MPRA

Predicting the equity premium with the implied volatility spread

Cao, Charles and Simin, Timothy and Xiao, Han

Smeal College of Business, Pennsylvania State University

December 2019

Predicting the equity premium with the implied volatility spread

Charles Cao[*]

Timothy Simin[†]

Han Xiao[‡]

[*] Cao is at the Smeal College of Business, The Pennsylvania State University, University Park, PA 16802, (814) 865-7891, qxc2@psu.edu.
[†] Simin (corresponding author) is at the Smeal College of Business, 345 Business bldg., The Pennsylvania State University, University Park, PA 16802, (814) 865-3457, tsimin@psu.edu.
[‡] Xiao is at the Smeal College of Business, The Pennsylvania State University, University Park, PA 16802, (814)-826-9875, hxx15@psu.edu.

Abstract

We show that the call-put implied volatility spread (*IVS*) outperforms many well-known predictors of the U.S. equity premium at return horizons up to six months over the period from 1996:1 to 2017:12. The predictive ability of the *IVS* is unrelated to the dividend yield and is useful in explaining the cross-section of returns. Decomposing the *IVS*, we find the longer run predictive ability of the *IVS* operates primarily through a cash flow channel. We also find the *IVS* is significantly related to indicators of aggregate market direction and expected market conditions. Our results are consistent with the *IVS* reflecting market sentiment as well as information about informed trading.

Introduction

Deviations from put-call parity captured by option implied volatilities contain information about future stock returns. The intuition behind this relation provided by the current state of the literature is that price pressure from informed trading in both the options and equity markets causes violations of put-call parity resulting in the option implied volatilities differing from their no-arbitrage values. The implied volatility spread (*IVS*) between the call and put options of individual firms is a measure of the informed trading deviation from put-call parity. While the *IVS* predicts individual stock returns (Cremers and Weinbaum, 2010) and aggregate short-run market returns over horizons up to one week (Atilgan et al., 2015), there is evidence that changes in implied volatilities (An et al., 2014) and the shape of the volatility smirk (Xing et al., 2010) predict portfolio returns for at least six months. The relations between implied volatilities and longer-horizon returns suggest that deviations from put-call parity may reflect information other than the short-term price pressure of informed traders.

In models where informed traders choose to trade in the options market first (Easley et al., 1998; An et al., 2014), options markets can predict stock returns. In these noisy rational expectations models of informed trading, the existence of noise traders allows informed traders to mask the information content of trades leading to inefficient, and hence predictable prices. If the information set of informed traders includes market sentiment, then trading in the options market may also reflect long-run information. The long-run information content of market sentiment is plausible behaviorally if people's optimism (pessimism) develops into a consensus view indicating the importance of sentiment may build over time, and rationally, if arbitrage forces, which are likely to eliminate short-run mispricing, fail at longer horizons. An example of this limit-to-arbitrage is the noise trader risk described in DeLong et al. (1990). Indeed, Brown and Cliff (2005)

find that while sentiment predicts longer-run returns, they find little predictive evidence for near-term returns. If the *IVS* captures the informed trader's assessment of market sentiment, then the *IVS* is an economically sensible predictor of the longer-run market risk premium.

Our first contribution is to demonstrate that the call-put implied volatility spread predicts the aggregate equity premium at return horizons well beyond one week. We first consider the in-sample predictive power of the *IVS* in comparison to other conventional variables investigated in Goyal and Welch (2008). We regress the equity premium on the individual lagged predictors for 1, 3, 6, and 12-month return horizons. Our results indicate that the *IVS* is a significant predictor of the market risk premium for up to 12 months. In this experiment, the *IVS* clearly dominates many of the other 14 predictors we examine.

Because in-sample prediction tests can suffer from several biases, including the Stambaugh (1999) coefficient bias caused by autocorrelated predictors, data snooping biases as in Ferson et al. (2003), and finite sample biases described in Nelson and Kim (1993), we test the out-of-sample performance of the *IVS* as a predictor. We evaluate the predictive performance of each predictor based on univariate ordinary least squares (OLS) regressions and the utility gains that a mean-variance utility investor would obtain by holding the portfolio formed based on the predictors relative to a portfolio based on the historical average benchmark forecast. A significant positive out-of-sample R-square (R_{OS}^2), in conjunction with a positive annualized utility gain, suggests that a predictor performs well out-of-sample. Compared to using the historical average return as a predictor, we find the *IVS* produces significantly smaller one-step-ahead mean square forecast errors and generates positive utility gains up to six months ahead. The *IVS* is the only predictor to produce significantly positive out-of-sample R-squares out of the set of 14 traditional predictors. We show that this result is robust to controls, model specification, and measurement of the *IVS*.

2

Our second contribution is to demonstrate that the *IVS* and the dividend yield contain different sources of predictive ability. We first show that the component of the risk premium that is uncorrelated with the *IVS* is predictable only by the dividend yield. When we orthogonalize the market risk premium relative to the other significant predictors, the *IVS* remains a significant predictor. These results indicate that the *IVS* and the dividend yield contain distinct information relevant to predicting future market returns. Next, we show that during expansionary periods, the dividend yield generates positive utility gains as a significant predictor of the market risk premium, while during recessions only the *IVS* remains a significant predictor while producing positive utility gains, suggesting an important difference in predictability between the *IVS* and the dividend yield. Finally, we compare the ability of the *IVS*, versus the dividend yield, to explain a large cross-section of portfolios by utilizing the predictors as instrumental variables in conditional versions of the CAPM, Fama and French three-factor, and the Carhart four-factor models. The conditional versions of these models, where the time-varying expected market risk premium is modeled as a linear function of *IVS*, produce 71%-85% fewer significant intercepts relative to their unconditional versions. Using the dividend yield to model the expected risk premium, we find at most 65% fewer significant alphas, and in the Carhart model conditioning on the dividend yield worsens the model's ability to explain the cross-section of returns.

Our third contribution is to show that the predictive ability of *IVS* captures more than just the short-term impacts of informed trading. Decomposing the predictive power of the *IVS* into the expected return, cash flow, and discount rate innovations, we find that the predictive ability of the *IVS* comes through the cash flow channel versus the expected return or discount rate channel. We then demonstrate that the predictive ability of *IVS* reflects information related to measures of market expectations and several measures of market uncertainty. Our proxies for market

expectations come from three surveys: the Gallup investor survey, the American Association of Individual Investors survey, and the crash confidence index from the Yale School of Management. We also include a statistical measure of sentiment from Baker and Wurgler (2006). Our proxies for aggregate uncertainty come from Jurado et al. (2015) and Baker et al. (2016). *IVS* is positively and significantly related to all the measures of market expectations. *IVS* is negatively and significantly related to all eight proxies of macroeconomic, political, and financial uncertainty. We take these results to indicate that *IVS* captures the impact of market sentiment on the cash flows of firms.

Historically, information from options markets is utilized to predict equity return volatility. Christoffersen et al. (2013) provide a comprehensive review of this literature. Increasingly, information extracted from option markets is used to predict market returns. For instance, Chordia et al, (2018) show how the order imbalance of stock index options predicts weekly S&P 500 returns. Chen and Liu (2018) predict market returns using an estimate of implied volatility from bid and ask prices of deep out-of-the-money put options on the S&P 500 Index. Feunou et al. (2012) show that the term structure of option-implied variances drives the equity premium. Bali et al. (2015) use a novel option-implied measure of risk to predict the equity risk premium and So et al. (2016) generate an option-implied measure of ambiguity to predict international market returns. The option-based predictor of the market risk premium that has garnered the most attention is the spread between option-implied and realized volatility known as the variance risk premium. A few of the influential papers using the variance risk premium include Bollerslev et al. (2009), Bekaert et al. (2013), Bollerslev et al. (2014), Zhou (2018), and Hollstein et al. (2019).

Our paper adds to a growing segment of this literature that shows that deviations from put-call parity captured by the *IVS* contain useful information for understanding equity returns. For

instance, Bali and Hovakimian (2009) show that implied volatility spreads are positively related to the cross-section of expected returns, and An et al. (2014) find an inverse relation between future returns and substantial changes in implied volatilities. Doran et al. (2013) use the *IVS* as a measure of informed trading relative to informed trading measures from the equity market. The *IVS* also roughly captures aggregate jump and tail risk (Bollerslev and Todorov, 2011; Kelly and Jiang, 2014) and may contain nonlinear risk information across different economic states such as rare disasters (Gabaix, 2012). Our contribution to this literature is to demonstrate that the *IVS* reflects market sentiment about the state of the economy, and through the cash flow channel, provides longer run predictability of the market risk premium.

The organization of the paper is as follows. In Section 1, we define the variables, data sets, and methodologies used. In Section 2, we provide in-sample and out-of-sample prediction results. In Section 3, we discuss three experiments to differentiate the source of the *IVS*' predictive ability from that of the other significant predictors. In Section 4, we explore why the *IVS* predicts future returns. Here we decompose returns based on Campbell and Shiller (1988) to see which components are predictable by the *IVS*. We also show how the *IVS* is related to market sentiment and show how the *IVS* and the variance risk premium are complementary predictors. We conclude in Section 5.

1. Data, variables, and methodology

1.1 Data

Our monthly data come from several sources. The option-implied volatility is from OptionMetrics. The conventional predictors, value-weighted CRSP stock returns, and risk-free rate are from Amit Goyal's website. The index for business cycles comes from the NBER. The common risk factors

for the Fama-French model and our test assets consist of monthly excess returns of 399 equally-weighted portfolios from Ken French's database. These include five sets of decile portfolios sorted on past variance, residual variance, net share issues, beta, or accruals; three sets of portfolios double-sorted into deciles based on size and book-to-market, size and operating profitability, or size and investment; and 49 industry portfolios. Due to restrictions in the option data, our main results are based on monthly data from 1996:1 to 2017:12. To be comparable with Atilgan et al. (2015), we apply the variance risk premium as a covariate to control for conditional variance risk, which comes from Hao Zhou's website. We test the source of the predictive ability of the *IVS* using conventional predictors from Amit Goyal's website and implied volatility computed using CRSP.[1]

1.2 The call/put implied volatility spread

We construct our monthly *IVS* measure using daily option-implied volatilities in OptionMetrics. The volatility surface represents the separately interpolated implied volatility surface for puts and calls, computed using a methodology based on a kernel smoothing algorithm with expirations of 30 to 730 calendar days, at deltas of 0.20 to 0.80 (negative deltas for puts). The underlying implied volatilities of individual options are computed using binomial trees that account for the early exercise of individual stock options and the dividends expected to be paid over the lives of the options. Using the volatility surface avoids having to make potentially arbitrary decisions on which strikes or maturities to include when computing an implied call or put volatility for each stock. Here we use implied volatilities with an absolute delta of 0.5*,* i.e., at-the-money options, and an expiration of 30 days. Our construction of the *IVS* differs from Atilgan et al. (2015) where they

---

[1] For detailed data construction and source information, see Ken French's Data Library
(https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html), Goyal's website (http://www.hec.unil.ch/agoyal/),
NBER Business Cycle indicator from http://www.nber.org/cycles.html, and. Zhou's website
(https://sites.google.com/site/haozhouspersonalhomepage/).

use OTM put options and ATM call options, i.e., their method uses the slope of the volatility smile to create the *IVS*.

The *IVS* for individual stocks is defined as the difference between their call option volatility and put option volatility. We then calculate the equally-weighted average of the *IVS*. Since we focus on predicting the return on the value-weighted market portfolio, it seems natural to use a value-weighted version of the *IVS* or directly use option-implied volatilities from the S&P 500 Index options (*SPX*) as in Atilgan et al. (2015). Our choice to utilize the equally-weighted *IVS* is supported by the intuition outlined in Rapach et al. (2016), who make the compelling argument that equal-weighting reflects more information concerning the trades of informed market participants relative to value-weighting given the skewness in the distribution of market capitalization. They argue that if informed trading is less important for large cap stocks, then equally weighting predictive measures that reflect the information content of the informed trading will provide a better predictive signal. We confirm the intuition of Rapach et al. (2016) in the out-of-sample analysis below and focus our results on an equally-weighted version of *IVS*.[2]

1.3 Common predictors

Following Goyal and Welch (2008), we define the equity premium as the logarithm of the CRSP value-weighted market return minus the logarithm of the prevailing short-term U.S. Treasury bill rate. The log of equity premium has a mean (standard deviation) of 0.015 (0.085) from 1996 to 2017.

