Forward Moments and Risk Premia Predictability

Panayiotis C. Andreou[*,†] Anastasios Kagkadis[‡] Dennis Philip[†]
Abderrahim Taamouti[†]

October 2016

[*]Department of Commerce, Finance and Shipping, Cyprus University of Technology, 140, Ayiou Andreou Street, 3603 Lemesos, Cyprus; Email: benz@pandreou.com

[†]Durham University Business School, Mill Hill Lane, Durham DH1 3LB, UK; Emails: dennis.philip@durham.ac.uk, abderrahim.taamouti@durham.ac.uk

[‡]Department of Accounting and Finance, Lancaster University Management School, Lancaster LA1 4YX, UK; Email: a.kagkadis@lancaster.ac.uk

Abstract

We rely on the recently established aggregation property of the second and third moments of returns to construct forward risk-neutral moments extracted from option prices. We show theoretically that according to standard affine no-arbitrage models, the forward moments should exhibit a factor structure, while the equity and the variance risk premia should also be affine functions of the same state variables. In light of this, we show that the factor extracted either from the forward variance or the forward skewness estimates exhibits strong predictability for the equity premium, both in-sample and out-of-sample. We also document that the forward skewness factor provides similar but often stronger predictability than the forward variance factor, and the combination of the two factors enhances the observed predictive performance. Finally, we find that the same forward moments factors that are designed to predict the equity premium, exhibit strong predictive power for the variance premium as well.

JEL Classification: G10, G11, G12

Keywords: Forward moments; Implied volatility surface; Partial least squares; Predictability of stock returns; Equity premium; Variance premium

1 Introduction

The forward-looking nature of the risk-neutral probability distribution has rendered the usage of option-implied moments of different maturities extremely popular for forecasting purposes among researchers. However, a new strand of the literature focuses on the predictive ability that can be offered by the term structure of *forward* implied moments rather than the implied moments themselves. More specifically, Bakshi, Panayotov and Skoulakis (2011) create measures of forward stock market variance and find that they are jointly successful in predicting future real activity, as well as stock market and treasury bill returns. Using a somewhat different specification, Luo and Zhang (2016) confirm the predictive ability of forward variances for stock market returns.

The contribution of this paper to the forward moments literature is twofold. First, we show theoretically that forward moments should exhibit predictive power for the equity as well as the variance premium. This is because under a standard affine no-arbitrage model forward moments are affine to the same factors driving the overall economy. Second, we create forward skewness measures and show empirically that, similar to forward variance, forward skewness exhibits strong predictability for the equity and variance premia. Moreover, the predictive power of forward skewness is even stronger than that of forward variance when predicting the equity premium especially at longer horizons.

The reason it is important to investigate the information content of forward moments on top of implied moments is because they provide, especially during turbulent periods, a better approximation of investors' current expectations about different future time periods. To see this, consider the example of the S&P 500 index forward and implied variance estimates on the 29^{th} September 2008. On that date, the annualized risk-neutral variance was 13.10% for the 3-month horizon, 10.05% for the 6-month horizon, 9.08% for the 9-month horizon and 8.77% for the 12-month horizon. This means that the level of investors' perceived variance was highly elevated and that the slope of the implied volatility term structure was negative, with investors worrying more about short-term rather than long-term market movements. However, the respective annualized forward variance estimates provide additional information. In particular, the forward variance for three-to-six months ahead was 7.01%, for six-to-nine months ahead it was 7.13% and for nine-to-twelve months ahead it was 7.85%. This means that while investors' perceived variance risk was the highest for the first quarter and much lower for the subsequent quarters, it was strictly increasing across the second, third and fourth quarters ahead. It is therefore evident, that in principle forward moments might contain information about investors' expectations that cannot be easily extracted from simple implied moments.

In order to create the forward moments estimates, we exploit the recently established Aggregation Property of Neuberger (2012). In particular, Neuberger (2012) provides alternative definitions of variance and skewness which have the characteristic that the sum of each moment of log returns across sequential time periods is equal to the respective moment over the overall time period examined. Most importantly, the only assumption that is required is that the asset price is a martingale. Furthermore, Kozhan, Neuberger, and Schneider (2013) provide a method for replicating these alternative definitions of variance and skewness via a positioning in call and put options.[1] Building on these two studies, we show how to use option prices to obtain not only forward variance but also forward skewness estimates.

Since according to our theoretical framework, the equity premium, the variance premium and all the forward moments are affine to the same common factors, it is reasonable to extract the relevant set of factors and use it to predict the equity and variance risk premia. To do so, we rely on the partial least squares methodology of Kelly and Pruitt (2013, 2015), which condenses the cross-section of a set of predictors according to its covariance with the predicted variable. Therefore, unlike the standard principal components method, the partial least squares technique employed in the paper extracts from the series of the predictive variables the factors that are more relevant for forecasting purposes.

[1] The standard definition of variance of log returns does not satisfy the Aggregation Property under the presence of jumps in the asset price process and hence cannot be replicated *exactly* by a positioning in call and put options. For more details the reader can refer to Neuberger (2012) and Bondarenko (2014).

Empirically, we find that one factor, extracted from the series of either forward variances or forward skewnesses and designed to predict the 1-month horizon equity premium, exhibits statistically and economically significant predictive power for future market returns for horizons up to twelve months ahead. In particular, in-sample predictability analysis shows that a one standard deviation increase in the forward variance factor is associated with an annualized excess monthly (quarterly) return of 11.85% (10.20%), while a one standard deviation increase in the forward skewness factor is associated with an annualized excess monthly (quarterly) return of 12.55% (13.34%). The respective R^2s for the forward variance factor are 4.48% and 8.77%, while for the forward skewness factor they are 5.02% and 14.99%. Furthermore, the results show that the predictability of the two forward moments factors remains intact after controlling for the simple implied moments or a wide range of alternative predictors.

The strong predictive power of the forward moments factors is also confirmed in the out-of-sample analysis. In particular, a predictive model with either of the factors clearly outperforms the historical average model across most horizons examined. For example, the forward variance model out-of-sample R^2 is 1.93% for monthly returns and 2.30% for quarterly returns, while the forward skewness model out-of-sample R^2 is 0.72% for monthly returns and 4.08% for quarterly returns. Results also show that the inclusion of the forward skewness factor to a predictive model that would otherwise include only the forward variance factor provides additional predictive power for the equity premium. The out-of-sample results appear to be also economically significant since a market-timing trading strategy that is based on the predictive power of the forward moments factors offers higher Sharpe ratios and increased utility (excess certainty equivalent return) to a mean-variance investor who would otherwise allocate her wealth by considering the recursively estimated historical average.

Finally, we show that the same forward variance and forward skewness factor that is used for the equity premium predictability is also a very successful predictor of the variance premium across horizons both in-sample and out-of-sample. This finding supports the theoretical framework presented in the paper and rationalizes the usage of proxies of the variance premium as predictors of the equity premium (see for example, Bollerslev, Tauchen and Zhou, 2009; Drechsler and Yaron, 2011; Bollerslev, Marrone, Xu and Zho, 2014).

The remainder of the paper is structured as follows. Section 2 presents the theoretical framework, while Section 3 describes the econometric methodology. Section 4 provides details regarding the data and the construction of the main variables used in the study. Section 5 discusses the empirical evidence from the equity premium predictability. Section 6 reports the results from the variance premium predictability. Finally, Section 7 concludes.

2 Theoretical Motivation

This section aims to provide a theoretical motivation of the empirical analysis considered in this paper. In particular, we show how the forward variance and forward high-order moments term structure can be used to recover the risk factors driving the equity risk premium.

We first consider the following notations. The one-period ahead excess returns derived from holding an equity are defined as follows:

$$r_{t+1} = p_{t+1} - p_t - r_{f,t},$$

where p_t is the log price of the equity at time t and $r_{f,t}$ is the log of one-period ahead risk-free rate. The excess returns over any horizon τ are given by:

$$r_{t,t+\tau} = \sum_{j=1}^{\tau} r_{t+j}.$$

Furthermore, we consider an economy driven by K state variables, say X_t, which satisfies the following three properties: (i) the joint distribution of r_{t+1} and X_{t+1} belongs to the family of affine jump-diffusion continuous-time (or discretized) models [see Duffie, Pan, and Singleton, 2000]; (ii) the risk-free rate $r_{f,t}$ is an affine function of X_t; and (iii) the stochastic discount factor is an exponential affine function of X_{t+1} and r_{t+1} [see Gourieroux and Monfort, 2007; Christoffersen et al., 2010]. Feunou et al. (2014) have formalized these properties to show that this class of models nests a wide array of discrete-time asset-pricing models. For example, the affine long-run risk models that are based on Epstein–Zin–Weil preferences fit these properties; see Bansal and Yaron (2004) and Eraker (2008) among others. Under this class of models, the equity premium over an investment horizon τ can be expressed as follows:

$$EP_t(\tau) = E^{\mathbb{P}}[r_{t,t+\tau}] = \beta_{a,0}(\tau) + \beta_a(\tau)^{\top} X_t, \tag{1}$$

where the coefficients $\beta_{a,0}(\tau)$ and $\beta_a(\tau)$ are functions of the underlying model's parameters; see Feunou et al. (2014). The estimation of risk–return trade-off equation in (1) is the main focus of the present paper. In this equation, the coefficient $\beta_a(\tau)$ characterizes the returns that are required by investors to bear the risk associated with variations in X_t. If the risk factors X_t were observable, then the coefficients $\beta_{a,0}(\tau)$ and $\beta_a(\tau)$ could be estimated directly via ordinary least squares (OLS). However, X_t is latent and this makes the estimation of Equation (1) infeasible.

