V_{t+\Delta}|V_t)).$$

To provide some insight into the time-series properties of the VIX, consider estimates for the simple square root model in equation (4) of the paper that are estimated using the entire sample. The long run mean θ is 19.678, which is very close to the VIX sample average. The mean reversion parameter κ is 5.582, which implies that volatility shocks have a half life of $ln(2)/5.582 = 0.1242$ years or approximately 31 days.

B The Two-Factor Model

An extensive literature argues that at least two factors are needed to describe volatility dynamics (e.g., Lee and Engle, 1999; Alizadeh, Brandt, and Diebold, 2002; Christoffersen, Heston, and Jacobs, 2009; Engle and Rangel, 2008). Typically, one factor is slow moving and captures the low-frequency movements in volatility. The other factor has faster mean reversion and captures high-frequency movements. We therefore extend the one-factor model and allow the VIX to mean-revert to a time varying stochastic mean which itself follows a square-root process. This model is known as the stochastic mean/central tendency model and has been widely used in the volatility and term structure literature.[1] It is given by:

(B.1) $$dVIX_t = \kappa(\theta_t - VIX_t)dt + \sigma\sqrt{VIX_t}dW_{VIX}$$

(B.2) $$d\theta_t = \overline{\kappa}(\overline{\theta} - \theta_t)dt + \overline{\sigma}\sqrt{\theta_t}dW_\theta$$

where θ_t and $\overline{\theta}$ are the stochastic mean and the long-run mean of the VIX, and κ and $\overline{\kappa}$ control the speed of mean-reversion to θ_t and $\overline{\theta}$, respectively. The two-factor model states that the VIX is mean-reverting at a rate of speed κ to a time varying stochastic mean θ_t, which itself is mean-reverting to the long-run mean $\overline{\theta}$ at a rate of speed $\overline{\kappa}$.

[1]See, among others, Jegadeesh and Pennacchi (1996), Balduzzi, Das, and Foresi (1998), Egloff, Leippold, and Wu (2010), Bates (2012), Mencia and Sentana (2013), and Ait-Sahalia, Karaman, and Mancini (2020). Alternatively volatility can be modeled using two additive components. See, for example, Bates (2000) and Christoffersen, Heston, and Jacobs (2009). Another strand of literature emphasizes the importance of incorporating jumps into the volatility dynamics. See, among others, Broadie, Chernov, and Johannes (2007) and Eraker, Johannes, and Polson (2003).

C The Wild Bootstrap

Following Neely, Rapach, Tu, and Zhou (2014), we apply the wild bootstrap to assess the statistical significance of return predictability.[2] Let

$$\hat{\epsilon}_{t+1} = r_{t+1} - \hat{\alpha}_{t+1} - \hat{\beta}_{t+1} x_t,$$

where $\hat{\alpha}_{t+1}$ and $\hat{\beta}_{t+1}$ are the 1-month OLS parameter estimates based on the original sample. We further assume that x_{t+1} follows an AR(1) process:

$$x_{t+1} = \rho_0 + \rho_1 x_t + v_{t+1},$$

and let

$$\hat{v}_{t+1} = x_{t+1} - \hat{\rho}_0^c - \hat{\rho}_1^c x_t,$$

where $\hat{\rho}_0^c$ and $\hat{\rho}_1^c$ are reduced-bias estimates of the AR parameters as in Neely, Rapach, Tu, and Zhou (2014).

Based on these AR parameters ($\hat{\rho}_0^c$ and $\hat{\rho}_1^c$) and residuals ($\hat{\epsilon}_{t+1}$ and $\hat{v}_{t+1}$), we generate a pseudo sample of (r^*_{t+1}, x^*_t) under the null hypothesis of no return predictability:

$$r^*_{t+1} = \bar{r} + \hat{\epsilon}_{t+1} w_t,$$
$$x^*_{t+1} = \hat{\rho}_0^c + \hat{\rho}_1^c x^*_t + \hat{v}_{t+1} w_t$$

where $\bar{r}$ is the sample mean of 1-month realized returns, w_t is a draw from the standard normal distribution, and $x^*_1 = x_1$. As in Neely, Rapach, Tu, and Zhou (2014), we multiply the residuals by w_t, which amounts to the wild bootstrap. Note that the wild bootstrap

[2] We also consider the pairs bootstrap and find similar results.

not only preserves the contemporaneous correlations in the data but also accounts for general forms of conditional heteroskedasticity.