---

[2] In the Appendix, we check the robustness of the equally-weighted *IVS* compared to several other possible specifications. In particular, we present the in- and out-of-sample results for versions of the *IVS* calculated using the implied volatility for options with 91, 182, 273, and 365 days to expiration and two versions of the *IVS* that account for a potential look-ahead bias caused by using historical dividend payments.

The set of predictor variables includes the following.

(1) Fundamental valuations:

- Logarithm of the dividend-price ratio ($\ln(DP)$): the log of 12-month moving sum of dividends paid on the S&P 500 Index minus the log of S&P 500 Index price.

- Logarithm of the dividend-yield ratio ($\ln(DY)$): the log of 12-month moving sum of dividends paid on the S&P 500 Index minus the log of lagged S&P 500 Index price.

- Logarithm of the earnings-price ratio ($\ln(EP)$): the log of 12-month moving sum of earnings on the S&P 500 Index minus the log of S&P 500 Index price.

- Logarithm of the dividend-payout ratio ($\ln(DE)$): the log of 12-month moving sum of dividends paid on the S&P 500 Index minus the log of 12-month moving sum of earnings on the S&P 500 Index.

- Stock Variance ($SVAR$): the sum of squared daily returns on the S&P 500 Index.

- Book-to-Market Ratio (BM): the ratio of book value to market value for the Dow Jones Industrial Average.

- Net Equity Expansion ($NTIS$): the ratio of 12-month moving sums of net issues by NYSE listed stocks divided by the total end-of-year market capitalization of NYSE stocks.

(2) Interest rate related variables:

- Treasury Bills (TBL): the 3-month Treasury bill secondary market rate.

- Long Term Yield (LTY): the long-term government bond yield.

- Long Term Rate of Returns (LTR): the rate of return on long-term government bond.

- Term Spread (TMS): the difference between the yield on long-term government bonds and the Treasury bill.

- Default Yield Spread (DFY): the difference between BAA and AAA-rated corporate bond yields.

- Default Return Spread (DFR): the difference between long-term corporate bond return and long-term government bond returns.

(3) Macroeconomic indicators:

- Inflation ($INFL$): the lag of the inflation calculated from the Consumer Price Index (All Urban Consumers).

- Investment to Capital Ratio (i/k): the ratio of aggregate (private nonresidential fixed) investment to aggregate capital for the whole economy.

In the Appendix, we report the detailed statistics of the predictor variables, the equity premium across quantiles, and the correlation matrix. Table 1 provides the mean, standard deviation, and autocorrelation summary statistics. The monthly equal-weighted *IVS* has the mean (standard error) of -0.008 (0.005), with a first-order autocorrelation of 0.372. The lagged predictor variables are often highly persistent, namely nine of the predictors have first-order autocorrelations of at least 0.90.

1.4 Methodology

In this subsection, we define the OLS regressions for the in-sample and step-ahead tests. For the step-ahead experiments, we test the hypothesis that the out-of-sample R-squared is significant based on Clack-West (2007) adjusted mean-square forecast error (MSFE) *F*-test. We also evaluate the step-ahead forecasts with a utility-based method as a measure of the economic value. By considering the difference between the forecasts of predictors versus the historical average forecast in the asset allocation decision, any utility gain can be interpreted as the portfolio management fee an investor would be willing to pay to have access to the predictive information.

1.4.1 OLS prediction

We first consider the in-sample predictive regression:

$$r_{t+h} = \alpha_{i,h} + \beta_{i,h} x_{i,t} + \varepsilon_{i,t+h}, \tag{1}$$

where $r_{t+h} = \left(\prod_{i=1}^{h} 1 + r_{t+i} \right) - 1$, for h = 1, 3, 6, 12 month, represents the return on a stock market index in excess of the risk-free interest rate from $t+1$ to $t+h$, and $x_{i,t}$ represents the ith predictor

where $X = \{IVS, \ln(DP), \ldots\}$ and $\varepsilon_{i,t+1}$ is the error term of the variable i's prediction for $t = 1, 2, \ldots,$ $T\text{-}h$. We use the autocorrelation-heteroskedasticity-consistent (HAC) standard errors from Newey and West (1987), with an automatic lag selection procedure as suggested in Ferson et al. (2003) to account for overlapping issues. The number of lags is chosen by computing the autocorrelations of the estimated residuals and truncating the lag length when the sample autocorrelations become "insignificant" at longer lags. Explicitly, we compute 12 sample autocorrelations and compare the values with a cutoff at two approximate standard errors: $2/\sqrt{T}$, where T is the sample size. The number of lags chosen is the minimum lag length at which no higher order autocorrelation is larger than two standard errors.

As in Goyal and Welch (2008), we generate out-of-sample forecasts of the equity premium using an expanding estimation window. More specifically, we first divide the total sample of T observations into a training period composed of the first n_1 observations and an out-of-sample portion composed of the $n_2 = T - n_1$ observations. The initial out-of-sample forecast of the equity premium using the predictor $x_{i,t}$ is given by:

$$\hat{r}_{i,m+1} = \hat{\alpha}_{i.m} + \hat{\beta}_{i.m} x_{i.m}, \tag{2}$$

where $\hat{\alpha}_{i.m}$ and $\hat{\beta}_{i.m}$ are the OLS estimates of α_i and β_i, respectively, generated by regressing $\{r_t\}_{t=2}^{m}$ on a constant and $\{x_{i,t}\}_{t=1}^{m-1}$. The out-of-sample forecast is given by:

$$\hat{r}_{i,m+2} = \hat{\alpha}_{i.m+1} + \hat{\beta}_{i.m+1} x_{i.m+1}, \tag{3}$$

where $\hat{\alpha}_{i.m+1}$ and $\hat{\beta}_{i.m+1}$ are generated by regressing $\{r_t\}_{t=2}^{m+1}$ on a constant and $\{x_{i,t}\}_{t=1}^{m}$. Proceeding in this manner through the end of the out-of-sample period, we generate a series of n_2 out-of-sample forecasts of the equity premium based on $x_{i,t}$, $\{\hat{r}_{i,t+1}\}_{t=m}^{T-1}$. Our training period is 1996:1 to

2006:12. Tests of forecast evaluation are done for the 2007:1–2017:12 period. Alternative divisions

of in-sample and out-of-sample periods produce similar results.

1.4.2 Forecast evaluation

We use the mean squared forecast error (*MSFE*), the out-of-sample R-squared $\left(R_{OS}^2 \right)$, and a utility-

based measure to evaluate the quality of the different forecasts. For a given predictor x_i, the *MSFE*

is given by:

$$MSFE_i = \frac{1}{n_2} \sum_{s=1}^{n_2} \left(r_{n_1+s} - \hat{r}_{i,n_1+s} \right)^2 , \tag{4}$$

where i is the index of predictors. We use the historical equity mean as a benchmark, i.e., we

assume a constant expected excess return $\bar{r}_{t+1}$, and calculate its *MSFE* as:

$$MSFE_0 = \frac{1}{n_2} \sum_{s=1}^{n_2} \left(r_{n_1+s} - \bar{r}_{n_1+s} \right)^2 . \tag{5}$$

The out-of-sample R_{OS}^2 statistic, suggested by Campbell and Thompson (2008), measures the

proportional reduction in *MSFE* for the predictive regression forecast relative to the historical

average:

$$R_{OS}^2 = 1 - \left(MSFE_i / MSFE_0 \right). \tag{6}$$

If $R_{OS}^2 > 0$, then $MSFE_i < MSFE_0$ and the predictor is more accurate than the historical mean. We

calculate the *p*-value for the Clark-West (2007) *MSFE*-adjusted statistic to evaluate the statistical

significance of R_{OS}^2. To test $H_0 : R_{OS}^2 \le 0$ against $H_A : R_{OS}^2 > 0$, we first calculate:

$$\tilde{d}_{i,n1+s} = \left(r_{n1+s} - \bar{r}_{n1+s} \right)^2 - \left[\left(r_{n_1+s} - \hat{r}_{i,n_1+s} \right)^2 - \left(\bar{r}_{n1+s} - \hat{r}_{i,n_1+s} \right)^2 \right], \tag{7}$$

then regress $\tilde{d}_{i,n1+s}$ on a constant. The *MSFE*-adjusted statistic is the *t*-statistic corresponding to the constant. When comparing forecasts from non-nested models, Diebold and Mariano (2002) and West (1996) show that the *MSFE*-adjusted statistic has an asymptotic standard normal distribution.

For our utility-based forecast evaluation, we consider a mean-variance investor with relative risk aversion γ who allocates wealth between stocks and risk-free bills based on the forecast of the equity premium. At time t, the investor allocates the following share of the portfolio to equities during $t+1$:

$$a_{i,t} = \left(\frac{1}{\gamma}\right)\left(\frac{\hat{r}_{i,t+1}}{\hat{\sigma}_{t+1}^2}\right), \tag{8}$$

where $\hat{\sigma}_{t+1}^2$ is a forecast of the stock return's variance. Over the forecast evaluation period, the investor realizes the average utility:

$$\hat{v}_i = \hat{\mu}_i - 0.5\gamma\hat{\sigma}_i^2, \tag{9}$$

where $\hat{\mu}_i$ and $\hat{\sigma}_i^2$ are the sample mean and variance of the portfolio formed on the basis of $\hat{r}_{i,t+1}$ and $\hat{\sigma}_{t+1}^2$ over the forecast evaluation period. If the investor instead relies on the historical average forecast of the equity premium, $\overline{r}_{i,t+1}$, the realized average utility over the forecast evaluation period is:

$$\hat{v}_0 = \hat{\mu}_0 - 0.5\gamma\hat{\sigma}_0^2, \tag{10}$$

where $\hat{\mu}_0$ and $\hat{\sigma}_0^2$ are the sample mean and variance of the portfolio formed on the basis of $\bar{r}_{i,t+1}$

and $\hat{\sigma}_{t+1}^2$ over the forecast evaluation period. We set the risk aversion parameter to 5 and define

the utility gain as $\Delta = \hat{v}_i - \hat{v}_0$.[3]

2. Prediction results

2.1 In-sample tests

Table 2 contains the results of in-sample regressions of the equity premium on the lagged predictor

variables up to 12 months ahead based on our predictive specification in equation (1). We use

Newey-West heteroscedasticity-autocorrelation consistent standard errors with lags chosen as in

Ferson et al. (2003) to calculate *t*-statistics. The results are consistent if we use Hansen-Hodrick *t*-

statistics. Market excess returns and predictors are standardized to have a mean zero and standard

deviation of one, thus the coefficients are comparable. For our 1996:1 to 2017:12 sample, we have

264, 262, 259, and 253 overlapping observations for estimating equation (1) at one-month (*h*=1),

three-months (*h*=3), semi-annual (*h*=6), and annual (*h*=12) horizons, respectively.

The results indicate that at the monthly horizon (columns (2) – (3)), *IVS* has significant

predictive ability for up to 12 months. *IVS* is one of four significant in-sample predictors, along

with ln(*DP*), ln(*DY*), and *SVAR* at the monthly frequency. A one standard error increase in *IVS*

predicts a 19 bps higher market excess returns in the next month. For the conventional predictors,

only ln(*DY*), Ln(*DP*), and *SVAR* exhibit significant coefficients for the one-month horizon. While

not tabulated, the R^2 of the implied volatility spread is the largest (3.3%) among Goyal and Welch

---

(2008) predictors. Campbell and Thompson (2008) argue that a monthly R^2 statistic of approximately 0.5% represents an economically meaningful degree of return predictability. The monthly R^2 statistics for the significant predictors are above this threshold.

At longer predictive horizons, the implied volatility spread maintains a strong predictive ability. At the three-month horizon, the coefficient for *IVS* is 0.26 in column (4) with a Newey-West *t*-statistic of 4.33. Here, a one standard deviation increase in *IVS* corresponds to an increase of 26 bps in excess market returns in the following quarter. Among remaining predictors, only ln(*DP*), ln(*DY*), *SVAR*, *BM*, and *NTIS* have significant predictions. Moreover, the *IVS*'s R^2 is 6.8% compared to 4.3% for the next highest R^2 (*NTIS*). The semi-annual and annual predictions of the implied volatility spread are substantially significant. Even though there are some predictors (e.g., ln(*DP*), ln(*DE*), and *BM*) that gain significant predictability after six months, the results come from high autocorrelation and overlapping effects. Most of the standard predictors fail to predict the equity premium at the longer horizons.