We next show how the risk factors X_t can be revealed using the term structure of risk-neutral *forward* moments (cumulants). Before doing so, we begin by recalling some results from Feunou et al. (2014) that show that in the context of the above mentioned affine models, the standard risk-neutral moments can be expressed as affine functions of risk factors X_t. In particular, one can show that the conditional variance of excess returns under risk-neutral measure, $\mathbb{Q}$, over a horizon τ is an affine function of risk factor X_t :

$$Var_t^{\mathbb{Q}}(\tau) = Var_t^{\mathbb{Q}}[r_{t,t+\tau}] = \beta_{vr,0}(\tau) + \beta_{vr}(\tau)^{\top} X_t, \tag{2}$$

where the coefficients $\beta_{vr,0}(\tau)$ and $\beta_{vr}(\tau)$ are functions of the underlying model's parameters. Equation (2) indicates that the variance at different maturities display a factor structure with dimension K. The above results can be generalized for any moment (cumulant) of order n

of excess returns over a horizon τ, say $M_{t,n}^{\mathbb{Q}}(\tau)$, for $n > 2$,

$$M_{t,n}^{\mathbb{Q}}(\tau) = \beta_{n,0}(\tau) + \beta_{n,X}(\tau)^{\top} X_t, \tag{3}$$

where again the coefficients $\beta_{n,0}(\tau)$ and $\beta_{n,X}(\tau)$ depend on the underlying model's parameters.

The novelty of the present paper is we use the risk-neutral *forward* moments (cumulants) to reveal and measure the risk factors X_t. The main reason of using the risk-neutral *forward* moments instead of the standard risk-neutral moments, as in Feunou et al. (2014), is because the former provide a better approximation of the term structure of uncertainty. In the following, we focus on the term structure of risk-neutral *forward* variance and skewness.

Formally, the risk-neutral *forward* variance between two maturities τ_1 and τ_2 is defined as follows:

$$Var_t^{\mathbb{Q}}(\tau_1, \tau_2) = Var_t^{\mathbb{Q}}\left[r_{t+\tau_1, t+\tau_2}\right].$$

Now, recall that Neuberger (2012) postulates that any real-valued function g of an adapted process Y has the aggregate property if for any $t \le \tau_1 \le \tau_2$, we have

$$E_t\left[g\left(X_{\tau_2} - X_t\right)\right] = E_t\left[g\left(X_{\tau_2} - X_{\tau_1}\right)\right] + E_t\left[g\left(X_{\tau_1} - X_t\right)\right]. \tag{4}$$

Assuming that the (*forward*) asset price P is a martingale, Neuberger (2012) and Kozhan, Neuberger, and Shneider (2013) define the following log and entropy variances, respectively, as:

$$G_{t,\tau_2}^{V} = E_t\left[\frac{F_{\tau_2}}{F_t} - 1 - \ln\left(\frac{F_{\tau_2}}{F_t}\right)\right] \text{ and } G_{t,\tau_2}^{E} = E_t\left[2\frac{F_{\tau_2}}{F_t}\ln\left(\frac{F_{\tau_2}}{F_t}\right) - \frac{F_{\tau_2}}{F_t} + 1\right],$$

where the functions inside the brackets have the aggregation property and converge to the second moment of returns. Observe that G_{t,τ_2}^{V} can be regarded as the implied variance of stock returns, i.e. $G_{t,\tau_2}^{V} = IV_{t,\tau_2}$. Similarly, Neuberger (2012) and Kozhan et al. (2013) define the skewness as follows:

$$G^S_{t,\tau_2} = E_t \left[6\frac{F_{\tau_2}}{F_t} \ln(\frac{F_{\tau_2}}{F_t}) - 2\frac{F_{\tau_2}}{F_t} + \ln(\frac{F_{\tau_2}}{F_t}) + 2 \right],$$

where the function inside the brackets has the aggregation property and converge to the third moment of returns. Notice that G^S_{t,τ_2} can be written as the difference between the two previously described variance measures. Consequently, G^S_{t,τ_2} can be regarded as the implied third moment of stock returns, i.e. $G^S_{t,\tau_2} = TM_{t,\tau_2}$.

Both implied variance and skewness are unbiased estimates of the true variance and skewness in the absence of any risk premia. Thus, from Equation (4), we can write, for any $t \leq \tau_1 \leq \tau_2$,

$$IV_{t,\tau_2} = IV_{t,\tau_1} + E_t \left[G^V_{\tau_1,\tau_2} \right], \tag{5}$$

$$TM_{t,\tau_2} = TM_{t,\tau_1} + E_t \left[G^S_{\tau_1,\tau_2} \right]. \tag{6}$$

Rearranging equations (5)-(6) we obtain:

$$FV_{t,\tau_1,\tau_2} = E_t \left[G^V_{\tau_1,\tau_2} \right] = IV_{t,\tau_2} - IV_{t,\tau_1}, \tag{7}$$

$$FS_{t,\tau_1,\tau_2} = E_t \left[G^S_{\tau_1,\tau_2} \right] = TM_{t,\tau_2} - TM_{t,\tau_1}, \tag{8}$$

where FV_{t,τ_1,τ_2} and FS_{t,τ_1,τ_2} are the time zero forward variance and skewness for the period u to t implied by the prices of OTM options at time zero. Consequently, from equations (2) and (7), we get:

$$FV_{t,\tau_1,\tau_2} = \beta_{vr,0}\left(\tau_1,\tau_2\right) + \beta_{vr}\left(\tau_1,\tau_2\right)^\top X_t, \tag{9}$$

where $\beta_{vr,0}\left(\tau_1,\tau_2\right) = \beta_{vr,0}\left(\tau_2\right) - \beta_{vr,0}\left(\tau_1\right)$ and $\beta_{vr}\left(\tau_1,\tau_2\right) = \beta_{vr}\left(\tau_2\right) - \beta_{vr,0}\left(\tau_1\right)$. Similarly, from equations (3) and (8), we get:

$$FS_{t,\tau_1,\tau_2} = \beta_{tm,0}\left(\tau_1,\tau_2\right) + \beta_{tm}\left(\tau_1,\tau_2\right)^\top X_t, \tag{10}$$

where $\beta_{tm,0}(\tau_1,\tau_2) = \beta_{tm,0}(\tau_2) - \beta_{tm,0}(\tau_1)$ and $\beta_{tm}(\tau_1,\tau_2) = \beta_{tm}(\tau_2) - \beta_{tm}(\tau_1)$.

Now, observe that the expressions of the risk-neutral forward variance and skewness (third moment) in equations (9) and (10) can be used to reveal the risk factors X_t, which in turn can be used to estimate the risk–return trade-off Equation in (1). This can be done with help of model-free measures of risk-neutral *forward* variance and skewness that are available directly from option prices. However, the measured risk-neutral *forward* variance and skewness differs from the true values. In particular, we have

$$FV_{t,\tau_1,\tau_2} = \widetilde{FV}_{t,\tau_1,\tau_2} + v_t(\tau_1,\tau_2) \text{ and } FS_{t,\tau_1,\tau_2} = \widetilde{FS}_{t,\tau_1,\tau_2} + s_t(\tau_1,\tau_2), \tag{11}$$

where the measurement errors $v_t(\tau_1,\tau_2)$ and $s_t(\tau_1,\tau_2)$ are assumed to be uncorrelated with the model-free measures of risk-neutral forward variance and skewness $\widetilde{FV}_{t,\tau_1,\tau_2}$ and $\widetilde{FS}_{t,\tau_1,\tau_2}$, respectively. We rely on the nonparametric approach of Bakshi and Madan (2000) to measure $\widetilde{FV}_{t,\tau_1,\tau_2}$ and $\widetilde{FS}_{t,\tau_1,\tau_2}$.