Using the pseudo sample of $(r^*_{t+1},\, x^*_t)$, we obtain long horizon returns and re-estimate the predictive regression and store the OLS t-statistic. We repeat this process 50000 times to get a finite sample distribution of the t-statistic under the null of no predictability.

D Power in Predictive Regressions

We conduct an analysis of the power of different t-statistics to examine their performance when the null hypothesis of no predictability is false, using the parameterization from the baseline expected volatility return. We simulate returns under the alternative hypothesis assuming various degrees of predictability (i.e., β takes different values):

$$r^{VIX}_{t+1} = \beta x_t + v_{t+1} \tag{D.1}$$

$$x_{t+1} = c + \rho x_t + u_{t+1}. \tag{D.2}$$

For each β, we again run 50000 simulations and calculate the rejection rates, which is the percentage of the simulated samples in which the t-statistics exceed the 95% critical values associated with the sampling distributions under the null hypothesis as shown in Panel C of Figure 7. Figure A1 in the Online Appendix compares the power performance of the OLS, HH, Hodrick and NW standard errors at the one-year and three-year horizons by plotting the rejection rates against β. With this parameterizion, the Hodrick standard error has the best performance. However, in unreported results, we find the rankings are sensitive to the parameter values used in simulation. We conclude that the performance ranking of the various t-statistics is less clear for power than for size, and depends on the parameterization.

References

Ait-Sahalia, Y., M. Karaman, and L. Mancini, "The Term Structure of Variance Swaps, Risk Premia and the Expectations Hypothesis", *Journal of Econometrics* 30 (2020), 2740–2768.

Alizadeh, S., M.W. Brandt, and F.X. Diebold, "Range-based estimation of stochastic volatility models", *Journal of Finance* 57 (2002), 1047–1091.

Balduzzi, P., S.R. Das, and S. Foresi, "The central tendency: A second factor in bond yields", *Review of Economics and Statistics* 80 (1998), 62–72.

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

Bates, David.S., "U.S. stock market crash risk, 1926-2010", *Journal of Financial Economics* 105 (2012), 229–259.

Broadie, M., M. Chernov, and M. Johannes, "Model specification and risk premia: Evidence from futures options", *Journal of Finance* 62 (2007), 1453–1490.

Christoffersen, P., S. Heston, and K. Jacobs, "The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well", *Management Science* 55 (2009), 1914–1932.

Egloff, D., M. Leippold, and L. Wu, "The term structure of variance swap rates and optimal variance swap investments", *Journal of Financial and Quantitative Analysis* 45 (2010), 1279–1310.

Engle, R.F., and J.G. Rangel, "The spline-GARCH model for low-frequency volatility and its global macroeconomic causes", *Review of Financial Studies* 21 (2008), 1187–1222.

Eraker, Bjørn, Michael Johannes, and Nicholas Polson, "The impact of jumps in volatility and returns", *Journal of Finance* 58 (2003), 1269–1300.

Gorton, G., and K.G. Rouwenhorst, "Facts and fantasies about commodity futures", *Financial Analysts Journal* 62 (2006), 47–68.

Hansen, L.P., and Robert.J. Hodrick, "Forward exchange rates as optimal predictors of future spot rates: An econometric analysis", *Journal of Political Economy* 88 (1980), 829–853.

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

Hong, H., and M. Yogo, "What does futures market interest tell us about the macroeconomy and asset prices?", *Journal of Financial Economics* 105 (2012), 473–490.

Jegadeesh, N., and G.G. Pennacchi, "The behavior of interest rates implied by the term structure of eurodollar futures", *Journal of Money, Credit and Banking* 28 (1996), 426–446.

Lee, G.J., and R.F. Engle, "A permanent and transitory component model of stock return volatility", Working Paper, University of California, San Diego (1999).