2.2 Out-of-sample tests

Table 3 reports one-month to 12-month ahead forecasts from the set of predictor variables relative to forecasts based on the historical average. Columns (1) – (3) include one-month ahead predictions, and Columns (4) – (12) represent predictions using three to twelve months ahead market excess returns. R^2_{OS} is the out of sample R^2 in equation (6), and *pval* is the corresponding *p*-value based on Clark-West (2007) *MSFE*-adjusted test in equation (8). Here Δ is the utility gain measure calculated from the difference between equations (9) and (10).

The significantly positive R^2_{OS} for the *IVS* for one-month ahead forecasts indicates that the *IVS* produces a significantly smaller mean-squared-error compared to the benchmark historical mean.

The higher the R^2_{OS} represents the better predictor performance. Moreover, the positive utility gain (Δ) implies that if the mean-variance investor allocates between the market portfolio and the risk-free bond based on the predictor, the utility will be higher than that of a portfolio allocation based on the benchmark historical average. A higher utility gain indicates a more valuable predictor.

The results imply that the *IVS* is the only predictor out of 15 predictors with a positive R^2_{OS} one-month ahead, while conventional predictors, for example ln(*DP*), underperformed out-of-sample. The *IVS* produces a significantly large increase (146 bps) in utility over the benchmark forecast at the one-month frequency, comparing to the next highest 87 bps (ln(*EP*)). This means that investors prefer to pay more fees on learning from the *IVS* and can get a higher utility level by holding a mean-variance portfolio based on the *IVS*. Both out-of-sample R^2 and utility gain measures support the predictive ability of the *IVS*.

Notably, the *IVS* continues to predict significantly better than the benchmark historical average in the three- (R^2_{OS} =3.18) and six-month ahead (R^2_{OS} =2.58) at the 5% significance level. The positive utility gain (310 bps and 44 bps, respectively) supports the results. However, unlike the in-sample prediction, the out-of-sample prediction finds no 12-month ahead prediction (R^2_{OS} =-1.41, and the utility gain is -218 bps). The results are consistent with results of Ang et al. (2014), who find changes in option-implied volatility provide predictability up to six months.

Consistent with Goyal and Welch (2008), the conventional predictors exhibit no significant out-of-sample predictability in terms of both R^2_{OS} and utility gains. Several of the conventional predictors occasionally exhibit either positive R^2_{OS} or positive utility gains at longer horizons; however, these two measures are either insignificant or contradictory. For example, *SVAR* is insignificantly positive (0.22 with a 23% *pval*) in three-month horizon yet the utility gain is

negative (-254 bps). The dividend yield (ln(*DY*)) produces strong forecasts at the six-month horizon (3.19 with 1% significance), nevertheless the utility gain is -43.34%. The results are consistent with the high autocorrelation underlying these predictors.

We also confirm the intuition of Rapach et al, (2016) that equally-weighting the *IVS* reflects more information on the informed buys or sells rather than value-weighting or the *IVS* constructed from the *SPX* index. We find that both the value-weighted version of the *IVS* and the *IVS* constructed from the *SPX* index options fail to provide any predictive ability as measured by R^2_{OS} or utility gains at any of the time horizons. These results suggest why our equally weighted version of the *IVS* exhibits predictive ability while others such as An et al. (2014), who use a value-weighted version of the *IVS*, and Atilgan et al. (2015) who create a version of the *IVS* utilizing implied volatilities from call and put options on the S&P 500 Index, do not.[4]

3. Does *IVS* capture new information?

In this section, we describe three experiments to differentiate the source of the *IVS*' predictive ability from that of the other significant predictors. First, we isolate the part of the equity premium that is unrelated to the *IVS* and check whether that component is predictable by the other predictor variables. If any of these variables can predict the part of the market risk premium that is orthogonal to the *IVS*, then we conclude they are capturing information that is different from what the *IVS* captures. For completeness, we reverse the process to see if the *IVS* can predict the part of the risk premium that is orthogonal to the other returns. Next, we use subsamples to determine if

---

[4] In the Appendix, we compare the in- and out-of-sample performance to several risk-based predictors that have shown promise as predictors of the market risk premium. These include the aggregate short interest index (SII) in Rapach et al. (2016), the time-varying tail risk of Kelly and Jiang (2014), the aggregate liquidity measure of Pástor and Stambaugh (2003), and the aggregate systematic risk (CATFIN) of Allen et al. (2012), which is based on bank-specific risks. The *IVS* outperforms these risk-based measures as a predictor of the market risk premium.

the *IVS* predicts the market risk premium at the same points in the business cycle as the other predictors. Finally, we compare the cross-sectional pricing ability of three linear factor models that utilize predictor variables as instruments for the conditional expected market risk premium.

3.1 Orthogonality comparison

We report the orthogonality tests of the information content in the predictors. In the first test, we use the residuals from a regression of market returns on lagged *IVS* as the part of returns that is orthogonal to *IVS*. We regress these orthogonalized returns on each of the other predictors individually to see if there is any predictable information remaining in returns after controlling for *IVS*. In the second test we reverse the process and first orthogonalize returns to the each of the other predictors and then regress those orthogonalized returns on *IVS* to see if there is any remaining information that is predictable by *IVS*. That is, in the first step, we regress:

$$r_{t+1} = \alpha_{i,h} + \beta_{i,h} x_{i,t} + \varepsilon_{i,t+1},$$ (11)

where $x_{i,t}$ represents one of the predictors. In the second step, we run the following regression:

$$\varepsilon_{i,t+1} = a_{i,t} + b_{i,t} \tilde{x}_{j,t} + u_{i,t+1},$$ (12)

where $\tilde{x}_{j,t}$ is the predictor $j \neq i$. If b_i is significant, then there is predictable information that is not being accounted for by x_i. Note that to make the magnitude of coefficients comparable among these regressions, we first standardize the level of residuals and predictors such that they are of mean zero and standard deviation one.

Table 4 contain the results from regressing returns that have been orthogonalized to the *IVS* on the other predictors. We focus on *DP*, *DY*, and *SVAR* since they are the only variables besides the *IVS* that significantly predict the market's risk premium at the 1-month frequency. Using

monthly data, we find that the dividend-to-price ratio and dividend yield are related to returns after controlling for the *IVS* while *SVAR* has no predictive ability after removing the predictable component of the risk premium related to the *IVS*. In Panel B, all the predictors contain information that is predictable by the *IVS*. These results indicate that the predictors involving dividends contain information important for predicting returns not contained in the *IVS*.

3.2 Predicting over the business cycle

Da et al. (2014) show that stock return predictability varies with the business cycle. To examine how the predictability of the *IVS* changes across business cycles, we partition our data into two sub-samples based on expansion and recession periods. Asset returns are functions of the state variables of the real economy. If these state variables are linked to economic fluctuations, then the time-varying expected returns and return predictability are expected and variables that measure and/or predict the state of the economy should display different patterns across business cycle. Recessions are identified according to the time periods in the NBER business cycle indicator, which is the dates from peak to trough, and expansions are periods outside the indicator. We have two recessions in our sample period: March 2001 to November 2001 (8 months), and December 2007 to June 2009 (18 months).

We find that in Table 5 the one-step-ahead predictive ability of *IVS* during expansionary periods is insignificant, but during recessionary periods the *IVS* is the best predictor based on (R^2_{OS}=15.65 with 6% *pval*) and utility gains (345 bps). For the conventional predictors, we find that the dividend yield (ln(*DY*)) is the strongest predictor during expansions (5.19 with 1% significance, and 321 bps utility gain), but is a poor predictor during recessions. This provides more evidence that the *IVS* is capturing different information relative to the dividend yield.

3.3 Explaining the cross-section of returns

A common use of variables shown to exhibit predictive ability is as informational variables in conditional versions of asset pricing models. In this subsection, we utilize the *IVS* as an information variable in conditional versions of linear asset pricing models. If the *IVS* contains information that is valuable for predicting the market risk premium, then utilizing this information in a conditional framework should improve the pricing ability of the model relative to the unconditional version of the model. Because the focus of our paper is predicting the equity risk premium, we consider three models that include the market as a factor: the CAPM, the three-factor model of Fama and French (FF3), and the four-factor model of Carhart. For the CAPM we estimate the parameters for the following model:

$$E_t(r_{t+1} \mid Z_t) = \alpha + \beta E_t(rm_{t+1} \mid Z_t), \tag{13}$$

where Z_t is the set of lagged information variables. For the multifactor models, we estimate:

$$E_t(r_{t+1} \mid Z_t) = \alpha + \beta E_t(rm_{t+1} \mid Z_t) + b'F, \tag{14}$$

where F is the $k \times 1$ vector of factor realizations not including the market risk premium and b is the $k \times 1$ vector of loadings. Based on our previous results and because *DY* and *DP* are highly correlated, we only include the dividend yield and *IVS* in the information set, i.e., $Z = \{\ln(DY), IVS\}$. For the unconditional versions of the models, $Z = \{\varnothing\}$.[5] We model the time variation in the conditional expected market risk premium as a linear function of the instruments. That is, we regress the market premium on the instruments as:

---

[5] We estimate the models via iterated GMM with a Newey-West spectral density using $T^{1/3} \cong 6$ lags. The estimators produce two-stage least-squares estimates of the coefficients. For the multifactor models, we do not include all the exogenous variables in the first stage estimate of the conditional expected market risk premium. We adjust the standard errors to account for the bias based on Gujarati (2003).

$$rm_{t+1} = \gamma_0 + \gamma_1'Z_t + u_{t+1}, \tag{15}$$

and then calculate the time-varying premium as $E_t\left(rm_{t+1} \mid Z_t\right) = \hat{\gamma}_0 + \hat{\gamma}_1'Z_t$.

Our test assets consist of monthly excess returns of 399 portfolios. All the portfolios are equally weighted and are available from Ken French's database. The test assets include five sets of decile portfolios sorted on past variance, residual variance, net share issues, beta, or accruals; three sets of portfolios double sorted into deciles based on size and book-to-market, size and operating profitability, or size and investment; and 49 industry portfolios. The risk-free rate, market return, SMB, and HML factors, as well as the momentum factor are all from the same source.

Table 6 contains the numbers and percentages of significant intercepts from the unconditional and conditional versions of the three models using different sets of information. For the CAPM and FF3 model, we see that modelling the time-varying market risk premium as a function of both *IVS* and the dividend yield reduces the number of significant alphas across the 399 test assets. For the CAPM, the unconditional version of the model produces 68 significant alphas indicating that the CAPM fails to price 17% of the portfolios. The conditional version of the CAPM using both the dividend yield and the *IVS* reduces the number of significant alphas to 65. Using only the dividend yield as an instrument has a large impact on the ability of the CAPM to explain the cross-section of returns as the number of significant alphas falls to 24 or 6% of the portfolios. Using only the *IVS* as a conditioning variable brings the number of significant alphas to 20 or only 5% of the portfolios, reducing the number of significant alphas by 71% relative to the unconditional version of the model.

A similar pattern emerges for the FF3 model. Here the conditional versions of the model all produce many fewer significant intercepts but utilizing the *IVS* as a sole instrument has the largest

impact. The unconditional FF3 model generates significant intercepts for 89 (22%) of the portfolios while using only the *IVS* as a conditioning variable for the market risk premium reduces that number by 85% with only 13 (3%) of the portfolios having a significant alpha. For the four factor Carhart model conditioning on *DY* alone or in conjunction with the *IVS* reduces the ability of the model to price these assets relative to the unconditional version. Using only the *IVS* to model the market risk premium reduces that number by 74% with 33 (8%) of the portfolios exhibiting a significant alpha.

Overall, when using the *IVS* to model the time-varying conditional expected market risk premium, we find a dramatic reduction in the number of significant alphas relative to unconditional versions of the models. This improvement in model performance is larger than when using the log of the dividend price ratio, which was the next best predictor of the market risk premium in our earlier experiments. We take these results as further evidence that the *IVS* contains information important for predicting returns which is different from that found in the other predictors.

4. The source of the *IVS* predictive power

In this section, we explore why the *IVS* predicts future returns on the market portfolio at longer horizons. Towards this end, we perform three experiments. First, we decompose the market return into three components related to expected returns, cash flows, and the discount rate and test which component(s) is predictable by the *IVS*. Next, we relate the *IVS* to several surveys of market participants to show market sentiment is related to the *IVS*. Finally, we test the relation between the *IVS* and several measures of aggregate market uncertainty.