Next, using equations (9)-(10) and stacking the measurements in Equation (11) across horizons $\tau = \tau_1, \tau_2, ..., \tau_q$, we obtain:

$$\widetilde{FV}_t + v_t = B_{0,vr} + B_{vr}X_t,$$

$$\widetilde{FS}_t + s_t = B_{0,tm} + B_{tm}X_t,$$

where the $q \times 1$ vectors $B_{0,vr}$ and $B_{0,tm}$ stack the constants $\beta_{vr,0}(\tau_1,\tau_2)$ and $\beta_{tm,0}(\tau_1,\tau_2)$, respectively, and the $q \times K$ matrices B_{vr} and B_{tm} stack the corresponding coefficients $\beta_{vr}(\tau_1,\tau_2)$ and $\beta_{tm}(\tau_1,\tau_2)$. We typically have more observations along the term structure than the underlying factors (i.e., $q > K$). We can then write,

$$X_t = -\bar{B}_{vr}B_{0,vr} + \bar{B}_{vr}\widetilde{FV}_t + \bar{B}_{vr}v_t, \tag{12}$$

$$X_t = -\bar{B}_{tm}B_{0,tm} + \bar{B}_{tm}\widetilde{FS}_t + \bar{B}_{tm}s_t, \tag{13}$$

where the $K \times q$ matrices $\bar{B}_{vr} = \left(B_{vr}^{\top} B_{vr}\right)^{-1} B_{vr}^{\top}$ and $\bar{B}_{tm} = \left(B_{tm}^{\top} B_{tm}\right)^{-1} B_{tm}^{\top}$ are the left inverse of B_{vr} and B_{tm}, respectively. Hence, equations (12)-(13) show that one can use the *forward* variance and skewness term structures separately as a signal for the underlying risk factors. Alternatively, we can also combine the *forward* variance and skewness term structures to extract the factors X_t. In the next section, we use the Partial Least Squares (PLS) methodology from Kelly and Pruitt (2013, 2015) to extract the factor(s) X_t using $\widetilde{FV}$ and $\widetilde{FS}$. Once the risk factors X_t are extracted from model-free measures of risk-neutral *forward* variance and skewness, they will be used to estimate our Equation of interest in (1).

3 Econometric Methodology

In this paper, we rely on the partial least squares (PLS) methodology of Kelly and Pruitt (2013, 2015) to extract one factor X fusing $\widetilde{FV}$ (or $\widetilde{FS}$). The main characteristic of the PLS method is that it extracts the factor structure of a set of predictive variables according to their covariance with the forecasted variable. In other words, using PLS we can identify a factor that drives the set of the predictive variables but is also relevant for forecasting the target variable.

To implement PLS, we first run N time-series regressions:

$$q_{i,t} = \phi_{i,0} + \phi_i xr^e_{t,t+1} + \epsilon_{i,t}, \tag{14}$$

where $q_{i,t}$ is each of the forward moments and $xr^e_{t,t+1}$ is the subsequent one-month ahead excess market return. Intuitively, the excess market return is used as a proxy for the unobservable forward moments factor, and ϕ_i is the loading of each forward moment to that factor. As a second step, we run T cross-sectional regressions:

$$q_{i,t} = \varphi_t + X_t \widehat{\phi}_i + \varepsilon_{i,t}, \tag{15}$$

where $\widehat{\phi}_i$ is the estimated coefficient from the first step for each of the forward moments. Intuitively, by regressing the forward moments at each time period on the corresponding factor loadings stemming from the first-step regressions, we can estimate the time-series of the forward moments factor ($\widetilde{FV}$ or $\widetilde{FS}$).

Notice that we utilize the one-month ahead excess market return as a proxy for the true forward moments factors. However, in the subsequent empirical analysis we show that the factors estimated through this procedure exhibit predictive power for long-horizon excess market returns and excess variance as well.

4 Data and Variables

This section provides details regarding the estimation of the main and alternative predictive variables used and the study, as well as summary statistics.

4.1 Options data and forward moments

To create a term structure of aggregate market forward moments, we use S&P 500 index options data from OptionMetrics. More specifically, we utilize the volatility surface file that provides implied volatilities for a given range of standardized deltas and maturities. The interpolated implied volatility surface is estimated based on a kernel smoothing algorithm using call and put options of different strike prices and maturities. We discard in-the-money options, i.e. options with an absolute value of delta higher than 0.5. Our sample period is 1996:01-2015:08 and the monthly time-series of forward moments is estimated using data on the last-but-one trading day of each month.[2]

On a given day, we estimate implied variance and skewness for three-, six-, nine- and twelve months ahead. Neuberger (2012) and Kozhan, Neuberger and Schneider (2013) show

[2]The one-day lag rule is used to account for the fact that, until the 4$^{\text{th}}$ March 2008, the data provided by OptionMetrics stem from closing prices that are recorded two minutes after the closure of the stock market (Battalio and Schultz, 2006). Moreover, it gives real-time investors the necessary time to analyze the options data.

that Equations (2)-(2) can be replicated exactly by a positioning in OTM call and puts options following the results of Bakshi and Madan (2000) and Carr and Madan (2001):

$$IV_t^T = \frac{2}{B_t^T}\left[\int_0^{F_t^T} \frac{P_t^T[K]}{K^2}dK + \int_{F_t^T}^{\infty} \frac{C_t^T[K]}{K^2}dK\right], \tag{16}$$

$$IS_t^T = \frac{6}{B_t^T}\left[\int_{F_t^T}^{\infty} \frac{K - F_t^T}{K^2 F_t^T} C_t^T[K]\,dK - \int_0^{F_t^T} \frac{F_t^T - K}{K^2 F_t^T} P_t^T[K]\,dK\right], \tag{17}$$

where $B_t^T = e^{-r(T-t)}$ is the price of a risk-free bond, F_t^T is the forward S&P 500 index level at time t, and $P_t^T[K]$ and $C_t^T[K]$ are the prices of a put and a call option respectively with strike price K and time to maturity $T - t$. The main issue regarding the usage of the above formulae is that they require a continuum of option prices while the available data is only discrete. Therefore, for each cross-section of implied volatilities we interpolate into the range of available moneyness levels using a smoothing cubic spline with smoothing parameter of 0.99 and extrapolate outside this range using the respective boundary values (see also Buss and Vilkov, 2012, and DeMiguel, Plyakha, Uppal and Vilkov, 2013). This way, we obtain a set of 1000 implied volatilities that cover the moneyness range from 0.0001 to 3. Finally, these implied volatilities are transformed into option prices and the trapezoidal approximation is used for the computation of the integrals in Equations (16) and (17).

Once we have the estimates of constant maturity implied moments, we use Equations (7)-(8) to create forward moments for three-to-six, six-to-nine and nine-to-twelve months ahead. Then, using the PLS method described in the previous section and the estimated forward moments we extract one factor from the forward variance vector and one factor from the forward skewness vector. Moreover, we keep the three-month ahead implied variance and skewness and use them as control variables in the subsequent empirical analysis.

4.2 Other variables

The remainder of the predictor variables include the aggregate dividend-price ratio (d-p, Fama and French, 1988, and Campbell and Shiller, 1988a,b), the market dividend-payout ratio (d-e, Campbell and Shiller, 1988a and Lamont, 1998), the yield term spread (TERM, Campbell, 1987 and Fama and French, 1989), the default spread (DEF, Keim and Stambaugh, 1986 and Fama and French, 1989), the relative short-term risk-free rate (RREL, Campbell, 1991), the stock market variance (SVAR, Guo, 2006) and the tail risk (TAIL, Kelly and Jiang, 2014). d-p is the difference between the log aggregate annual dividends and the log level of the S&P 500 index, while d-e is the difference between the log aggregate annual dividends and the log aggregate annual earnings. TERM is the difference between the 10-year bond yield and the 1-year bond yield, while DEF is the difference between BAA and AAA corporate bonds yields from Moody's. RREL is the difference between the 3-month t-bill rate and its moving average over the preceding twelve months. SVAR is the sum of squared daily returns of the S&P 500 index. Finally, TAIL captures the probability of extreme negative market returns and is constructed by applying Hill's (1975) estimator to the whole NYSE/AMEX/NASDAQ cross-section (share codes 10 and 11) of daily returns within a given month. Data on monthly market prices, dividends, and earnings are obtained from Robert Shiller's website. All interest rate data are obtained from the FRED database of the Federal Reserve Bank of St. Louis. The stock market variance data come from Amit Goyal's website.

As a proxy for the equity premium we use CRSP value-weighted index excess market returns. Excess returns are estimated by subtracting from the monthly log-return the (log of) the one-month Treasury bill rate obtained from Kenneth French's website. Continuously compounded excess market returns for longer horizons are created by taking cumulative sums of monthly excess market returns.

4.3 Summary statistics

Table 1 provides descriptive statistics for the predictive variables used in the study. FV exhibits positive skewness and high excess kurtosis, while FS exhibits slightly negative skewness and even higher kurtosis. Both variables exhibit modest first-order autocorrelation coefficients (0.64 and 0.53 respectively). IV and IS, on the other hand, exhibit more extreme higher moments and are also more persistent with autocorrelation coefficients of 0.86 and 0.75 respectively. With the exception of SVAR and TAIL, the remainder of the predictors are highly persistent, with autocorrelation coefficients that are close to unity.