Mencia, J., and E. Sentana, “Valuation of VIX derivatives”, *Journal of Financial Economics* 108 (2013), 367–391.

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

Pearson, N.D., and T. Sun, “Exploiting the conditional density in estimating the term structure: An application to the Cox, Ingersoll, and Ross model”, *Journal of Finance* 49 (1994), 1279–1304.

Table A1: Robustness: Fully Collateralized Returns

This table reports results for forecasting VIX futures returns on a fully collateralized position following the implementation in Gorton and Rouwenhorst (2006) and Hong and Yogo (2012), respectively.

Panel A: Gorton-Rouwenhorst								
Horizon (months)	1	3	6	12	24	36	48	60
Slope	0.47	0.32	0.24	0.18	0.17	0.18	0.12	0.12
Hodrick t	(3.48)	(2.83)	(2.11)	(1.57)	(1.62)	(1.80)	(1.55)	(2.01)
HH t	(3.09)	(2.25)	(1.85)	(1.75)	(2.36)	(3.22)	(2.83)	(4.74)
Adj. R^2	3.74%	5.43%	5.95%	8.07%	18.17%	34.52%	29.20%	51.33%
Panel B: Hong-Yogo								
Horizon (months)	1	3	6	12	24	36	48	60
Slope	0.47	0.32	0.24	0.18	0.17	0.18	0.12	0.12
Hodrick t	(3.48)	(2.83)	(2.12)	(1.58)	(1.63)	(1.80)	(1.55)	(2.01)
HH t	(3.10)	(2.26)	(1.86)	(1.75)	(2.37)	(3.23)	(2.83)	(4.74)
Adj. R^2	3.75%	5.45%	5.99%	8.13%	18.29%	34.63%	29.28%	51.41%

Table A2: Forecasting VIX Futures Returns with Other Predictors

This table reports slopes (scaled by 1,000 for RV, VIX, VRP, VVIX, LTV, and Slope) from predictive regressions of VIX futures returns on alternative predictors suggested by the existing literature. T-statistics are computed according to Hansen and Hodrick (1980) and Hodrick (1992). The predictors are realized variance (RV), the VIX index, the variance risk premium (VRP), the VVIX index, the tail risk index (LTV), and the VIX term structure (Slope). The sample periods are as follows: VIX, April 2004 to November 2022; RV and VRP, April 2004 to December 2020; VVIX: August 2006 to November 2022; Slope, April 2004 to December 2017; LTV, April 2004 to December 2019.