4.1 Stock return decomposition

Following Campbell and Shiller (1988), we decompose the stock return into innovations of expected return, cash flow, and the discount rate, controlling for dividend-price ratio and other conventional variables in an vector autoregression (VAR) framework. Then, we analyze whether investors, who trade based on the *IVS*, can anticipate future stock returns by anticipating shocks in discount rate and/or cash flow. Taking the set of conventional predictors in Goyal and Welch (2008) as a proxy for the market information set, this analysis provides a deeper understanding of the sources of the forward-looking information in the *IVS*.

To understand the Campbell and Shiller (1988) method, we begin with the definition of the log stock return, $r_{t+1} = \log\left[\left(P_{t+1} + D_{t+1}\right)/P_t\right]$, where P_t (D_t) is the monthly stock price (dividend). The Campbell-Shiller log-linear approximation of r_{t+1} is given by:

$$r_{t+1} \approx k + \rho p_{t+1} + \left(1 - \rho\right) d_{t+1} - p_t, \tag{16}$$

where p_t (d_t) is the log stock price (divided), $\rho = \dfrac{1}{1 + \exp\left(d - p\right)}$. where $(d - p)$ is the mean of log dividend ratio, and $k = -\log\left(\rho\right) - \left(1 - \rho\right)\log\left[\left(1/\rho\right) - 1\right]$. Thus, we can rewrite equation (16) as:

$$p_t \approx k + \rho p_{t+1} + \left(1 - \rho\right) d_{t+1} - r_{t+1}. \tag{17}$$

Solving equation (17) forward and imposing the boundary condition, the stock price decomposition is given by:

$$p_t = \frac{k}{1 - \rho} + \sum_{j=0}^{\infty} \rho^j \left(1 - \rho\right) d_{t+1} - \sum_{j=0}^{\infty} \rho^j r_{t+1+j}. \tag{18}$$

Letting E_t denote the expectation operator conditional on information through month t, equations (16) and (18) imply the following decomposition for the log stock return innovation:

$$r_{t+1} - E_t r_{t+1} = \left(E_{t+1} - E_t\right) \sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j} - \left(E_{t+1} - E_t\right) \sum_{j=0}^{\infty} \rho^j r_{t+1+j}. \tag{19}$$

According to equation (19), the stock return innovation can be decomposed into cash flow news and discount rate news components:

$$\eta_{t+1}^r = \eta_{t+1}^{CF} - \eta_{t+1}^{DR}, \tag{20}$$

where:

$\eta_{t+1}^r = r_{t+1} - E_t r_{t+1}$ (expected return shock),

$\eta_{t+1}^{CF} = \left(E_{t+1} - E_t\right) \sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j}$ (cash flow shock), and

$\eta_{t+1}^{DR} = \left(E_{t+1} - E_t\right) \sum_{j=0}^{\infty} \rho^j r_{t+1+j}$ (discount rate shock).

We use a VAR framework to extract the cash flow and discount shock components. Consider the following VAR(1) model:

$$Y_{t+1} = A Y_t + U_{t+1}, \tag{21}$$

where $Y_t = (r_t, d_t - p_t, z_t)$, the N-vector z_t contains predictor variables, and A is the $(N + 2)$-by-$(N + 2)$ matrix of VAR slope coefficients, and U_t is the vector of mean-zero innovations. Denoting e_1 as $(N + 2)$ vector with one as its first element and zeros for the remaining elements, the stock return innovation and discount rate shocks can be expressed as:

$$\eta_{t+1}^r = e_1' U_{t+1}, \tag{22}$$

and

$$\eta_{t+1}^{DR} = e_1' \rho A \left(I - \rho A\right)^{-1} U_{t+1}. \tag{23}$$

The cash flow shock is then residually defined using equation (20):

$$\eta_{t+1}^{CF} = \eta_{t+1}^r + \eta_{t+1}^{DR}. \tag{24}$$

In terms of equation (20), the expected stock return for $t + 1$ based on information through t is given by:

$$E_t r_{t+1} = e_1' A Y_t. \tag{25}$$

Using $r_{t+1} = E_t r_{t+1} + \eta_{t+1}^r$ and equation (20), the log stock return can then be decomposed as:

$$r_{t+1} = E_t r_{t+1} + \eta_{t+1}^{CF} - \eta_{t+1}^{DR}. \tag{26}$$

With sample observations for Y_t for $t = 1, \ldots, T$, we can use OLS to estimate A and U_{t+1} ($t=1, \ldots, T - 1$) for the VAR model given by equation (21). Thus we can estimate all coefficients, denoted as $\hat{A}$, $\hat{U}_{t+1}$, and $\hat{\rho}$. Plugging into equations (22), (23), (24), and (25) yields $\hat{\eta}_{t+1}^r$, $\hat{\eta}_{t+1}^{DR}$, $\hat{\eta}_{t+1}^{CF}$, and $\hat{E}_t r_{t+1}$, respectively.

We analyze the source of *IVS*'s predictive power for future stock returns by examining its ability to predict the individual components comprising the total stock return. We begin with an in-sample predictive regression model for the market excess return on *IVS*:

$$r_{t+1} = \alpha + \beta IVS_t + \varepsilon_{t+1}. \tag{27}$$

We then consider the following predictive regression models for the estimation of the individual components on the RHS of equation (26):

$$\begin{aligned} r_{t+1} &= \alpha + \beta IVS_t + \varepsilon_{t+1} \\ \hat{\eta}_{t+1}^{CF} &= \beta_{CF} IVS_t + \varepsilon_{t+1}^{CF} \\ \hat{\eta}_{t+1}^{DR} &= \beta_{DR} IVS_t + \varepsilon_{t+1}^{DR} \end{aligned}. \tag{28}$$

The properties of OLS imply the following relation between the OLS estimation of β in equation (27) and those of β_E, β_{CF}, and β_{DR} in equation (28): $\hat{\beta} = \hat{\beta}_{\hat{E}} + \hat{\beta}_{CF} - \hat{\beta}_{DR}$. By comparing the estimated slope coefficients in equation (28), we can ascertain the extent to which *IVS*'s ability to predict total stock returns relates to its ability to anticipate the individual components on the RHS of equation (26).

Table 7 reports OLS estimation when the expected return shock, cash flow shock, and discount rate shock are estimated based on individual VAR constructed from market excess returns, dividend price ratio, and one of conventional predictors. We always include the dividend price ratio in the VAR to properly estimate the cash flow and discount rate shocks. The results in Table 7 show that nearly all the $\hat{\beta}_{CF}$ estimates are significant except for *NTIS* and *DFY*, indicating that innovations in cash flows contribute to the size of $\hat{\beta}_{IVS}$. Moreover, the magnitude for the coefficients of the innovation in the discount rate makes up 50% of that in the regression of the *IVS*. A weaker situation occurs for $\hat{\beta}_{DR}$. Innovations in cash flows explain most of the predictive power in the *IVS*. In contrast, the $\hat{\beta}_{E}$ coefficients are smaller and much less often significant. The intuition is the information underlying the option market influences market returns primarily through changes of cash flows in these markets.

4.2 Market sentiment and the IVS

Greenwood and Shleifer (2014) argue that investor's beliefs about future market returns are correlated with predictors such as the dividend yield. Since sophisticated traders incorporate sentiment while trading both equity and options, we test whether aggregated market sentiment is related to deviations in put-call parity that results in the predictive power captured by the *IVS*. Market sentiment is difficult to measure quantitatively.[6]Greenwood and Shleifer (2014) propose using surveys of investor's subjective forecast to the future economy. We use three survey-based return expectations as proxies of market sentiment: (1) the Gallup investor survey, (2) the American Association of Individual Investors (AAII) survey, and (3) the crash confidence index

---

[6] Studies use option open interest (Bessembinder et al., 1996) and turnover to identify investor sentiment, assuming that trading activity is driven by beliefs about market trends (Hong and Stein, 2007).

from the Yale School of Management. The time horizon of survey expectations for the Gallup survey is the next 12 months. The AAII and crash index survey reports six-month expectations. For comparison, we include a version of the statistically derived sentiment index of Baker and Wurgler (2006) from Huang et al. (2015), which is adjusted for estimation error.

The Gallup survey (1996 – 2012) asks participants whether they are "very optimistic," "optimistic," "neutral," "pessimistic," or "very pessimistic" about stock returns over the next year. The Gallup survey is widely used in the economics and finance literature given its large sample size and consistent methodology. Our measure of expectations based on Gallup survey is:

$$Gallup = Bullish\% - Bearish\%,$$

where *Bullish%* denotes the percentage of investors who are at least "optimistic" about the future stock market performance and *Bearish%* denotes those who are no more than "pessimistic." A non-zero Gallop measure indicates disparity among optimistic and pessimistic investors.

The American Association of Individual Investors Investor Sentiment Survey (1987 – 2016) measures the percentage of individual investors who are bullish, neutral, or bearish on the stock market for the next six months. The survey is administered weekly to members of the American Association of Individual Investors. The main differences between American Association of Individual Investors survey and the Gallup survey include: (1) a shorter prediction horizon, (2) fewer response levels, (3) different survey subjects, and (4) a shorter survey frequency. We construct two time series of sentiment using this survey. First, we subtract the percentage of "bearish" investors from the percentage of "bullish" investors and denote this as AAII. We also take the moving average over the past eight weeks of the Bullish index in the survey (Bull8MA). We use monthly averages of the weekly data. The former measure is similar to that of the Gallup survey in terms of the intuition, whereas the latter measure captures the momentum of the

optimistic sentiment. The higher moving average of bullish index, the stronger aggregate sentiment in better future stock market performance.

The Investor Behavior Project at Yale University led by Robert Shiller releases surveys of individual investor confidence in the stock market. We use the one-year individual Crash Confidence Index. The Crash Confidence Index is the percentage of institutional respondents who think that the probability of a market crash is strictly less than 10%. Data are available only sporadically between 1989 and July 2001. After that, the surveys are conducted monthly. The main difference between crash confidence index and the other two surveys is that we use institutional investors' responses to construct an aggregate measure of market expectations.

We also include a version of the ubiquitous market sentiment measure of Baker and Wurgler (2006) suggested by Huang et al. (2015) called the Aligned Investor Sentiment Index.[7] The Index adjustment utilizes a partial least squares method to focus the Baker and Wurgler measure on information that more closely relates to expected market returns by removing a substantial amount of approximation errors that are not relevant for forecasting returns.

If market sentiment contributes to market return prediction and is measured with little noise, then investor sentiment and predictors will have a significant positive correlation. To test this, we regress the *IVS* on the contemporaneous standardized sentiment measures from surveys, i.e.,

$$IVS_t = \alpha + \beta Sent_{i,t} + \varepsilon_{i,t}, \tag{29}$$

where *Sent* denotes the ith market sentiment measures defined above. The results in Table 8 support the prediction that the *IVS* is related to the market participant's view of future market performance. For each survey, there is a positive and significant relation between market sentiment and the *IVS*. The positive and significant relation also exists between the *IVS* and the aligned sentiment index.

---

[7] We thank Goufu Zhou for making the Aligned Investor Sentiment Index available at his webpage:
http://apps.olin.wustl.edu/faculty/zhou/PLS_BW_investor_sentiment_indexes.xls

We take these results, along with those of Table 7, as evidence that the *IVS* captures investors' perspective of future cash flows or cash flow growth.

4.4 Macroeconomic uncertainty and the IVS

The results Tables 7 and 8 focus on the connection between the longer-run predictive ability of the *IVS* and the sentiment concerning the future performance of the equity markets. We conjecture that the *IVS* also reflects reactions of option investors to innovations in the aggregate level of uncertainty that impacts firms through cash flows, investment opportunities, and earning management, as well as through monetary and fiscal policy shocks that have a significant impact on market returns. In this subsection, we explore whether the *IVS* reflects information concerning linkages between the state of the economy and financial markets (Dew-Becker et al., 2018).

For this experiment, we test the contemporaneous relation between the *IVS* and broad measures of macroeconomic and financial uncertainty from Jurado et al. (2015). *MUx* and *FUx* are macroeconomic uncertainty and financial uncertainty about next *x*-month horizon (*x*=1, 3, 12), respectively. The econometric estimates of their uncertainty measure do not rely on model structure or economic indicators because they use innovations from a factor-augmented VAR with a large set of time series. Specifically, *MUx* and *FUx* are created using 132 macroeconomic and 147 financial variables, respectively.