Figure 1 plots the estimated forward variance and forward skewness factors together with the implied variance and implied skewness. The top panels show that both FV and FS exhibit several spikes in the period of the Asian financial crisis (1997:07 to 1997:12) and the Russian default (1998:08 to 1998:09), while FS exhibits also a big spike in 2000:02 just before the all-time high level of the NASDAQ index. Both factors remain relatively stable in the subsequent years and become again very volatile in the period of the Lehman Brothers' collapse (2008:09 to 2009:03) and the subsequent European sovereign debt crisis. For example, FS exhibits a big spike between 2010:05 and 2010:07 which corresponds to the first bailout agreement for Greece. In general, it can be seen that FV and FS exhibit similarities across time but their peaks and downs can differ in both timing and magnitude. The bottom panels show that IV and IS exhibit very similar but opposite patterns. More importantly, we can observe that the pattern of IV is quite different from that of FV and this is also true for IS and FS. The above relations are also apparent in Table 2, which presents the correlation coefficients among the predictive variables. As expected, FV and FS are positively correlated (0.62), while IV and IS are very highly negatively correlated (-0.96).[3] The correlation between FV and IV is 0.10, while the respective correlation between FS and IS is only 0.05. It is, therefore, apparent that the information embedded in our

[3]Recall that our skewness measure is the third non-central moment and is not scaled by variance. The respective correlation coefficient presented in Neuberger (2012) is -0.95.

forward moments factors is very different from that contained in the implied moments. FV and FS exhibit also quite low correlations with the rest of the predictors, the highest being the correlation between FS and SVAR (-0.26).

5 Equity Premium Predictability

5.1 In-sample analysis

We gauge the predictive power of the estimated forward moments factors, by running multiple-horizon regressions of excess stock market returns of the following form:

$$xr^e_{t,t+h} = \alpha_h + \boldsymbol{\beta}'_h \mathbf{z}_t + \varepsilon_{t,t+h}, \tag{18}$$

where $xr^e_{t,t+h} = \left(\frac{1200}{h}\right)\left[xr^e_{t+1} + xr^e_{t+2} + ... + xr^e_{t+h}\right]$ is the annualized h-month excess return of the CRSP value-weighted index and $\mathbf{z}_t$ is the vector of predictors. The regression analysis covers the period 1996:01-2015:08 and for each forecasting horizon we lose h observations. To avoid spurious statistical inference stemming from overlapping observations, we employ Newey and West (1987) as well as Hodrick (1992) standard errors with lag length equal to the forecasting horizon. The beta coefficients reported in the subsequent tables have been scaled and can be interpreted as the percentage annualized excess market return caused by a one standard deviation increase in each regressor.

Table 3 reports the results of 1-, 2-, 3-, 6-, 9-, and 12-month horizon univariate predictive regressions for the forward variance factor and the forward skewness factor. It can be seen that both FV and FS are highly significant until the 6-month horizon, while FS continues to be significant at either the 5% or the 1% level until the 12-month horizon.[4] Moreover, the slope coefficients are economically significant: for example, a one standard deviation

[4] Recall that, despite predicting the equity premium at different horizons, the forward moments factors that are used throughout the paper are the same and are designed to predict the 1-month ahead equity premium.

increase in FV predicts an annualized excess monthly return of 11.85%, while a one standard deviation increase in FS predicts an annualized excess monthly return of 12.55%. The economic significance gradually tapers off for horizons longer than three months ahead. The R^2 values show a similar patter. In particular, the 1-month ahead R^2 of FV is a sizeable 4.48% that grows to 8.77% for the 3-month horizon and then gradually decreases. Similarly, the 1-month ahead R^2 of FS is 5.02% and grows to 14.99% for the 3-month horizon before it starts decreasing.

Tables 4 and 5 assess the robustness of the forecasting power of FV and FS to the presence of IV, IS and our set of alternative economic predictors. Table 4 shows that FV remains significant at either the 5% or 1% level in all cases when horizons up to six months ahead are examined. From the rest of the variables considered, only d-p and RREL exhibit some modest predictability, while IS becomes significant at longer horizons but only when Newey-West standard errors are considered. Similarly, Table 5 demonstrates that, consistent with the univariate analysis, FS remains significant at either the 5% or 1% level in all but one case (the bivariate model with d-p at the 12-month horizon when Hodrick standard errors are considered) across all horizons examined. As in the case of FV, we find some evidence of predictability for d-p and RREL, while IV and IS turn occasionally significant but only when Newey-West standard errors are employed.

Overall, the results of this section suggest that the both the forward variance and the forward skewness factor exhibit statistically and economically significant predictability for the equity premium with quite high coefficients of determination. Moreover, the predictability of forward skewness appears to be relatively stronger than that of forward variance especially for horizons longer than six months ahead. Finally, the predictability of the forward moments remains intact when controlling for the implied variance, implied skewness and a set of alternative economic predictors.

5.2 Out-of-sample analysis

While in-sample (IS) predictability tests exhibit higher statistical power (Inoue and Kilian, 2004), an examination of the out-of-sample (OS) predictability of the forward moments factors is of particular importance for several reasons. First, OS predictability tests avoid potential over-fitting problems (Goyal and Welch, 2003, 2008). Second, they employ only data that are available to investors in real-time when making their forecasts. Third, they are not affected by the small-sample biases of the PLS method, discussed in Kelly and Pruitt (2013). Therefore, in this Section we evaluate the OS performance of the forward moments factors for 1-, 2-, 3-, 6-, 9-, and 12-month horizons.

Following Goyal and Welch (2003, 2008), Campbell and Thompson (2008), Rapach, Strauss and Zhou (2010), Kelly and Pruitt (2013) and Huang, Jiang, Tu and Zhou (2015), among others, we estimate the model in Equation (18) recursively using observations $1, ..., s$. Next, based on the estimated parameters, we form for each time period $s = s_0, ..., T - h$, with T being the total number of months in our sample period and h the forecasting horizon, the OS forecasts for the expected excess market return using the concurrent values of the predictive variables examined:

$$\widehat{xr}^{e}_{s,s+h} = \widehat{\alpha}_s + \widehat{\boldsymbol{\beta}}'_s \mathbf{z}_s. \tag{19}$$

This way we create a series of T_{OS} OS forecasts starting from 2000:01. It is important to note that when the predictive variable is one of the forward moments factors, the PLS method employs only data that are known at time s. The OS forecasts of each predictive model are compared to a series of recursively estimated historical averages, which correspond to OS forecasts of a restricted model with only a constant as a regressor. The evaluation of the OS predictive performance is based on three measures.

The first measure is the OS R^2, denoted by R^2_{OS}, which takes the form:

$$R^2_{OS} = 1 - \frac{MSE_U}{MSE_R}, \tag{20}$$

where $MSE_U = \frac{1}{T_{OS}} \sum_{s=s_0}^{T-h} \left(xr^e_{s,s+h} - \widehat{xr}^e_{s,s+h} \right)^2$ is the mean squared error of the unrestricted model and $MSE_R = \frac{1}{T_{OS}} \sum_{s=s_0}^{T-h} \left(xr^e_{s,s+h} - \widetilde{xr}^e_{s,s+h} \right)^2$ is the mean squared error of the restricted model, with $\widetilde{xr}^e_{s,s+h}$ being the recursively estimated historical average. R^2_{OS} takes positive values whenever the unrestricted model outperforms the restricted model in terms of predictive power (i.e. $MSE_U < MSE_R$).

The second measure of OS performance is the MSE-F test from McCracken (2007):

$$MSE - F = (T_{OS} - h + 1) \frac{MSE_R - MSE_U}{MSE_U}, \tag{21}$$

which tests the null hypothesis that the restricted model's MSE is less than or equal to the unrestricted model's MSE. McCracken (2007) shows that the F-statistic follows a non-standard normal distribution and provides appropriate critical values using Monte Carlo simulations.

Finally, the third measure is the encompassing test of Clark and McCracken (2001):

$$\begin{aligned} ENC - NEW \quad = \quad & \frac{(T_{OS} - h + 1)}{T_{OS}} \\ & \frac{\sum_{s=s_0}^{T-h} \left[\left(xr^e_{s,s+h} - \widehat{xr}^e_{s,s+h} \right)^2 - \left(xr^e_{s,s+h} - \widehat{xr}^e_{s,s+h} \right) \left(xr^e_{s,s+h} - \widetilde{xr}^e_{s,s+h} \right) \right]}{MSE_U}, \end{aligned} \tag{22}$$

which examines whether the restricted model encompasses the unrestricted model, meaning that the unrestricted model does not improve the forecasting ability of the restricted model. Appropriate critical values based on Monte Carlo simulations are also provided by the authors.

Results for the predictive models with the two forward moments factors are presented in Panel A of Table 6, while results for the predictive models with the alternative predictors are presented in Panel B of Table 6. It can be seen that FV exhibits positive and quite large R^2_{OS} values for horizons up to six months ahead, ranging from 0.45% to 2.30%. Moreover, the outperformance of the unrestricted model based on FV is statistically significant at the

5% level in all but one case (MSE-F statistic for the 6-month horizon where the significance is at the 10% level) when considering those horizons. In a similar vein, FS exhibits positive and even larger R^2_{OS} values, ranging from 0.72% to 4.08%, for all horizons up to nine months ahead. The outperformance of the FS model compared to the historical average model for those horizons is also statistically significant at the 5% level in all but one case (MSE-F statistic for the 1-month horizon where the significance is at the 10% level). The above results from univariate OS predictability tests are in line with the IS results presented in the previous section, since both forward moments factors exhibit significant equity premium predictability, with FS having slightly stronger power than FV.