Horizon (months)		1	3	6	12	24	36	48	60
	Slope	-0.41	-0.67	-0.88	-0.65	-0.47	-0.39	-0.41	-0.30
VRP	Hodrick t	(-0.73)	(-2.26)	(-2.79)	(-2.35)	(-1.68)	(-1.38)	(-1.77)	(-1.91)
	HH t	(-0.53)	(-1.41)	(-2.31)	(-2.26)	(-2.47)	(-2.27)	(-2.92)	(-3.05)
	Adj. R^2	-0.37%	0.57%	3.23%	4.24%	6.39%	7.31%	14.29%	15.86%
	Slope	-0.13	-0.19	-0.29	-0.33	-0.24	-0.17	-0.24	-0.18
RV	Hodrick t	(-0.32)	(-0.86)	(-1.78)	(-2.39)	(-2.00)	(-1.18)	(-1.93)	(-2.02)
	HH t	(-0.40)	(-0.69)	(-1.28)	(-2.10)	(-2.37)	(-1.68)	(-3.22)	(-3.55)
	Adj. R^2	-0.43%	0.02%	1.83%	6.83%	10.04%	7.01%	24.70%	28.64%
	Slope	-0.14	-0.21	-0.31	-0.32	-0.23	-0.18	-0.23	-0.17
VIX	Hodrick t	(-0.44)	(-1.19)	(-2.16)	(-2.53)	(-1.97)	(-1.26)	(-1.90)	(-2.01)
	HH t	(-0.53)	(-0.94)	(-1.61)	(-2.32)	(-2.52)	(-1.93)	(-3.51)	(-3.97)
	Adj. R^2	-0.33%	0.43%	3.23%	8.99%	13.00%	10.98%	32.01%	36.63%
	Slope	0.0207	0.0144	0.0112	0.0082	0.0066	0.0068	0.0053	0.0046
Unscaled ER^V	Hodrick t	(3.55)	(3.26)	(2.59)	(1.88)	(1.62)	(1.69)	(1.62)	(1.93)
	HH t	(3.27)	(2.47)	(2.16)	(2.00)	(2.37)	(3.09)	(3.15)	(4.66)
	Adj. R^2	4.22%	6.19%	7.32%	9.31%	16.49%	29.42%	31.69%	47.20%
	Slope	0.014	-0.013	-0.026	-0.015	-0.012	-0.011	-0.010	-0.008
Scaled VRP	Hodrick t	(0.54)	(-0.96)	(-2.26)	(-1.42)	(-1.18)	(-1.24)	(-1.41)	(-1.64)
	HH t	(0.44)	(-0.71)	(-1.98)	(-1.53)	(-1.89)	(-1.94)	(-2.00)	(-2.13)
	Adj. R^2	-0.41%	-0.27%	1.55%	1.03%	2.37%	3.32%	4.75%	5.84%
	Slope	-0.92	-0.95	-1.35	-1.36	-0.89	-0.50	-0.26	-0.37
VVIX	Hodrick t	(-1.19)	(-1.41)	(-2.20)	(-2.58)	(-2.20)	(-1.61)	(-1.67)	(-2.89)
	HH t	(-0.99)	(-1.19)	(-2.01)	(-2.64)	(-2.65)	(-1.82)	(-1.18)	(-2.18)
	Adj. R^2	-0.01%	1.11%	6.09%	13.65%	14.00%	6.27%	2.92%	10.18%
	Slope	3.03	-2.26	-5.68	-6.44	-4.58	-3.35	-3.32	-2.24
LTV	Hodrick t	(0.87)	(-0.97)	(-2.48)	(-3.13)	(-2.34)	(-1.95)	(-2.40)	(-2.67)
	HH t	(0.84)	(-0.63)	(-1.81)	(-2.99)	(-3.45)	(-3.62)	(-4.97)	(-4.52)
	Adj. R^2	-0.17%	-0.05%	5.24%	16.90%	23.36%	24.67%	42.50%	39.66%
	Slope	-43.40	-31.72	-22.53	-13.74	-11.70	-11.21	-5.77	-3.86
Slope	Hodrick t	(-4.60)	(-3.70)	(-2.96)	(-1.84)	(-1.77)	(-2.04)	(-1.27)	(-1.19)
	HH t	(-3.52)	(-2.97)	(-2.30)	(-1.80)	(-2.40)	(-3.22)	(-1.61)	(-1.54)
	Adj. R^2	6.55%	8.91%	7.69%	6.29%	12.11%	22.93%	10.09%	9.45%

Table A3: Forecasting S&P 500 Returns with Other Predictors

This table reports slopes (scaled by 1,000 for RV, VIX, VRP, VVIX, LTV, and Slope) from predictive regressions of S&P 500 returns on alternative predictors suggested by the existing literature. T-statistics are computed according to Hansen and Hodrick (1980) and Hodrick (1992). The predictors are realized variance (RV), the VIX index, the variance risk premium (VRP), the VVIX index, the tail risk index (LTV), and the VIX term structure (Slope). The sample periods are as follows: VIX, April 2004 to November 2022; RV and VRP, April 2004 to December 2020; VVIX: August 2006 to November 2022; Slope, April 2004 to December 2017; LTV, April 2004 to December 2019.