We also consider measures of policy uncertainty from Baker et al. (2016). Their measure is based on newspaper coverage frequency. They find that policy uncertainty is associated with greater stock price volatility and reduced investment and employment. They count key word frequency within three main sources to construct the measure: (1) 10 large newspapers, (2) reports by the Congressional Budget Office (CBO), and (3) the Federal Reserve Bank of Philadelphia's

Survey of Professional Forecasters. We apply their news measure (*News*, using Newspaper source) and baseline measure (*baseline,* using all three sources).

We estimate and test for significance of the slope parameter in the following model:

$$IVS_t = \alpha + \beta Uncertainty_{i,t} + \varepsilon_{i,t}, \tag{30}$$

where *Uncertainty* ∈ {*MU1*, *MU3*, *MU12*, *FU1*, *FU3*, *FU12*, *Baseline*, *News*}.[8] Here, *Uncertainty* and *IVS* are standardized to have a mean zero and standard deviation of one. The sample period is 1996:1 to 2017:12. The results in Table 9 indicate a consistently negative relation between the *IVS* and aggregate levels of macroeconomic, financial, and political uncertainty. We interpret these findings as further evidence that the *IVS* captures general market sentiment.

4.5 Controlling for the variance risk premium

Bollerslev et al. (2009) demonstrate the ability of the variance risk premium, i.e., the difference between the expected variance under the risk-neutral and the physical measure, to predict future market returns. Atilgan et al. (2015) find that after controlling for conditional variance, the volatility spread (OTM put vs. ATM call based on the SPX) cannot predict the market excess returns longer than a one-week horizon. Since both the *IVS* and the *VRP* derive their predictive abilities by utilizing information from the options markets, we explore whether there is any incremental predictive power in the *IVS* beyond the *VRP*.

In Table 10, we demonstrate that our *IVS* measure calculated using individual stock option implied volatilities is robust to accounting for the variance risk premium. In Panel A, we present the in-sample prediction results when we include the variance risk premium (*VRP*) in the *IVS* predictive regressions. We conduct one- to twelve-month ahead predictions and evaluate the in-

---

[8] We download the data from: https://www.sydneyludvigson.com/s/MacroFinanceUncertainty_2019Feb_update.zip and http://www.policyuncertainty.com/media/US_Policy_Uncertainty_Data.xlsx.

sample prediction results. The coefficients of *IVS* are 0.14, 0.19, 0.19, and 0.09 for one- to twelve-month horizons and are only slightly smaller than those in Table 2. The coefficients for *IVS* are significant up to three months and the coefficients on *VRP* are significant up to twelve months out.

In Panel B of Table 10, we present the out-of-sample results for a predictive regression using only *VRP*, a predictive regression using both the *VRP* and *IVS*, a combination of the forecasts of *VRP* and the *IVS* as a simple average, and the discounted mean square prediction error (*DMSPE*), which combines *IVS* and *VRP* using the weights defined in Rapach et al. (2010). Following Rapach et al. (2010), DMSPE combines the *m* forecast of *i = IVS* and *VRP* using the following weights:

$\omega_{i,t} = \varphi_{i,t}^{-1} \Big/ \sum_{j=1}^{N} \varphi_{i,t}^{-1}$, where $\varphi_{i,t} = \sum_{s=m}^{t-1} \theta^{t-1-s}(r_{s+1} - \hat{r}_{i,s+1})^2$ is the discounted squared forecast using $\theta = 0.75$ as the discount factor. *VRP* exhibits significant out-of-sample R^2 and positive utility gains over the entire sample and for expansions and recessionary periods. Including the *IVS* in the predictive regression continues to produce significant out-of-sample R^2s, but we find larger utility gains overall and in both sub-samples. For the full sample and recessionary sample, including the *IVS* nearly doubles the increase in utility. Averaging the forecasts of the *IVS* and the *VRP* using a simple average or the DMSPE method also increases the out-of-sample R^2s and utility gains.

While both the *VRP* and the *IVS* exhibit significant out-of-sample predictive ability, we are unable to deduce whether the *VRP* or the *IVS* is a better predictor of the market risk premium. We can conclude that they are complementary predictors. An investor who combines the predictability of these two variables will see large gains in utility, especially during recessionary periods.

5. Conclusion

We show that the *IVS* can robustly predict the equity risk premium over the 1996 to 2017 sample period. We first consider in-sample predictive performance using monthly data to predict 1, 3, 6,

and 12 months ahead equity risk premium. We find that in-sample results are significant under most forward periods, even after controlling for the variance risk premium. We also examine the out-of-sample prediction of the *IVS*. We find that the *IVS* is a strong predictor of the equity risk premium under various specifications, in-sample and out-of-sample, and for longer horizons than has been demonstrated in the literature. Part of the predictive ability is due to using an equally weighted version of the *IVS* suggested by the intuition of Rapach et al. (2016). We also show that the predictive ability of the *IVS* is the strongest during recessions, while the predictive ability of the dividend yield is the strongest during expansions. Utilizing the *IVS* as an information variable in conditional versions of the CAPM significantly improves the pricing of portfolios that are difficult to price unconditionally.

The source of prediction is due to forward-looking information underlying the *IVS*. We show that the predictive ability of the *IVS* is primarily related to cash flow innovations relative to innovations in expected returns or innovations in the discount rate. We also demonstrate a significant relation between the *IVS* and measures of aggregate market performance, as well as uncertainty concerning the macroeconomy and financial markets. Together the evidence of longer-term predictive ability and relation to aggregate market expectations indicate that the *IVS* captures information important for predicting the market risk premium beyond the short-term impacts arising from informed trading in the options and equity markets. The *IVS* also appears to capture general market sentiment.

References

Allen, L., Bali, T. G., & Tang, Y. (2012). Does systemic risk in the financial sector predict future economic downturns? Review of Financial Studies, 25, 3000-3036. https://doi.org/10.1093/rfs/hhs094

An, B. J., Ang, A., Bali, T. G., & Cakici, N. (2014). The joint cross section of stocks and options. Journal of Finance, 69, 2279-2337. https://doi.org/10.1111/jofi.12181

Atilgan, Y., Bali, T. G., & Demirtas, K. O. (2015). Implied volatility spreads and expected market returns. Journal of Business and Economic Statistics, 33, 87-101. https://doi.org/10.1080/07350015.2014.923776

Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. Quarterly Journal of Economics, 131, 1593-1636. https://doi.org/10.1093/qje/qjw024

Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. Journal of Finance, 61(4), 1645-1680. https://doi.org/10.1111/j.1540-6261.2006.00885.x

Bali, T. G., Cakici, N., & Chabi-Yo, F. (2015). A new approach to measuring riskiness in the equity market: Implications for the risk premium. Journal of Banking and Finance, 57, 101-117. https://doi.org/10.1016/j.jbankfin.2015.03.005.

Bali, T. G., & Hovakimian, A. (2009). Volatility spreads and expected stock returns. Management Science, 55, 1797-1812. https://doi.org/10.1287/mnsc.1090.1063.

Bekaert, G., Hoerova, M., Lo Duca, M., (2013). Risk, uncertainty, and monetary policy. Journal of Monetary Economics, 60(7), 771-788. https://doi.org/10.1016/j.jmoneco.2013.06.003.

Bessembinder, H., Chan, K., & Seguin, P. J. (1996). An empirical examination of information, differences of opinion, and trading activity. Journal of Financial Economics. 40, 105-134. https://doi.org/10.1016/0304-405X(95)00839-7.

Bliss, R. R., & Panigirtzoglou, N. (2004). Option-implied risk aversion estimates. Journal of Finance, 59, 407-446. https://doi.org/10.1111/j.1540-6261.2004.00637.x.

Bollerslev, T., Marrone, J., Xu, L., & Zhou, H. (2014). Stock Return Predictability and Variance Risk Premia: Statistical Inference and International Evidence. Journal of Financial and Quantitative Analysis, 49, 3, 633–661. https://doi.org/10.1017/S0022109014000453.

Bollerslev, T., Tauchen, G., Zhou, H., (2009). Expected stock returns and variance risk premia. Review of Financial Studies 22, 4463–4492. https://doi.org/10.1093/rfs/hhp008.

Bollerslev, T., & Todorov, V. (2011). Tails, fears, and risk premia. Journal of Finance, 66, 2165-2211. https://doi.org/10.1111/j.1540-6261.2011.01695.x.

Brown, G., & Cliff, M. (2005). Investor Sentiment and Asset Valuation Journal of Business, 78(2), 405-440. http://dx.doi.org/10.1086/427633.

Campbell, J. Y., & Thompson, S. B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average?. Review of Financial Studies, 21, 1509-1531. https://doi.org/10.1093/rfs/hhm055.

Campbell, J. Y., & Shiller, R. J. (1988). Stock prices, earnings, and expected dividends. Journal of Finance, 43, 661-676. https://doi.org/10.1111/j.1540-6261.1988.tb04598.x.

Chen, J., & Liu, Y., (2018). Bid and Ask Prices of Index Put Options: Which Predicts the Underlying Stock Returns? SSRN Working paper. https://ssrn.com/abstract=3285854.

Christoffersen, P., Jacobs, K., & Chang, B. Y. (2013). Forecasting with option-implied information. In Handbook of economic forecasting (Vol. 2, pp. 581-656). Elsevier. https://doi.org/10.1016/B978-0-444-53683-9.00010-4.

Chordia, T., A. Kurov, D. Muravyev, and A. Subrahmanyam, (2019), Index Option Trading Activity and Market Returns. Management Science, forthcoming. http://dx.doi.org/10.2139/ssrn.2798390.

Clark, T. E., & West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. Journal of Econometrics, 138, 291-311. https://doi.org/10.1016/j.jeconom.2006.05.023.

Cremers, M., & Weinbaum, D. (2010). Deviations from put-call parity and stock return predictability. Journal of Financial and Quantitative Analysis, 45, 335-367. https://doi.org/10.1017/S002210901000013X.

Da, Z., Jagannathan, R., and Shen, J. (2014). Growth Expectations, Dividend Yields, and Future Stock Returns. NBER Working Paper. 10.3386/w20651.

DeLong, J. B., A. Shleifer, L. H. Summers, and R. J. Waldmann. (1990). Noise trader risk in financial markets. Journal of Political Economy 98:703–38. https://doi.org/10.1086/261703.

Dew-Becker, I., Giglio, S., & Kelly, B. (2019). Hedging Macroeconomic and Financial Uncertainty and Volatility., Yale ICF Working Paper. http://dx.doi.org/10.2139/ssrn.3284790.

Diebold, F. X., & Mariano, R. S. (2002). Comparing predictive accuracy. Journal of Business Economics and Statistics 20, 134-144. DOI: 10.1080/07350015.1995.10524599.

Doran, J. S., Fodor, A., & Jiang, D. (2013). Call-put implied volatility spreads and option returns. Review of Asset Pricing Studies, 3, 258-290. https://doi.org/10.1093/rapstu/rat006.

Easley, D., O'Hara, M., & Srinivas, P. (1998). Option Volume and Stock Prices: Evidence on Where Informed Traders Trade. Journal of Finance, *53*(2), 431-465. https://doi.org/10.1111/0022-1082.194060.

Ferson, W. E., Sarkissian, S., & Simin, T. (2003). Spurious regressions in financial economics?. Journal of Finance. 58, 1393-1413. https://doi.org/10.1111/1540-6261.00571.

Feunou, B., Fontaine, J., Taamouti, A., & Tédongap, R., (2012). Risk Premium, Variance Premium and the Maturity Structure of Uncertainty. Review of Finance, 18 (1), 219-269. https://doi.org/10.1093/rof/rft004.

Gabaix, X. (2012). Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance. Quarterly Journal of Economics, 127, 645-700. https://doi.org/10.1093/qje/qjs001.

Greenwood, R., & Shleifer, A. (2014). Expectations of returns and expected returns. Review of Financial Studies. 27, 714-746. https://doi.org/10.1093/rfs/hht082.

Goyal, A. & Welch, I. (2008). A comprehensive look at the empirical performance of equity premium prediction. Review Financial Studies 21, 1455-1508. https://doi.org/10.1093/rfs/hhm014.

Gujarati, D.N. (2003) Basic Econometrics. 4th Edition, McGraw-Hill, New York.

Hollstein, F., Prokopczuk, M., Tharann, B., & Wese Simen, C., (2019). Predicting the equity market with option-implied variables. European Journal of Finance, 25, 10, 937-965. https://doi.org/10.1080/1351847X.2018.1556176.

Hong, H., & Stein, J. C. (2007). Disagreement and the stock market. Journal of Economic Perspectives, 21, 109-128. DOI: 10.1257/jep.21.2.109.