Two additional observations are in order. First, a bivariate model that includes both FV and FS exhibits stronger predictability across all horizons than a model that takes into consideration only FV. More specifically, the bivariate model exhibits R^2_{OS}s which range from 1.51% to 7.12% when considering horizons up to nine months ahead, while the respective null hypotheses from the MSE-F and the ENC-NEW tests are rejected even more decisively. The last rows of Panel A compare the predictive ability of the bivariate model with that of the FV model (instead of the historical average model) and show a clear and statistically significant outperformance of the model that includes FS on top of FV. Second, Panel B shows that the OS predictability of the forward moments factors is much better than that of the simple implied moments or the alternative equity premium predictors. In particular, the only alternative variables that exhibit positive R^2_{OS}s are IS, which shows some modest predictability only for the 6- and 9-month horizons, and d-p which shows strong predictability when considering long horizons of nine and twelve months ahead.

Overall, the results from out-of-sample predictability confirm that the forward moments factors exhibit strong predictive performance across several horizons. This performance does not seem to be an artifact of econometric biases, can be exploited even by using only real-time data and is much stronger than what is offered by simple implied moments or other traditional predictors. Finally, the inclusion of the forward skewness factor to a predictive model that already utilizes the forward variance factor appears to be beneficial for predictive purposes.

5.3 Asset allocation

In this section, we assess the economic significance of the documented out-of-sample predictability of the forward moments factors. To this end, following Campbell and Thompson (2008), Ferreira and Santa-Clara (2011), Huang, Jiang, Tu and Zhou (2015) and Rapach, Ringgenberg and Zhou (2016), among others, we create a market-timing strategy that relies on the 1-month horizon OS forecasting power of the estimated factors and the alternative predictors.

More specifically, we consider a mean-variance investor who allocates her wealth every month between the market index and the risk-free asset. At the end of each month s, the investor makes a forecast for the 1-month ahead excess market return[5] using the procedure described in the previous section. Moreover, she forms an estimate of the market returns variance using all available data up to time s. Based on these estimates, the investor forms her portfolio weights as follows:

$$\omega_s = \frac{\widehat{xr}^e_{s,s+1}}{\gamma \widehat{\sigma}^2_{s,s+1}}, \tag{23}$$

where $\widehat{xr}^e_{s,s+1}$ is the OS forecast of the 1-month ahead excess stock market return, γ is the risk aversion coefficient, which is set as equal to three, and $\widehat{\sigma}^2_{s,s+1}$ is the estimate of the variance of the stock market return computed as the historical variance for the period $1, ..., s$. Following Campbell and Thompson (2008), we impose realistic leverage values by constraining the portfolio weight on the market index ω_s to lie between 0 and 1.5.

[5] In this section, the term return refers to simple return and not to logarithmic return.

The realized return from the above market-timing strategy can be represented by:

$$R_{p;s,s+1} = \omega_s R_{m;s,s+1} + (1 - \omega_s) R_{f;s,s+1}, \tag{24}$$

where $R_{m;s,s+1}$ denotes the simple market return and $R_{f;s,s+1}$ denotes the return of the riskless asset. Therefore, iterating this procedure forward, we create a series of realized portfolio returns based on the OS forecasting power of each forecasting model and we compare each strategy with a strategy based on the recursively estimated historical average (HAV).

For each trading strategy, we estimate the mean portfolio return, the standard deviation, and its Sharpe ratio. Moreover, since the Sharpe ratio weights equally the mean and volatility of the portfolio returns, we compute a certainty equivalent return (ΔCER) in excess of the HAV strategy:

$$\Delta CER = E(R_{p;s,s+1}) - E(\bar{R}_{p;s,s+1}) + \frac{\gamma}{2}\left[Var(\bar{R}_{p;s,s+1}) - Var(R_{p;s,s+1})\right], \tag{25}$$

where γ is the risk aversion coefficient, $R_{p;s,t+1}$ is the portfolio return of the predictive regression model strategy and $\bar{R}_{p;s,t+1}$ is the portfolio return of the HAV strategy. ΔCER essentially represents the change in the investor's utility resulting from her choice to follow the predictive regression strategy instead of the HAV strategy. It can also be interpreted as the annual "fee"that an investor is willing to pay to invest in the predictive regression strategy instead of investing in the HAV strategy. All measures are in annualized terms.

The results from the asset allocation exercise are reported in Table 7. The trading strategy that utilizes FV exhibits a very similar volatility to the HAV strategy (12.12% versus 12.16%) but much higher average return (3.94% versus 1.25%). This translates to a Sharpe ratio of 0.32, compared to 0.10 provided by the HAV strategy, and a positive ΔCER of 2.70%, which shows that the utility provided by the FV strategy is significantly higher than the utility provided by the strategy associated with the recursively estimated historical average. The results for the FS strategy are similar to those for the FV strategy, albeit slightly weaker. In particular, the strategy associated with FS exhibits a volatility of 13.67% and an average return of 2.83%, with these values generating a Sharpe ratio of 0.21 and a ΔCER of 0.99%. It is important to note that while, similar to the 1-month ahead OS predictability results, the FS strategy performs slightly worse than the FV strategy, the utilization of both variables into the predictive model leads to a strategy that exhibits better performance than the FV strategy. In particular, a strategy that relies on both FV and FS has a volatility of 12.18% and an average return of 4.94%. These values generate a Sharpe ratio of 0.41 (four times higher than the HAV strategy Sharpe ratio), and a ΔCER of 3.68%. Turning to the rest of the predictors, some of them exhibit good strategy performance despite their overall bad 1-month ahead OS predictability. More specifically, d-p, d-e and RREL exhibit high Sharpe ratios of 0.43, 0.33 and 0.35, and high ΔCERs of 3.84%, 2.81% and 2.95% respectively. TERM and DEF strategies also perform better than the HAV strategy with Sharpe ratios of 0.12 and 0.13, and positive ΔCERs of 0.30 and 0.59 accordingly. On the other hand, the two implied moments and the rest of the predictors do not outperform the naive HAV strategy.

Overall, even though the forward moments factors do not exhibit their highest OS predictability at the 1-month horizon, this section shows that a trading strategy that relies on the 1-month ahead predictive power of either the forward variance or the forward skewness factor outperforms by far a similar strategy that utilizes the recursively estimated historical average. Moreover, a strategy that relies on both forward moments factors provides even better performance and is outperformed only by the strategy that uses the aggregate dividend-price ratio.

6 Variance Premium Predictability

A series of recent papers, such as Bollerslev, Tauchen and Zhou (2009), Drechsler and Yaron (2011) and Bollerslev, Marrone, Xu and Zho (2014) investigate the predictive power of the variance premium for future market returns. In fact, Feunou et al. (2014) show that within the framework provided in Section 2, the variance premium over an investment horizon, $VP(t,\tau)$, should be an affine function of the same state variables, X_t, that drive the equity premium:

$$VP(t,\tau) = E^{\mathbb{Q}}\left[\sigma^2_{t,t+\tau}\right] - E^{\mathbb{P}}\left[\sigma^2_{t,t+\tau}\right] = \beta_{v,0}(\tau) + \beta_v(\tau)^\top X_t. \tag{26}$$

Therefore, in this section we investigate whether the forward moments factors that were used in the previous sections and were designed to predict the equity premium, exhibit significant forecasting power for the variance premium as well. Since the true variance premium is unobservable we measure the ex post excess variance of market returns:

$$xv^e_{t,t+\tau} = E^{\mathbb{Q}}\left[\sigma^2_{t,t+\tau}\right] - \sigma^2_{t,t+\tau}, \tag{27}$$

where $\sigma^2_{t+\tau} = \sum_{j=1}^{\tau} Var_t\left[xr^e_{t+j}\right]$. For a given horizon, excess variance is estimated as the difference between the implied variance extracted from S&P 500 index option prices and the S&P 500 index realized variance of the respective horizon. Realized variance data come from Hao Zhou's website.

Similarly to Sections 5.1 and 5.2, we focus on 1-, 2-, 3-, 6-, 9-, and 12-month horizon predictive regressions. Results from in-sample univariate predictability tests are reported in Panel A of Table 8, while results from out-of-sample univariate predictability tests are presented in Panel B of Table 8. Panel A shows that, similarly to the case of the equity premium, FV is a strong predictor of the variance premium as well. In particular, it is always significant at either the 5% or the 1% level when considering horizons up to six months ahead, while it continues to be significant at longer horizons when the Hodrick standard errors are considered.[6] A similar significance, albeit slightly lower for short horizons and slightly

[6] The reason Hodrick standard errors exhibit such a strong significance is due to the nature of the predicted variable. In particular, the 1-month ahead variance premium which is used for the computation of the Hodrick standard errors across horizons, is much smaller in magnitude than the 1-month ahead equity premium. Therefore, even if the regression coefficient is also relatively smaller, the resulting t-statistic ends up being extremely high. For more details about the methodology the reader can refer to Hodrick (1992) and Ang and Bekaert (2007).

higher for longer horizons, is also apparent for FS. Moreover, for both factors the regression coefficients are economically significant. For example, a one standard deviation increase in FV forecasts an annualized monthly excess variance of 1.20%, while a one standard deviation increase in FS forecasts an annualized monthly excess variance of 0.93%. The R^2s are quite high taking their maximum values at the 1-month horizon, with 8.39% for FV and 5.00% for FS, and gradually tappering off as the horizon increases.