Horizon (months)		1	3	6	12	24	36	48	60
	Slope	0.12	0.23	0.22	0.16	0.10	0.09	0.08	0.08
VRP	Hodrick t	(0.65)	(1.94)	(2.60)	(2.56)	(1.87)	(1.68)	(1.90)	(2.77)
	HH t	(0.73)	(2.16)	(2.31)	(2.04)	(1.82)	(1.66)	(1.73)	(2.11)
	Adj. R^2	-0.24%	2.06%	3.70%	3.81%	3.85%	4.81%	5.89%	9.47%
	Slope	0.03	0.05	0.08	0.09	0.05	0.04	0.04	0.04
RV	Hodrick t	(0.24)	(0.76)	(1.61)	(2.25)	(2.24)	(1.49)	(1.87)	(2.49)
	HH t	(0.43)	(0.86)	(1.51)	(2.01)	(1.88)	(1.38)	(1.67)	(1.98)
	Adj. R^2	-0.42%	0.34%	3.04%	6.84%	6.72%	5.08%	8.31%	12.73%
	Slope	0.03	0.06	0.08	0.08	0.05	0.04	0.04	0.04
VIX	Hodrick t	(0.37)	(1.07)	(1.87)	(2.46)	(2.23)	(1.58)	(1.91)	(2.57)
	HH t	(0.61)	(1.16)	(1.70)	(2.17)	(1.94)	(1.50)	(1.76)	(2.14)
	Adj. R^2	-0.29%	0.96%	4.11%	8.63%	8.56%	7.80%	11.55%	17.76%
	Slope	-0.004	-0.002	-0.002	-0.001	-0.001	-0.001	-0.001	-0.001
Unscaled ER^V	Hodrick t	(-2.42)	(-2.19)	(-1.90)	(-1.23)	(-1.51)	(-1.76)	(-1.71)	(-2.36)
	HH t	(-2.79)	(-1.66)	(-1.19)	(-0.87)	(-1.29)	(-1.57)	(-1.53)	(-2.07)
	Adj. R^2	2.98%	2.85%	2.49%	1.78%	6.33%	12.99%	13.08%	21.70%
	Slope	0.001	0.006	0.006	0.004	0.003	0.002	0.002	0.002
Scaled VRP	Hodrick t	(0.11)	(1.50)	(2.32)	(1.88)	(1.27)	(1.37)	(1.58)	(2.80)
	HH t	(0.10)	(1.41)	(1.84)	(1.54)	(1.46)	(1.36)	(1.36)	(1.80)
	Adj. R^2	-0.51%	0.45%	1.43%	1.16%	1.38%	1.74%	2.07%	4.47%
	Slope	0.23	0.24	0.28	0.30	0.23	0.17	0.09	0.10
VVIX	Hodrick t	(1.28)	(1.52)	(1.76)	(2.54)	(2.32)	(3.19)	(3.17)	(4.52)
	HH t	(1.15)	(1.27)	(1.50)	(2.01)	(2.33)	(2.34)	(1.46)	(1.84)
	Adj. R^2	0.16%	1.44%	3.79%	8.67%	11.53%	10.53%	4.70%	8.53%
	Slope	-0.29	0.74	1.47	1.47	0.96	0.87	0.77	0.63
LTV	Hodrick t	(-0.27)	(0.94)	(2.08)	(2.74)	(2.31)	(2.27)	(2.39)	(4.14)
	HH t	(-0.39)	(0.94)	(1.94)	(2.56)	(2.37)	(3.00)	(3.16)	(3.26)
	Adj. R^2	-0.46%	0.62%	6.82%	14.15%	13.65%	20.75%	24.36%	29.09%
	Slope	9.01	5.70	3.33	1.35	2.79	2.90	2.03	1.51
Slope	Hodrick t	(3.22)	(2.31)	(1.83)	(0.89)	(1.84)	(2.06)	(1.72)	(2.79)
	HH t	(3.43)	(2.27)	(1.28)	(0.62)	(1.95)	(2.70)	(1.95)	(1.89)
	Adj. R^2	6.21%	5.58%	2.58%	0.40%	9.67%	19.60%	13.74%	13.89%

Table A4: Forecasting S&P 500 Returns with the VRP, RV and VIX. Extended Sample 1990-2020

We report results on forecasting S&P 500 index returns with realized variance (RV), the VIX, and the VRP. The sample period is from January 1990 to December 2020.