Huang, D., Jiang, F., Tu, J., & Zhou, G. (2015). Investor sentiment aligned: A powerful predictor of stock returns. Review of Financial Studies, 28(3), 791-837. https://doi.org/10.1093/rfs/hhu080.

Jurado, K., Ludvigson, S. C., & Ng, S. (2015). Measuring uncertainty. American Economic Review, 105, 1177-1216. DOI: 10.1257/aer.20131193.

Kelly, B., & Jiang, H. (2014). Tail risk and asset prices. Review of Financial Studies. 27, 2841-2871. https://doi.org/10.1093/rfs/hhu039.

Ludvigson, S. C., & Ng, S. (2015). The empirical risk–return relation: A factor analysis approach. Journal of Financial Economics, 83, 171-222. https://doi.org/10.1016/j.jfineco.2005.12.002.

Newey, W. K., & West, K. D. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. Econometrica, 55, 703-708. DOI: 10.2307/1913610.

Nelson, C. & Kim, M. (1993). Predictable Stock Returns: The Role of Finite sample Bias. Journal of Finance, 48, 641-661. https://doi.org/10.1111/j.1540-6261.1993.tb04731.x.

Pástor, Ľ., & Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. Journal of Political Economy, 111, (3), 642-685. https://doi.org/10.1086/374184.

Rapach, D. E., Ringgenberg, M. C., & Zhou, G. (2016). Short interest and aggregate stock returns. Journal of Financial Economics, 121, 46-65. https://doi.org/10.1016/j.jfineco.2016.03.004.

Rapach, D. E., Strauss, J. K., & Zhou, G. (2010). Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. Review of Financial Studies. 23, 821-862. https://doi.org/10.1093/rfs/hhp063.

So, R., Driouchi, T. & Trigeorgis, L., (2016). Option Market Ambiguity, Excess Returns and the Equity Premium. SSRN Working paper. https://ssrn.com/abstract=2663317.

Stambaugh, R. F. (1999). Predictive regressions. Journal of Financial Economics, 54, 375-421. https://doi.org/10.1016/S0304-405X(99)00041-0.

West, K. D. (1996). Asymptotic inference about predictive ability. Econometrica, 64, 1067-1084. DOI: 10.2307/2171956.

Xing, Y., Zhang, X., & Zhao, R. (2010). What does the individual option volatility smirk tell us about future equity returns?. Journal of Financial and Quantitative Analysis, 45(3), 641-662. https://doi.org/10.1017/S0022109010000220.

Zhou, H., (2018). Variance risk premia, asset predictability puzzles, and macroeconomic uncertainty. Annual Review of Financial Economics, 10, 481-497. https://doi.org/10.1146/annurev-financial-110217-022737.

Table 1

Summary statistics. This table displays summary statistics of market returns, implied volatility spread (*IVS*), and 14 predictors from Goyal and Welch (2008), where AR(1) is the 1st order autocorrelation, respectively. Market excess returns are the logarithm of the return on the value-weighted CRSP index minus logarithm of the prevailing short-term interest rate. *IVS* is the equal-weighted average of call and put option implied volatility spread across individual stocks. ln(*DP*) is the log dividend-price ratio, ln(*DY*) is the log dividend yield, ln(*EP*) is the log earnings-price ratio, ln(*DE*) is the log dividend-payout ratio, *SVAR* is sum of squared daily returns on the S&P 500, *BM* is the book-to-market ratio for the Dow Jones Industrial Average, *NTIS* is net equity expansion, *TBL* is the interest rate on a three-month Treasury bill, *LTY* is the long-term government bond yield, *LTR* is the return on long-term government bonds, *TMS* is the long-term government bond yield minus the Treasury bill rate, *DFY* is the difference between Moody's BAA- and AAA-rated corporate bond yields, *DFR* is the long-term corporate bond return minus the long-term government bond return, and *INFL* is inflation calculated from the *CPI* for all urban consumers. The sample contains 264 observations for the monthly sample for the 1996:1 to 2017:12 period.

	Mean	Std. Dev.	AR(1)
Market excess returns	0.005	0.043	-0.043
IVS	-0.008	0.005	0.372
ln(*DP*)	-4.016	0.209	0.949
ln(*DY*)	-4.010	0.209	0.949
ln(*EP*)	-3.156	0.381	0.929
ln(*DE*)	-0.860	0.436	0.940
SVAR	0.003	0.005	0.438
BM	0.268	0.073	0.926
NTIS	0.002	0.019	0.936
TBL	0.022	0.021	0.983
LTY	0.045	0.014	0.958
LTR	0.006	0.030	-0.182
TMS	0.023	0.013	0.939
DFY	0.010	0.004	0.888
DFR	0.001	0.018	-0.087
INFL	0.002	0.004	0.038

Table 2

In-sample predictability. This table reports the estimation results of in-sample predictive regressions. The prediction model is:

$$r_{t+h} = \alpha_i + \beta_i X_{i,t} + \varepsilon_{i,t+h},$$

where r_{t+h} is h-month-ahead market excess returns, including 1-month ahead predictions and overlapping 3, 6, and 12-month predictions, and X_i represents each predictor in the first column. See the notes of Table 1 for variable definitions. The columns labeled t_{NW} contain t-statistics using Newey-West heteroscedasticity-autocorrelation consistent standard errors with lags chosen as in Ferson et al. (2003). Market excess returns and predictors are standardized to have a mean zero and standard deviation of one. The sample period is 1996:1 to 2017:12.

Predictor	$h = 1$ β	t_{NW}	$h = 3$ β	t_{NW}	$h = 6$ β	t_{NW}	$h = 12$ β	t_{NW}
IVS	0.19	3.02	0.26	4.33	0.25	4.09	0.13	2.03
ln(DP)	0.13	2.10	0.20	3.35	0.31	5.21	0.43	7.48
ln(DY)	0.14	2.37	0.22	3.54	0.32	5.39	0.44	7.80
ln(EP)	0.06	0.93	0.04	0.58	0.02	0.34	0.08	1.19
ln(DE)	0.01	0.19	0.07	1.07	0.13	2.10	0.14	2.22
SVAR	-0.16	-2.57	-0.15	-2.51	-0.03	-0.53	0.05	0.74
BM	0.07	1.19	0.13	2.05	0.24	4.03	0.31	5.22
NTIS	0.10	1.59	0.21	3.42	0.27	4.52	0.30	4.95
TBL	-0.05	-0.76	-0.02	-0.28	-0.05	-0.80	-0.11	-1.73
LTY	-0.08	-1.26	-0.06	-0.98	-0.08	-1.28	-0.06	-1.02
LTR	0.03	0.53	-0.02	-0.36	0.03	0.55	0.02	0.32
TMS	-0.01	-0.14	-0.04	-0.63	0.00	-0.08	0.11	1.76
DFY	-0.08	-1.23	-0.08	-1.24	0.01	0.14	0.08	1.24
DFR	0.08	1.23	0.06	0.91	0.06	0.90	0.06	1.01
INFL	0.09	1.50	-0.05	-0.79	-0.13	-2.15	-0.14	-2.16

Table 3

Out-of-sample predictability. This table reports estimation results of out-of-sample predictive regression for next 1-, 3-, 6-, and 12-month (cumulative) market excess returns. Out-of-sample R^2 (R_{OS}^2 (%)) is defined as one minus the ratio of mean square forecast error (MSFE) of predictive regression versus MSFE of historical average up to the current period. p represents the p-value associated with Clark-West MSFE-adjusted statistic, testing the null hypothesis, H_0: $R_{OS}^2 \leq 0$, and the alternative H_1: $R_{OS}^2 > 0$; that is, the historical average MSFE is less than or equal to the predictive regression MSFE. Δ (%) is the annualized utility gain, the difference between predicted utility and historical average utility, calculated from the CARA utility of holding mean-variance portfolios. See the notes of Table 1 for variable definitions and IVS (VW) is the value-weighted version of the IVS where the weights are based on market capitalization. SPX is the IVS constructed from the index options (SPX). The sample period is 1996:1 to 2017:12.

	$h = 1$			$h = 3$			$h = 6$			$h = 12$		
	R_{OS}^2	pval	Δ	R_{OS}^2	pval	Δ	R_{OS}^2	pval	Δ	R_{OS}^2	pval	Δ
IVS	3.23	0.08	1.46	3.18	0.04	3.10	2.58	0.01	0.44	-1.41	0.12	-2.18
ln(DP)	-5.01	0.83	-6.18	-9.30	0.82	-13.83	0.93	0.03	-45.81	-2.26	0.15	-62.34
ln(DY)	-3.10	0.65	-3.76	-6.94	0.67	-10.92	3.19	0.01	-43.34	0.43	0.12	-54.19
ln(EP)	-8.04	0.65	0.87	-15.90	0.82	0.40	-7.64	0.08	-13.23	-15.32	0.24	-36.06
ln(DE)	-6.93	0.43	-0.28	-8.35	0.33	0.80	-15.40	0.51	-10.17	-28.82	0.71	-26.48
SVAR	-3.04	0.27	-7.35	0.22	0.23	-2.54	-4.86	0.27	-1.47	-5.36	0.27	0.27
BM	-1.10	0.78	-0.94	-0.49	0.28	-0.02	-2.92	0.01	-13.00	-12.00	0.02	-15.90
NTIS	-1.45	0.26	-2.87	-0.04	0.07	-1.82	-4.48	0.00	1.30	-5.17	0.08	7.68
TBL	-0.92	0.59	-0.49	-1.72	0.50	-0.18	-10.98	0.61	-11.62	-16.35	0.64	-13.65
LTY	-0.58	0.43	-0.03	-1.17	0.41	0.38	-1.98	0.11	3.97	-4.95	0.28	3.13
LTR	-1.53	0.96	-1.13	-1.45	0.78	-1.34	-1.07	0.79	-0.06	-1.08	0.94	-0.27
TMS	-0.61	0.72	-0.76	-0.67	0.55	0.22	-10.90	0.90	-12.24	-19.00	0.79	-16.47
DFY	-4.50	0.22	-0.70	-10.99	0.09	-1.88	-11.41	0.00	1.16	-11.45	0.11	4.52
DFR	-5.21	0.79	-4.79	-5.64	0.90	-6.13	-3.56	0.65	-7.90	-2.60	0.74	-1.18
INFL	-0.26	0.34	-2.17	-1.31	0.67	-1.95	-0.04	0.21	-2.81	-0.84	0.31	-5.22
IVS (VW)	-1.08	0.79	-0.91	-2.26	0.97	-2.33	-0.26	0.32	-1.28	-0.57	0.47	-1.96
SPX	-0.90	0.54	-1.21	-0.45	0.59	-0.42	-4.48	0.97	-6.29	-6.77	0.92	-4.97

Table 4

Information content of *IVS*. The table presents the predictive power for all predictors, using an "orthogonal two-step" method. The first step is to use the predictor *IVS* to predict the equity premium, i.e.,

$$r_{t+1} = \alpha + \beta x_t + \varepsilon_{t+1}.$$

Using ε_{t+1} as the part of the equity premium not explained by *x,* the second step is to regress the residual on other predictors, i.e.,

$$\varepsilon_{t+1} = \delta + \gamma \tilde{x}_t + \omega_{t+1}.$$

In Panel A, we report the orthogonality regression of the call-put option implied volatility spread residual on other predictors, and Panel B Column we report a test of whether *IVS* can explain other predictor's residuals. The sample period is 1996:1 to 2017:12.

Panel A: Other predictor on *IVS*'s residual			
Second Step: $\varepsilon_{t+1}^{IVS} = \delta + \gamma X_t + \omega_{t+1}$	γ	t_{NW}	R^2 (%)
ln(DP)	0.17	2.80	2.91
ln(DY)	0.19	3.17	3.70
SVAR	-0.09	-1.46	0.81
Panel B: *IVS* on other predictor's residual			
Second Step: $\varepsilon_{t+1}^{X} = \delta + \gamma IVS_t + \omega_{t+1}$	γ	t_{NW}	R^2 (%)
ln(DP)	0.22	1.99	4.53
ln(DY)	0.23	2.10	4.88
SVAR	0.13	1.86	1.61

Table 5

Expansions and recessions. We use the business cycles defined by the NBER to separate the sample into expansion and recession periods. See the notes to Table 1 for the variable definitions. Out-of-sample R^2 is defined as one minus the ratio of mean square forecast error (MSFE) of prediction versus MSFE of historical average up to the current period. Clark-West p-values are reported The Clark-West MSFE-adjusted statistic tests the null hypothesis, H_0: $R_{OS}^2 \leq 0$, and the alternative H_1: $R_{OS}^2 > 0$; that is, the historical average MSFE is less than or equal to the predictive regression MSFE. The sample period is 1996:1 to 2017:12.