Panel B shows that the forward moments factors are strong predictors of the variance premium also out-of-sample. More specifically, the R^2_{OS} values are positive in all but one case (12-month horizon for FV), with values ranging from -0.18% to 6.18% for FV and from 1.64% to 5.85% for FS. Moreover, the outperformance of the two forward moments models compared to the restricted (historical average) model is always significant at the 5% level when considering horizons up to nine months ahead, irrespective of the statistic examined. In general, the predictability of FV appears stronger than that of FS for horizons up to six months ahead, while the opposite is true for longer horizons. This result is in line with those of Panel A.

Summarizing, this section provides empirical evidence showing that the same forward variance and skewness factors that are designed to predict the 1-month ahead equity premium are also strong predictors of the variance premium at various horizons. This finding supports the theoretical framework according to which both risk premia are affine functions of the same risk factors and rationalizes the usage of variance premium proxies as predictors for the equity premium.

7 Conclusion

We rely on the recently established aggregation property of the second and third moments of returns to construct forward risk-neutral moments extracted from option prices. We show theoretically that according to standard affine no-arbitrage models, the forward moments should exhibit a factor structure, while the equity and the variance risk premia should also be affine functions of the same state variables. In light of this, we use a partial least squares technique to extract the factor that maximizes the covariance between the forward moments and the equity premium. We show empirically that the factor extracted either from the forward variance or the forward skewness estimates exhibits strong predictability for the equity premium, both in-sample and out-of-sample. Moreover, it is robust to and outperforms in terms of predictive power simple measures of risk-neutral variance and skewness as well as a wide series of traditional predictors. We also document that the forward skewness factor provides similar but often stronger predictability than the forward variance factor and the combination of the two factors enhances the observed predictive performance. Finally, we find that the same forward moments factors that are designed to predict the equity premium, exhibit strong predictive power for the variance premium as well. Overall, this paper provides a new approach for capturing the information embedded in the option prices of different maturities and highlights the important predictive power of forward risk-neutral moments, and especially skewness, for the equity and variance premia.

References

[1] Ang, A. and G. Bekaert. 2007. Stock return predictability: is it there? Review of Financial Studies 20, 651-707.

[2] Bakshi, G., N. Kapadia and D. Madan. 2003. Stock return characteristics, skew laws, and the differential pricing of individual equity options. Review of Financial Studies 16, 101-143.

[3] Bakshi, G. and D. Madan. 2000. Spanning and derivative-security valuation. Journal of Financial Economics 55, 205-238.

[4] Bakshi, G., G. Panayotov, and G. Skoulakis. 2011. Improving the predictability of real economic activity and asset returns with forward variances inferred from option portfolios. Journal of Financial Economics 100, 475-495.

[5] Bansal, R. and A. Yaron. 2004. Risks for the long run: a potential resolution of asset pricing puzzles. Journal of Finance 59, 1481–1509.

[6] Bates, D. S. 2006. Maximum likelihood estimation of latent affine processes. Review of Financial Studies 19, 909–965.

[7] Battalio, R. and P. Schultz. 2006. Options and the bubble. Journal of Finance 61, 2071-2102.

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

[9] Bollerslev, T., G. Tauchen and H. Zhou. 2009. Expected stock returns and variance risk premia. Review of Financial Studies 22, 4463-4492.

[10] Bondarenko, O. 2014. Variance trading and market price of variance risk. Journal of Econometrics 180, 81-97.

[11] Buss, A. and G. Vilkov. 2012. Measuring equity risk with option-implied correlations. Review of Financial Studies 25, 3113-3140.

[12] Campbell, J. Y. 1987. Stock returns and the term structure. Journal of Financial Economics 18, 373-399.

[13] Campbell, J. Y. 1991. A variance decomposition for stock returns. Economic Journal 101, 157-179.

[14] Campbell, J. Y. and R. J. Shiller. 1988a. Stock prices, earnings and expected dividends. Journal of Finance 43, 661-676.

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

[16] Campbell, J. Y. and S. Thompson. 2008. Predicting excess stock returns out of sample: can anything beat the historical average? Review of Financial Studies 21, 1509-1531.

[17] Carr, P. and D. Madan. 2001. Optimal positioning in derivative securities. Quantitative Finance 1, 19-37.

[18] Christoffersen, P., R. Elkamhi, B. Feunou and K. Jacobs. 2010. Option valuation with conditional heteroskedasticity and nonnormality. Review of Financial Studies 23, 2139–2183.

[19] Christoffersen, P., K. Jacobs, C. Ornthanalai and Y. Wang. 2008. Option valuation with long-run and short-run volatility components. Journal of Financial Economics 90, 272–297.

[20] Clark, T. and M. McCracken. 2001. Tests of equal forecast accuracy and encompassing for nested models. Journal of Econometrics 105, 85-110.

[21] Clark, T. and K. West. 2007. Approximately normal tests for equal predictive accuracy in nested models. Journal of Econometrics 138, 291-311.

[22] DeMiguel, V., Y. Plyakha, R. Uppal and G. Vilkov. 2013. Improving portfolio selection using option-implied volatility and skewness. Journal of Financial and Quantitative Analysis 48, 1813–1845.

[23] Drechsler, I. and A. Yaron. 2011. What's vol got to do with it. Review of Financial Studies 24, 1-45.

[24] Duffie, D., J. Pan and K. Singleton. 2000. Transform analysis and asset pricing for affine jump-diffusions. Econometrica 68, 1343–1376.

[25] Eraker, B. 2008. Affine general equilibrium models. Management Science 54, 2068-2080.

[26] Fama, E. and K. R. French. 1988. Dividend yields and expected stock returns. Journal of Financial Economics 22, 3-25.

[27] Fama, E. and K. R. French. 1989. Business conditions and expected returns on stocks and bonds. Journal of Financial Economics 25, 23-49.

[28] Ferreira, M. A. and P. Santa-Clara. 2011. Forecasting stock market returns: the sum of the parts is more than the whole. Journal of Financial Economics 100, 514-537.

[29] Feunou, B., J.-S. Fontaine, A. Taamouti, and R. Tédongap. 2014. Risk premium, variance premium, and the maturity structure of uncertainty. Review of Finance 18, 219-269.

[30] Gourieroux, C. and A. Monfort. 2007 Econometric specification of stochastic discount factor models. Journal of Econometrics 136, 509–530.

[31] Goyal, A. and P. Santa-Clara. 2003. Idiosyncratic risk matters! Journal of Finance 58, 975-1008.

[32] Goyal, A. and I. Welch. 2003. Predicting the equity premium with dividend ratios. Management Science 49, 639-654.

[33] Goyal, A. and I. Welch. 2008. A comprehensive look at the empirical performance of equity premium prediction. Review of Financial Studies 21, 1455-1508.

[34] Guo, H. 2006. On the out-of-sample predictability of stock returns. Journal of Business 79, 645-670.

[35] Hill, B. 1975. A simple general approach to inference about the tail of a distribution. Annals of Statistics 3, 1163-1174.

[36] Hodrick, R. J. 1992. Dividend yields and expected stock returns: alternative procedures for inference and measurement. Review of Financial Studies 2, 357-386.

[37] Huang, D., F. Jiang, J. Tu and G. Zhou. 2015. Investor sentiment aligned: a powerful predictor of stock returns. Review of Financial Studies 28, 791-837.

[38] Inoue, A. and L. Kilian. 2004. In-sample or out-of-sample tests of predictability: which one should we use? Econometric Reviews 23, 371-402.

[39] Keim, D. and R. Stambaugh. 1986. Predicting returns in the stock and bond markets. Journal of Financial Economics 17, 357-390.

[40] Kelly, B. and H. Jiang. 2014. Tail risk and asset prices. Review of Financial Studies 27, 2841-2871.

[41] Kelly, B. and S. Pruitt. 2013. Market expectations in the cross-section of present values. Journal of Finance 68, 1721-1756.

[42] Kelly, B. and S. Pruitt. 2015. The three-pass regression filter: a new approach to forecasting using many predictors. Journal of Econometrics 186, 294–316.

[43] Kozhan, R., A. Neuberger and P. Schneider. 2013. The skew risk premium in the equity index market. Review of Financial Studies 26, 2174-2203.

[44] Lamont, O. 1998. Earnings and expected returns. Journal of Finance 53, 1563-1587.

[45] Luo, X. and J. E. Zhang. 2016. Expected stock returns and forward variance. Journal of Financial Markets, forthcoming.

[46] McCracken, M. 2007. Asymptotics for out of sample tests of Granger causality. Journal of Econometrics 140, 719-752.