Panel A: VRP								
Horizon (months)	1	3	6	12	24	36	48	60
Slope	0.07	0.17	0.11	0.04	0.01	0.01	-0.01	-0.02
Hodrick t	(0.34)	(2.03)	(1.92)	(0.88)	(0.32)	(0.26)	(-0.28)	(-0.41)
HH t	(0.80)	(2.85)	(2.33)	(0.88)	(0.29)	(0.23)	(-0.26)	(-0.40)
Adj. R^2	-0.10%	2.46%	1.95%	0.17%	-0.21%	-0.20%	-0.18%	-0.05%
Panel B: RV								
Horizon (months)	1	3	6	12	24	36	48	60
Slope	0.001	-0.016	0.018	0.024	0.013	-0.007	-0.008	0.001
Hodrick t	(0.01)	(-0.24)	(0.32)	(0.69)	(0.53)	(-0.23)	(-0.34)	(0.04)
HH t	(0.01)	(-0.38)	(0.50)	(0.79)	(0.52)	(-0.22)	(-0.30)	(0.03)
Adj. R^2	-0.27%	-0.20%	-0.09%	0.33%	0.04%	-0.19%	-0.11%	-0.30%
Panel C: VIX								
Horizon (months)	1	3	6	12	24	36	48	60
Slope	0.036	0.058	0.080	0.054	0.024	-0.004	-0.013	-0.004
Hodrick t	(0.31)	(0.66)	(1.12)	(1.11)	(0.64)	(-0.10)	(-0.41)	(-0.15)
HH t	(0.58)	(0.99)	(1.47)	(1.06)	(0.56)	(-0.09)	(-0.34)	(-0.12)
Adj. R^2	-0.18%	0.39%	2.16%	1.73%	0.47%	-0.26%	0.10%	-0.25%

Table A5: VIX Futures Returns Based on Univariate Sorts

We report average VIX futures returns for tercile portfolios sorted on the variance risk premium, the VVIX, and the tail risk factor (LTV). The Satterthwaite t-statistics for testing the equality of the means of the L and H groups are reported in parentheses.

Panel A: Sort on VRP								
Horizon (months)	1	3	6	12	24	36	48	60
L	-0.079	-0.058	-0.047	-0.056	-0.052	-0.052	-0.056	-0.057
2	-0.047	-0.043	-0.050	-0.052	-0.056	-0.057	-0.057	-0.061
H	-0.079	-0.103	-0.103	-0.089	-0.085	-0.081	-0.076	-0.076
H-L	-0.0002	-0.045	-0.057	-0.034	-0.033	-0.029	-0.020	-0.019
t-statistic	(-0.01)	(-2.09)	(-3.64)	(-3.02)	(-5.17)	(-6.42)	(-5.21)	(-7.22)
Panel B: Sort on VVIX								
Horizon (months)	1	3	6	12	24	36	48	60
L	-0.054	-0.050	-0.038	-0.039	-0.042	-0.061	-0.064	-0.063
2	-0.045	-0.051	-0.052	-0.052	-0.068	-0.063	-0.066	-0.067
H	-0.090	-0.085	-0.095	-0.095	-0.080	-0.075	-0.071	-0.073
H-L	-0.036	-0.035	-0.058	-0.056	-0.038	-0.014	-0.007	-0.010
t-statistic	(-1.06)	(-1.52)	(-3.53)	(-5.51)	(-5.93)	(-2.84)	(-1.80)	(-3.56)
Panel C: Sort on LTV								
Horizon (months)	1	3	6	12	24	36	48	60
L	-0.061	-0.059	-0.050	-0.047	-0.045	-0.045	-0.050	-0.057
2	-0.101	-0.073	-0.052	-0.048	-0.057	-0.060	-0.060	-0.061
H	-0.053	-0.074	-0.094	-0.097	-0.089	-0.084	-0.080	-0.077
H-L	0.008	-0.015	-0.045	-0.049	-0.043	-0.040	-0.030	-0.021
t-statistic	(0.20)	(-0.62)	(-2.87)	(-4.96)	(-7.46)	(-9.95)	(-9.44)	(-8.40)

Figure A1: Finite-Sample Power in Predictive Regressions

This figure plots the power curves for various computations of t-statistics. The results are based on a parameterization calibrated to the baseline expected return. The top panel is for the one-year horizon and the bottom panel is for the three-year horizon.