Predictor	Expansion			Recession		
	R_{OS}^2	*pval*	Δ (%)	R_{OS}^2	*pval*	Δ (%)
IVS	-6.17	0.48	1.15	15.65	0.06	3.45
ln(DP)	4.33	0.00	2.70	-17.37	0.99	-61.25
ln(DY)	5.19	0.00	3.21	-14.06	0.99	-47.34
SVAR	2.61	0.07	0.42	-10.52	0.35	-58.76

Table 6
Conditioning on *IVS* and dividend yield. This table contains the number and percentage of significant intercepts from regressing the 399 monthly excess portfolio returns described in the text on excess market returns for conditional and unconditional versions of the CAPM, Fama and French three-factor model (FF3), and the Carhart four-factor model (Carhart). All the portfolios are equally weighted. The factors for three models and the portfolio data are all available from Ken French's web page. The row labeled Unconditional contains number and percentage of the sample of alphas significantly different from zero using Newey-West (1987) standard errors from the unconditional versions of the models. The rows labeled $E[R_M|\{DY,IVS\}]$ contain the results for conditional versions of the where the market risk premium is a linear function of the instruments both *DY* and *IVS*. The data are for the period 1996:1 to 2016:12.

Model	Number of |t-stat|≥ 1.96	Percentage of Sample
CAPM Unconditional	68	0.17
CAPM with $E[R_M|\{DY,IVS\}]$	65	0.16
CAPM with $E[R_M|DY]$	24	0.06
CAPM with $E[R_M|IVS]$	20	0.05
FF3 Unconditional	89	0.22
FF3 with $E[R_M|\{DY,IVS\}$	35	0.09
FF3 with $E[R_M|DY]$	34	0.08
FF3 with $E[R_M|IVS]$	13	0.03
Carhart Unconditional	125	0.31
Carhart with $E[R_M|\{DY,IVS\}]$	168	0.42
Carhart with $E[R_M|DY]$	198	0.50
Carhart with $E[R_M|IVS]$	33	0.08

Table 7

Decomposing market returns. This table contains the sources of predictive power in call-put implied volatility spread. Based on Campbell and Shiller (1988), we first regress the equity premium on the logarithm of dividend-price ratio and other conventional predictors. The innovation is the difference between raw discount data and expected discount variables. Lastly, we regress three innovations on the call-put implied volatility spread respectively. See the notes in Table 1 for the variable definitions. The sample period is 1996:1 to 2017:12.

Predictor	Expected Return Channel		Cash Flow Channel		Discounted Rate Channel	
	$\hat{\beta}_E$	pval	$\hat{\beta}_{CF}$	pval	$\hat{\beta}_{DR}$	pval
ln(DP)	-0.04	0.04	0.12	0.01	-0.12	0.06
ln(DP),SVAR	0.04	0.38	0.10	0.01	-0.06	0.22
ln(DP), BM	-0.04	0.04	0.10	0.01	-0.13	0.04
ln(DP), NTIS	0.00	0.96	0.03	0.23	-0.15	0.03
ln(DP), TBL	-0.04	0.07	0.11	0.02	-0.12	0.04
ln(DP), LTY	-0.04	0.04	0.12	0.03	-0.11	0.06
ln(DP), LTR	-0.04	0.07	0.12	0.01	-0.11	0.07
ln(DP), TMS	-0.03	0.15	0.12	0.01	-0.10	0.12
ln(DP), DFY	0.03	0.20	0.08	0.32	-0.08	0.04
ln(DP), DFR	-0.05	0.02	0.12	0.01	-0.12	0.04
ln(DP), INFL	-0.04	0.09	0.12	0.02	-0.11	0.05

Table 8

Sentiment and the implied volatility spread. This table reports the results of regressions of the implied volatility spread (*IVS*) on the measures of sentiment calculated from survey-based data. The Gallup Spread series spans 1996:10 to 2011:12, with a notable gap between 2009:11 and 2011:2. AAII Bull8MA and the AAII Spread is for the period1996:1 to 2016:12. Crash confidence is available between 1989:10 and 2016:12. BW AIS is the Baker and Wurgler (2006) sentiment index aligned for forecasting by Huang et al. (2015) from 1996:1 to 2016:12. We regress each sentiment measure on the contemporaneous *IVS*. The represented results are coefficients estimate of β and three statistics. t_{NW} is calculated using Newey-West heteroscedasticity-autocorrelation consistent standard errors with twelve-month lags, t_{HH} corrects overlapping observation bias with twelve-month lags, and R^2 (%) associates with OLS regression.

Sentiment Measure	β	t_{NW}	t_{HH}	R^2 (%)
Gallup Spread	0.45	5.20	6.07	20.40
AAII Bull8MA	0.15	2.69	2.26	2.15
AAII Spread	0.23	2.04	2.20	12.45
Crash	0.35	2.76	2.25	11.98
BW AIS	0.29	3.32	2.93	8.45

Table 9

***IVS* and macroeconomic, financial, and political uncertainty.** This table contains test results of the relationship between uncertainty measures and the implied volatility spread (*IVS*). *MUx* and *FUx* are macroeconomic uncertainty and financial uncertainty about next *x*-month horizon (*x*=1, 3, 12), respectively (Jurado et al., 2015). Political uncertainty about the next one month includes *Baseline* and *News*, economic policy uncertainty (*EPU*) measures based on Baker et al. (2016). The results are coefficient estimates of *β* and corresponding *t*-statistics and R^2. t_{NW} is calculated using Newey-West (1987) heteroscedasticity-autocorrelation consistent standard errors with twelve-month lags modified by *Ferson et al. (*2003). The prediction model is:

$$IVS_t = \alpha + \beta\ Uncertainty_{it} + \varepsilon_{i,t},$$

where *Uncertainty$_{it}$* is the *i^{th}* uncertainty measure, and *Uncertainty* $\in$ *{MU1, MU3, MU12, FU1, FU3, FU12, Baseline, News}*. Uncertainty and *IVS* are standardized to have a mean zero and standard deviation of one. The sample period is 1996:1 to 2017:12.

Uncertainty	β	t_{NW}	R^2 (%)
MU1	-0.46	-2.25	21.42
MU3	-0.46	-2.19	20.90
MU12	-0.44	-2.16	19.38
FU1	-0.39	-2.56	14.94
FU3	-0.38	-2.60	14.48
FU12	-0.36	-2.66	12.91
Basline (*h*=1)	-0.28	-3.45	7.70
News (*h*=1)	-0.22	-2.71	4.97

Table 10

Controlling for variance risk premium. Panel A reports estimation results of in-sample predictive regressions including the *IVS* and the *VRP*. The prediction model is:

$$r_{t+1} = \alpha_i + \beta_i\,IVS_{i,t} + \gamma_i VRP_t + \varepsilon_{i,t+1}.$$

where r_{t+1} is one-month ahead market excess returns, and *VRP* is the variance risk premium. Panel B contains out-of-sample results for a predictive regression using only *VRP*, a predictive regression using both the *VRP* and *IVS*, the forecast of combining the forecasts of *VRP* and the *IVS* as a simple average and the discounted mean square prediction error (DMSPE) which combines *IVS* and *VRP* using the weights defined in Rapach et al. (2010). Here t_{NW} is calculated using Newey-West heteroscedasticity-autocorrelation consistent standard errors with 12-month lags modified by Ferson et al. (2003). R^2 is the percentage of adjusted R^2. Market excess returns and predictors are standardized to have a mean zero and standard deviation of one. Out-of-sample R^2 is defined above. Clark-West *p*-values are reported The Clark-West MSFE-adjusted statistic tests the null hypothesis, H_0: $R_{OS}^2 \leq 0$, and the alternative H_1: $R_{OS}^2 > 0$; that is, the historical average MSFE is less than or equal to the predictive regression MSFE. The sample period is 1996:1 to 2017:12.

Panel A: In-sample Prediction with *VRP*					
	β	t_{NW}	β_{VRP}	t_{NW}	R^2
One-month	0.14	2.23	0.21	3.51	7.76
Three-month	0.19	2.01	0.30	4.39	15.17
Six-month	0.19	1.68	0.25	2.94	11.85
Twelve-month	0.09	0.86	0.15	1.72	3.85

Panel B: Out-of-sample Prediction with *VRP*								
	Overall			Expansion			Recession		
	R_{OS}^2	*pval*	Δ(%)	R_{OS}^2	*pval*	Δ(%)	R_{OS}^2	*pval*	Δ (%)
VRP	5.13	0.02	2.29	3.94	0.03	1.14	7.08	0.10	8.65
VRP + IVS	4.57	0.02	4.44	2.19	0.05	1.82	8.51	0.06	19.89
Simple average	6.59	0.02	4.31	3.06	0.04	1.81	12.44	0.05	19.04
DMSPE	6.72	0.02	4.82	3.07	0.04	2.04	12.76	0.05	21.30

Online Appendix

Robustness

We perform several robustness checks of our results. First, we explore several different ways of calculating the *IVS* to ensure our results are not due to any biases. We then utilize a robust testing methodology to confirm our results. Finally, we demonstrate that our findings are robust to controls and compare our results with other predictors not in the Goyal and Welch (2008) data set.

Alternative Definitions of the IVS

Our *IVS* measure is calculated using option implied volatilities for at-the-money options with 30 days-to-expiration. If the option is near expiration, there is limited forward-looking information to pin down forecasts market excess returns. On the other hand, options with 30 days until expiration are the most actively traded. In Table A1, we present the in- and out-of-sample results for versions of the *IVS* calculated using the implied volatility for options with 91, 182, 273, and 365 days to expiration. For in-sample prediction, the *IVS* created using longer time to expiration options continue to be significant positive predictors of the equity premium but the significance declines with time to expiration. For the out-of-sample experiment, we see that these versions of the *IVS* fail to produce any utility gains beyond using the average equity premium and out-of-sample R^2's are all negative except for the *IVS* using options with 91 days to expiration. We conclude that the higher trading activity of the options with 30 days to expiration is important for predicting the equity risk premium.

Our primary *IVS* measure is calculated using monthly data with implied volatilities being estimated using the binominal tree model with historical dividend payments. The use of historical dividend payments may cause a look-ahead bias in our estimation of the *IVS*. To

account for the potential bias in the *IVS* we first compare our monthly version of *IVS* to a version

calculated using only the last five days of implied volatility in a given month. Using the last

week of the month reduces the probability of our sample containing a stock that paid a dividend.

In Table A1 of the appendix, we show that *IVS* created using the last five days of the month

performs on par with our monthly measure both in and out-of-sample.

Table A1: Different implied volatility spread definitions

The table reports the predictive ability of the *IVS* constructed from monthly equal-weighted at-the-money options with more than 30 days-to-expiration, as well as *IVS* from the last five days in a month. The rows labeled DTEx report results for the *IVS* formed using x=91,182,273, and 365 days to expiration. The row labeled SPX contains the results for the *IVS* constructed from the index options (SPX). Panel A reports the in-sample prediction for one to twelve-month horizons. Panel B is the out-of-sample results with multiple horizons. See Table 2 and 3 for detailed methodologies and statistics definitions. The sample period is 1996:1 to 2017:12

	Panel A: In-sample Results											
	$h=1$			$h=3$			$h=6$			$h=12$		
	β	t_{NW}	R^2 (%)	β	t_{NW}	R^2 (%)	β	t_{NW}	R^2 (%)	β	t_{NW}	R^2 (%)
DTE 91	0.16	2.63	2.58	0.24	4.04	5.92	0.24	4.02	5.94	0.15	2.40	2.26
DTE 182	0.14	2.20	1.82	0.20	3.36	4.17	0.19	3.18	3.79	0.11	1.81	1.29
DTE 273	0.13	2.06	1.60	0.19	3.16	3.71	0.19	3.02	3.43	0.11	1.79	1.26
DTE 365	0.12	1.92	1.40	0.18	3.00	3.35	0.18	2.97	3.34	0.12	1.90	1.42
Last 5 days	0.20	3.28	3.96	0.33	5.54	10.60	0.24	3.88	5.56	0.13	2.14	1.79

	Panel B: Out-of-sample Results											
	$h=1$			$h=3$			$h=6$			$h=12$		
	R^2_{OS}	pval	Δ (%)	R^2_{OS}	pval	Δ (%)	R^2_{OS}	pval	Δ (%)	R^2_{OS}	pval	Δ (%)
DTE 91	0.72	0.16	-1.50	0.01	0.07	-0.02	-0.37	0.03	1.18	-6.10	0.15	-1.87
DTE 182	-0.59	0.23	-0.69	-1.48	0.14	-1.27	-2.70	0.11	-4.52	-7.78	0.34	-5.26
DTE 273	-0.85	0.26	-0.63	-1.86	0.17	-1.42	-2.71	0.15	-4.93	-7.57	0.39	-5.99
DTE 365	-1.19	0.29	-1.86	-2.86	0.21	-1.57	-2.97	0.17	-4.71	-7.14	0.40	-5.82
Last 5 days	-2.67	0.16	-0.05	-9.60	0.18	-1.17	-6.73	0.04	-5.87	-6.23	0.21	-21.37

We more thoroughly address the possible look-ahead bias caused from using realized

dividends in Table A2. Here we delete all options with stocks having dividend payments

(identified by declaration dates) before the expiration of the option. We then calculate the

implied volatility spread to predict market excess returns using the sample of firms that did not

pay a dividend. The results in Table A2 of the appendix indicate that eliminating options on

dividend paying stock only marginally impacts the distribution of the *IVS* and its predictive

ability remains significant with similar economic value in- and out-of-sample and across

business cycles. We conclude that our previous results are not an artifact of a look-ahead bias in

how we calculate the *IVS*.