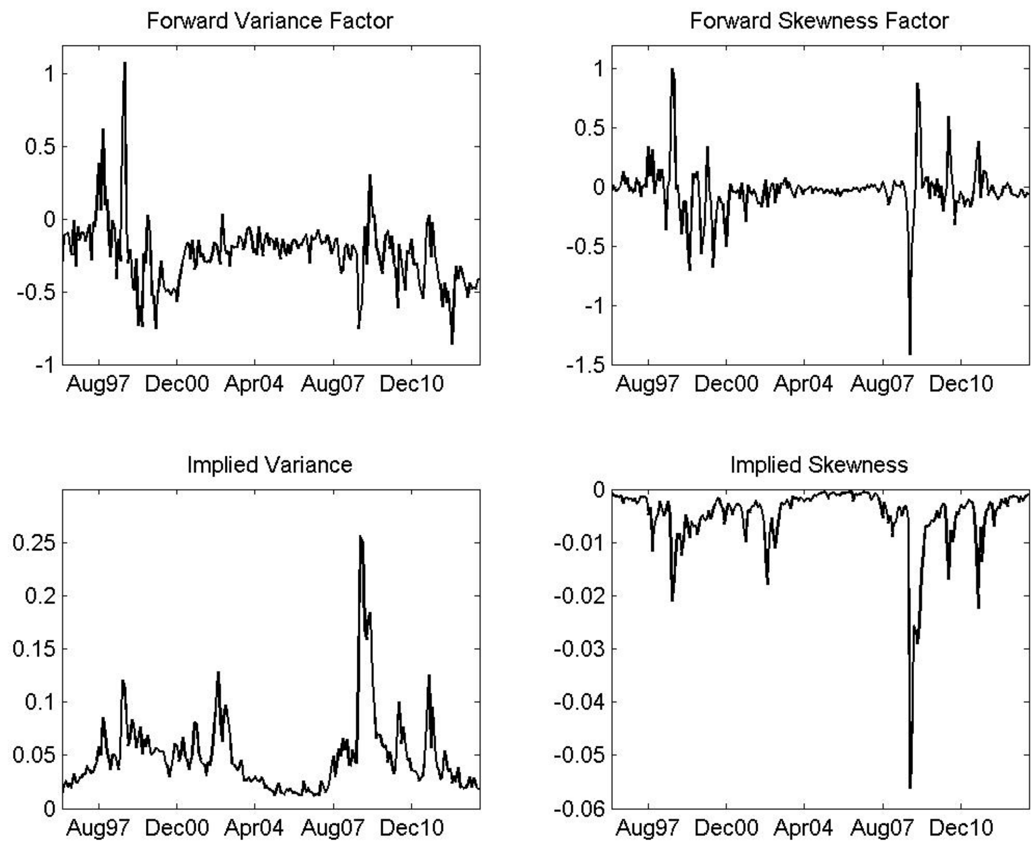
[47] Neuberger, A. 2012. Realized skewness. Review of Financial Studies 25, 3423-3455.

[48] Newey, W. and K. West. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703-708.

[49] Rapach, D. E., M. C. Ringgenberg and G. Zhou. 2016. Short interest and aggregate stock returns. Journal of Financial Economics, forthcoming.

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

Figure 1: Forward moments factors and implied moments



This figure plots the monthly time series of the forward variance factor, forward skewness factor, implied variance and implied skewness for the period 1996:01-2015:08.

Table 1: Descriptive statistics

	FV	**FS**	**IV**	**IS**	**d-p**	**d-e**	**TERM**	**DEF**	**RREL**	**SVAR**	**TAIL**
Mean	-0.25	-0.04	0.05	-0.01	-4.03	-0.88	0.01	0.01	-0.00	0.00	0.42
Median	-0.24	-0.03	0.04	-0.00	-4.02	-1.01	0.02	0.01	-0.00	0.00	0.42
Maximum	1.09	1.01	0.26	-0.00	-3.32	1.38	0.03	0.03	0.01	0.06	0.51
Minimum	-0.86	-1.41	0.01	-0.06	-4.50	-1.24	-0.00	0.01	-0.03	0.00	0.29
St. Dev.	0.22	0.21	0.04	0.01	0.22	0.46	0.01	0.00	0.01	0.01	0.04
Skewness	1.18	-0.14	2.80	-4.55	0.09	3.21	-0.12	2.97	-0.96	6.27	-0.47
Kurtosis	9.74	17.05	14.22	30.46	3.78	14.06	1.70	13.94	4.39	54.92	3.41
$\boldsymbol{\rho}(\mathbf{1})$	0.64	0.53	0.86	0.75	0.98	0.98	0.98	0.96	0.97	0.70	0.55

This table reports descriptive statistics of the forecasting variables used in the study. The forecasting variables are the forward variance factor (FV), forward skewness factor (FS), implied variance (IV), implied skewness (IS), dividend-price ratio (d-p), dividend payout ratio (d-e), yield term spread (TERM), default spread (DEF), relative short-term risk-free rate (RREL), stock market variance (SVAR) and tail risk (TAIL) (Panel B). The sample period is 1996:01-2015:08. $\rho(1)$ is the first-order autocorrelation coefficient.

Table 2: Correlation coefficients

	FV	FS	IV	IS	d-p	d-e	TERM	DEF	RREL	SVAR	TAIL
FV	1.00										
FS	0.62	1.00									
IV	0.10	-0.03	1.00								
IS	-0.09	0.05	-0.96	1.00							
d-p	0.07	0.18	0.26	-0.33	1.00						
d-e	0.13	0.11	0.58	-0.50	0.47	1.00					
TERM	-0.12	0.02	0.14	-0.13	0.40	0.32	1.00				
DEF	-0.04	-0.03	0.69	-0.68	0.60	0.74	0.33	1.00			
RREL	-0.03	0.02	-0.44	0.35	-0.16	-0.44	-0.32	-0.41	1.00		
SVAR	-0.01	-0.26	0.83	-0.86	0.29	0.38	0.09	0.59	-0.34	1.00	
TAIL	0.10	0.05	-0.44	0.44	-0.03	-0.10	0.02	-0.30	0.01	-0.44	1.00

This table reports correlation coefficients of the forecasting variables used in the study. The forecasting variables are the forward variance factor (FV), forward skewness factor (FS), implied variance (IV), implied skewness (IS), dividend-price ratio (d-p), dividend payout ratio (d-e), yield term spread (TERM), default spread (DEF), relative short-term risk-free rate (RREL), stock market variance (SVAR) and tail risk (TAIL). The sample period is 1996:01-2015:08.

Table 3: Forward moments predictability

Horizon	1-month	2-month	3-month	6-month	9-month	12-month
FV	11.85	10.52	10.20	6.14	4.04	3.35
	(3.20)***	(3.56)***	(4.04)***	(2.91)***	(1.91)*	(1.63)
	[2.63]***	[2.96]***	[3.40]***	[2.46]**	[1.82]*	[1.42]
$\mathbf{R^2}$ **(%)**	4.48	6.35	8.77	5.67	3.56	3.14
FS	12.55	12.93	13.34	8.47	5.71	4.38
	(3.19)***	(4.73)***	(5.68)***	(5.63)***	(3.12)***	(2.96)***
	[2.21]**	[2.87]***	[3.51]***	[3.36]***	[2.94]***	[2.49]**
$\mathbf{R^2}$ **(%)**	5.02	9.59	14.99	10.81	7.09	5.37

This table reports the in-sample results for univariate predictive regressions of the CRSP value-weighted index excess return on the forward variance factor (FV) and forward skewness factor (FS). The sample period is 1996:01-2015:08. Reported coefficients indicate the percentage annualized excess return resulting from a one standard deviation increase in each predictor variable. Newey and West (1987) and Hodrick (1992) t-statistics with lag length equal to the forecasting horizon are reported in parentheses and square brackets respectively. ***, ** and * denote significance at the 1%, 5% and 10% levels.