Table A2: *IVS* with and without dividend

The table investigates the effect of look-ahead bias inherited from dividend. We delete daily *IVS* within the 30-
day window before the dividend declaration date corresponding to 30 day-of-expiration options. *IVS*: No look-
ahead bias is the measure without dividend look ahead bias. Panel A summarizes the descriptive statistics. Panel
B reports in-sample and out-of-sample prediction results. Column (1) and (2) are in-sample prediction with
Newey-West heteroscedasticity-autocorrelation consistent standard errors with 12-month lags modified by Ferson
et al. (2003) *t* statistics in brackets. Column (3) and (4) are out-of-sample prediction results with p-value in the
parentheses. Panel C further separates the sample into expansion and recession. See Table 2 and 3 for detailed
methodologies and statistics definitions. The sample period is 1996:1 to 2017:12.

Panel A: Summary Statistics				
	Mean	Std. Dev.	Skew	Kurt
IVS	-0.008	0.005	-0.786	5.145
IVS: No look-ahead bias	-0.009	0.005	-0.740	4.906
Panel B: In-sample and Out-of-sample				
	β	R^2 (%)	R_{OS}^2(%)	Δ (%)
IVS	0.80	3.00	3.21	6.05
	[2.66]		(0.08)	
IVS: No look-ahead bias	0.79	2.80	2.85	5.88
	[2.54]		(0.09)	
Panel C: Out-of-sample Across Business Cycles				
	Expansion		Recession	
	R_{OS}^2(%)	Δ (%)	R_{OS}^2(%)	Δ (%)
IVS	-6.34	0.98	15.58	34.54
	(0.51)		(0.06)	
IVS: No look-ahead bias	-7.01	0.51	15.62	36.11
	(0.60)		(0.06)	

Inference Methodology

Since we use overlapping market excess returns for the longer-run forecast exercises, we utilize a

robust Wald test that addresses persistence in predictors as in Kostakis, Magdalinos, and

Stamatogiannis (2014). We apply the IVX-Wald statistic to the in-sample estimates in Equation

(1). In Table A3, the IVX-Wald statistics of the implied volatility spread are significant at all

horizons, while among Goyal and Welch (2008) predictors, only ln(DP), SVAR, and NTIS are

significant. These results indicate that the predictive power of *IVS* is not a statistical artifact of persistence caused by using overlapping data.

Table A3: In-sample prediction using robustness inference

The table reports in-sample prediction results over one-, three-, six- and twelve-month ahead market excess returns. We use robustness inference Wald test based on Kostakis *et al.* (2014) and present the p-value associated with the Wald statistics (Column (2), (4), (6), and (8)). Column (1)(3)(5)(7) are in-sample prediction coefficients. See the notes in Table 1 for detailed variable definitions. The sample period is 1996:1 to 2017:12.

	h = 1		h = 3		h = 6		h = 12	
	β	p-IVX	β	p-IVX	β	p-IVX	β	p-IVX
IVS	0.19	0.00	0.26	0.00	0.25	0.00	0.13	0.23
ln(DP)	0.13	0.14	0.20	0.11	0.31	0.07	0.43	0.04
ln(DY)	0.14	0.09	0.22	0.09	0.32	0.06	0.44	0.04
ln(EP)	0.06	0.36	0.04	0.62	0.02	0.74	0.08	0.59
ln(EDE)	0.01	1.00	0.07	0.67	0.13	0.48	0.14	0.48
SVAR	-0.16	0.01	-0.15	0.04	-0.03	0.55	0.05	0.82
BM	0.07	0.42	0.13	0.17	0.24	0.04	0.31	0.02
NTIS	0.10	0.04	0.21	0.02	0.27	0.01	0.30	0.03
TBL	-0.05	0.76	-0.02	0.68	-0.05	0.50	-0.11	0.24
LTY	-0.08	0.29	-0.06	0.30	-0.08	0.28	-0.06	0.35
LTR	0.03	0.61	-0.02	0.54	0.03	0.55	0.02	0.79
TMS	-0.01	0.59	-0.04	0.74	0.00	0.99	0.11	0.38
DFY	-0.08	0.11	-0.08	0.30	0.01	0.81	0.08	0.73
DFR	0.08	0.22	0.06	0.21	0.06	0.16	0.06	0.12
INFL	0.09	0.13	-0.05	0.48	-0.13	0.04	-0.14	0.01

Comparing the IVS with Risk-based Predictors

Goyal and Welch (2008) investigate a set of conventional predictors which are based on either firms' fundamental value or economic variables. They conclude that these variables rarely predict market excess returns. Recently, several risk-based predictors that have shown promise as predictors of the market risk premium. These include:

1) The aggregate short interest index (SII) in Rapach, Ringgenberg, and Zhou, (2016).

2) Time-varying tail risk of Kelly and Jiang (2014).[9]

---

[9] Following Kelly and Jiang (2014), we define tail risk conditional on exceeding some extreme lower "tail threshold," u_t, and given information up to current time t, we assume that an asset's return obeys the tail probability distribution. The common time-varying component of return tails, λ_t, is estimated month-by-month by applying a power law estimator to the set of daily return observations for all stocks in month t. Applied to the pooled cross-section each month, it takes the form

3) Aggregate liquidity measure of Pástor and Stambaugh (2003).[10]

4) And aggregate systematic risk (CATFIN) of Allen, Bali, and Tang (2012) which is based on bank specific risks.[11]

In Table A4, for the monthly in-sample prediction, we find that the coefficient for SII is negative and insignificant when we use either a value or equally weighted version and it remains insignificant in the multivariable model that includes SSI and *IVS*. For the one-month ahead prediction, SSI fails to produce better prediction than the mean of the market risk premium but does indicate some utility gain.

Table A4: Short-selling constraints

The table compares the prediction results of short-selling interest indicator with *IVS*. Short-selling interest indicator is defined as equal-weighted (*SII EW*) and value-weighted (*SII VW*) monthly short interest among individual firms (Rapach *et al.*, 2016). Panel A reports one-month ahead univariate and multivariate in-sample prediction of *IVS* and SII, where the first two entries are OLS estimated coefficients, and we present t_{NW} in the brackets calculated using Newey-West heteroscedasticity-autocorrelation consistent standard errors with 12-month lags modified by Ferson *et al.* (2003). Panel B reports out-of-sample one-month predictability of *IVS* and SII. See Table 2 and 3 for detailed methodology and variable definitions. The sample is 1996:1 to 2014:12.

	Panel A: In-sample			Panel B: Out-of-sample		
	Univariate		Multivariable	R_{OS}^2 (%)	p	Δ (%)
IVS	0.93		0.89	5.20	0.07	8.38
	[3.37]		[4.01]			
SII (EW)		-0.38	-0.08	-0.15	0.40	3.22
		[-0.81]	[-0.36]			
SII (VW)			-0.09	-2.56	0.64	1.95
			[-0.18]			
R^2(%)	3.80	0.30	-0.40	3.50		

In Table A5, we present the one-, three-, six-, and twelve-month in-sample and out-of-sample predictions for tail risk, liquidity risk and aggregate systematic risk. In sample, none of the risk-

$$\lambda_t^{Hill} = \frac{1}{K_t} \sum_{k=1}^{K_t} log\left(\frac{R_{k,t}}{u_t}\right)$$

where $R_{k,t}$ is the kth daily return that falls below an extreme value threshold u_t during month t, and K_t is the total number of such exceedances within month t. We take the fifth percentile of the cross-section each period as u_t.

[10] http://finance.wharton.upenn.edu/~stambaug/liq_data_1962_2018.txt
[11] http://faculty.msb.edu/tgb27/CATFIN_2017.xlsx

based measures perform well based on Newey-West t-statistics or the IVX p-values except for tail-risk at the 3-month horizon. For out-of-sample prediction none of the risk-based measures show promise as a predictor of the market risk premium given consistently negative out-of-sample R^2's.

Table A5: *IVS* vs Risk-based predictors

This table reports the estimation results of in-sample and out-of-sample evaluation of predictive regressions. In Panel A, the prediction model is

$$r_{t+h} = \alpha_i + \beta_i X_{i,t} + \varepsilon_{i,t+h}$$

where r_{t+h} is h-horizon ahead market excess returns, and X_i represents each predictor in the first column. *Tail Risk* is defined as Kelly and Jiang (2014) using cross sectional returns and exogenously select returns over fifth percentile of the cross-section each period as the "tail threshold", from 1996 to 2017. *CATFIN* is bank-specifically systematic risks measured using the average of three *VaR* measures, from Bali's website. *LIQ* is Pastor-Stambaugh aggregate liquidity innovation measure from Stambaugh website. See the notes of Table 1 for variable definitions. Column (2) (4) (6) (8) report the β estimates. t_{NW} in the brackets is calculated using Newey-West standard errors with 12-month lags modified by Ferson *et al.* (2003). IVX-Wald test p value in the parentheses is calculated based on Kostakis *et al.* (2014). R^2 in Column (3) (5) (7) (9) are the adjusted R^2. Panel B contains the out-of-sample results. See Table 3 for detailed description. Market excess returns and predictors are standardized to have a mean zero and standard deviation of one. The sample period is 1996:1 to 2017:12.

Panel A: In-sample Prediction Comparing with Risk-Based Predictors								
	$h = 1$		$h = 3$		$h = 6$		$h = 12$	
	β	R^2	β	R^2	β	R^2	β	R^2
IVS	0.19	3.38	0.26	6.76	0.25	6.14	0.13	1.62
	[3.02]		[4.33]		[4.09]		[2.03]	
	(0.00)		(0.00)		(0.00)		(0.23)	
Tail Risk	0.06	0.35	0.14	1.91	0.09	0.82	0.09	0.73
	[0.91]		[2.24]		[1.46]		[1.36]	
	(0.15)		(0.01)		(0.13)		(0.38)	
CATFIN	-0.08	0.67	-0.08	0.66	-0.05	0.22	-0.06	0.30
	[-0.86]		[-1.31]		[-0.76]		[-0.87]	
	(0.15)		(0.20)		(0.40)		(0.51)	
LIQ	0.05	0.22	0.10	0.96	0.08	0.64	0.10	1.06
	[0.66]		[1.58]		[1.29]		[1.64]	
	(0.45)		(0.08)		(0.18)		(0.10)	

Panel B: Out-of-sample Prediction Comparing with Risk-Based Predictors								
	$h = 1$		$h = 3$		$h = 6$		$h = 12$	
	R_{OS}^2	p	R_{OS}^2	p	R_{OS}^2	p	R_{OS}^2	p
IVS	3.23	0.08	3.18	0.04	2.58	0.01	-1.41	0.12
Tail Risk	-0.58	0.48	-1.04	0.15	-2.56	0.05	-3.66	0.17
CATFIN	-2.01	0.42	-1.76	0.34	-2.27	0.33	-5.16	0.32
LIQ	-0.50	0.70	1.15	0.14	-0.03	0.34	-2.11	0.68