Table 4: Forward variance predictability controlling for other predictors

	FV	Z	R² (%)	FV	Z	R² (%)	FV	Z	R² (%)	FV	Z	R² (%)	FV	Z	R² (%)	FV	Z	R² (%)
Horizon	**1-month**			**2-month**			**3-month**			**6-month**			**9-month**			**12-month**		
IV	12.04	-1.69	4.57	10.48	0.46	6.36	10.13	0.66	8.81	5.81	3.53	7.53	3.72	3.42	6.08	3.09	2.86	5.41
	(3.10)***	(-0.31)		(3.43)***	(0.10)		(3.76)***	(0.15)		(2.67)***	(1.57)		(1.75)*	(1.88)*		(1.46)	(1.59)	
	[2.59]**	[-0.29]		[2.90]***	[0.09]		[3.31]***	[0.13]		[2.24]**	[0.81]		[1.62]	[0.97]		[1.28]	[0.96]	
IS	11.84	-0.20	4.48	10.33	-2.32	6.66	10.08	-1.44	8.94	5.75	-4.94	9.32	3.69	-4.53	8.00	3.07	-3.68	6.91
	(3.12)***	(-0.04)		(3.41)***	(-0.51)		(3.75)***	(-0.31)		(2.56)**	(-2.36)**		(1.73)*	(-3.42)***		(1.44)	(-3.02)***	
	[2.63]***	[-0.04]		[2.91]***	[-0.49]		[3.34]***	[-0.29]		[2.25]**	[-1.13]		[1.63]	[-1.35]		[1.28]	[-1.34]	
d-p	11.43	5.59	5.47	10.04	6.15	8.51	9.69	6.28	12.07	5.46	7.51	14.10	3.32	8.04	17.53	2.59	8.30	22.26
	(2.97)***	(1.10)		(3.25)***	(1.38)		(3.63)***	(1.47)		(2.85)***	(2.16)**		(2.04)**	(2.68)***		(2.12)**	(3.16)***	
	[2.53]**	[1.12]		[2.83]***	[1.26]		[3.25]***	[1.29]		[2.19]**	[1.62]		[1.51]	[1.83]*		[1.11]	[1.96]*	
d-e	11.90	-0.34	4.48	10.44	0.66	6.37	10.02	1.34	8.92	5.76	2.74	6.78	3.66	2.79	5.22	2.96	2.87	5.40
	(3.18)***	(-0.06)		(3.53)***	(0.14)		(3.88)***	(0.33)		(2.88)***	(0.84)		(1.85)*	(1.09)		(1.55)	(1.62)	
	[2.61]***	[-0.06]		[2.92]***	[0.13]		[3.32]***	[0.28]		[2.26]**	[0.66]		[1.61]	[0.76]		[1.23]	[0.84]	
TERM	12.19	2.87	4.74	10.85	2.80	6.79	10.53	2.92	9.48	6.47	3.08	7.08	4.44	3.61	6.35	3.85	4.64	9.09
	(3.20)***	(0.83)		(3.62)***	(0.94)		(4.16)***	(1.04)		(3.06)***	(1.13)		(2.15)**	(1.36)		(2.07)**	(1.89)*	
	[2.66]***	[0.81]		[2.99]***	[0.78]		[3.42]***	[0.80]		[2.54]**	[0.84]		[1.95]*	[1.00]		[1.61]	[1.32]	
DEF	11.77	-2.17	4.63	10.47	-1.39	6.46	10.19	-0.22	8.77	6.22	1.95	6.24	4.15	2.61	5.04	3.49	3.26	6.11
	(3.16)***	(-0.37)		(3.50)***	(-0.27)		(4.05)***	(-0.05)		(2.99)***	(0.53)		(2.03)**	(0.90)		(1.82)*	(1.64)	
	[2.62]***	[-0.38]		[2.94]***	[-0.25]		[3.39]***	[-0.04]		[2.51]**	[0.40]		[1.88]*	[0.61]		[1.49]	[0.86]	
RREL	12.07	7.35	6.20	10.74	7.50	9.57	10.42	7.61	13.64	6.32	7.52	14.19	4.23	7.42	15.53	3.51	6.45	14.79
	(3.32)***	(1.73)*		(3.69)***	(2.03)**		(4.14)***	(2.19)**		(2.90)***	(2.09)**		(1.85)*	(1.83)*		(1.56)	(1.69)*	
	[2.67]***	[1.87]*		[3.01]***	[1.93]*		[3.47]***	[1.92]*		[2.53]**	[1.78]*		[1.90]*	[1.69]*		[1.49]	[1.49]	
SVAR	11.84	-8.86	6.98	10.51	-5.15	7.87	10.18	-4.93	10.82	6.13	-0.04	5.67	4.06	1.49	4.04	3.37	1.61	3.86
	(3.40)***	(-1.88)*		(3.68)***	(-1.88)*		(4.22)***	(-2.13)**		(2.90)***	(-0.02)		(1.89)*	(1.12)		(1.59)	(1.24)	
	[2.63]***	[-1.24]		[2.96]***	[-0.98]		[3.40]***	[-1.00]		[2.47]**	[-0.01]		[1.83]*	[0.42]		[1.43]	[0.56]	
TAIL	11.68	1.76	4.58	10.69	-1.61	6.50	10.37	-1.62	8.99	6.29	-1.29	5.92	3.91	1.10	3.82	3.08	2.19	4.46
	(3.19)***	(0.46)		(3.63)***	(-0.59)		(4.09)***	(-0.65)		(2.84)***	(-0.53)		(1.89)*	(0.41)		(1.61)	(0.85)	
	[2.60]***	[0.40]		[3.01]***	[-0.43]		[3.51]***	[-0.48]		[2.67]***	[-0.45]		[1.90]*	[0.42]		[1.41]	[0.90]	

This table reports the in-sample results for bivariate predictive regressions of the CRSP value-weighted index excess return on the forward variance factor (FV) and each of the other predictors used in the study. The remainder of the forecasting variables are the implied variance (IV), implied skewness (IS), dividend-price ratio (d-p), dividend payout ratio (d-e), yield term spread (TERM), default spread (DEF), relative short-term risk-free rate (RREL), stock market variance (SVAR) and tail risk (TAIL). The sample period is 1996:01-2015:08. Reported coefficients indicate the percentage annualized excess return resulting from a one standard deviation increase in each predictor variable. Newey and West (1987) and Hodrick (1992) t-statistics with lag length equal to the forecasting horizon are reported in parentheses and square brackets respectively. ***, ** and * denote significance at the 1%, 5% and 10% levels.

Table 5: Forward skewness predictability controlling for other predictors

	FS	Z	R^2 (%)	FS	Z	R^2 (%)	FS	Z	R^2 (%)	FS	Z	R^2 (%)	FS	Z	R^2 (%)	FS	Z	R^2 (%)
Horizon	**1-month**			**2-month**			**3-month**			**6-month**			**9-month**			**12-month**		
IV	12.55	-0.01	5.02	12.99	1.95	9.81	13.40	2.12	15.37	8.62	4.37	13.68	5.85	3.98	10.53	4.49	3.29	8.40
	(3.21)***	(-0.00)		(4.94)***	(0.65)		(5.68)***	(0.87)		(6.87)***	(2.90)***		(4.23)***	(2.23)**		(3.81)***	(1.78)*	
	[2.22]**	[-0.00]		[2.88]***	[0.39]		[3.53]***	[0.44]		[3.48]***	[1.02]		[3.06]***	[1.15]		[2.59]**	[1.13]	
IS	12.66	-1.94	5.14	13.15	-3.94	10.47	13.51	-3.05	15.77	8.82	-5.90	16.04	6.02	-5.18	12.90	4.63	-4.19	10.26
	(3.13)***	(-0.54)		(4.83)***	(-1.32)		(5.46)***	(-1.25)		(5.72)***	(-5.54)***		(4.38)***	(-4.23)***		(3.93)***	(-3.21)***	
	[2.23]**	[-0.35]		[2.89]***	[-0.83]		[3.55]***	[-0.62]		[3.58]***	[-1.37]		[3.16]***	[-1.56]		[2.67]***	[-1.55]	
d-p	11.75	4.27	5.58	12.06	4.70	10.81	12.45	4.74	16.82	7.21	6.64	17.20	4.28	7.53	18.98	2.86	7.99	22.60
	(2.64)***	(0.84)		(4.06)***	(1.14)		(4.56)***	(1.28)		(5.83)***	(2.07)**		(3.61)***	(2.65)***		(4.31)***	(3.11)***	
	[1.98]**	[0.82]		[2.61]***	[0.94]		[3.17]***	[0.95]		[2.66]***	[1.39]		[2.10]**	[1.67]*		[1.62]	[1.86]*	
d-e	12.56	-0.14	5.02	12.86	0.63	9.61	13.20	1.23	15.12	8.18	2.62	11.83	5.41	2.69	8.65	4.07	2.82	7.56
	(3.15)***	(-0.03)		(4.79)***	(0.17)		(5.67)***	(0.38)		(6.80)***	(0.95)		(3.22)***	(1.13)		(2.89)***	(1.62)	
	[2.20]**	[-0.03]		[2.86]***	[0.13]		[3.44]***	[0.26]		[3.10]***	[0.63]		[2.64]***	[0.72]		[2.23]**	[0.82]	
TERM	12.52	1.18	5.06	12.91	1.29	9.68	13.31	1.44	15.16	8.42	2.18	11.52	5.64	2.99	9.03	4.29	4.13	10.14
	(3.16)***	(0.34)		(4.69)***	(0.42)		(5.61)***	(0.52)		(5.82)***	(0.84)		(3.25)***	(1.17)		(3.33)***	(1.70)*	
	[2.20]**	[0.34]		[2.87]***	[0.37]		[3.51]***	[0.41]		[3.34]***	[0.60]		[2.92]***	[0.84]		[2.45]**	[1.19]	
DEF	12.48	-2.27	5.18	12.89	-1.42	9.70	13.33	-0.24	14.99	8.53	1.95	11.38	5.79	2.61	8.57	4.47	3.24	8.30
	(3.16)***	(-0.42)		(4.51)***	(-0.35)		(5.63)***	(-0.07)		(6.31)***	(0.59)		(3.56)***	(0.94)		(3.71)***	(1.65)*	
	[2.20]**	[-0.40]		[2.86]***	[-0.25]		[3.51]***	[-0.04]		[3.43]***	[0.40]		[3.03]***	[0.61]		[2.58]**	[0.85]	
RREL	12.38	6.67	6.43	12.76	6.86	12.28	13.16	6.97	19.08	8.28	7.14	18.48	5.51	7.16	18.25	4.21	6.25	16.30
	(3.27)***	(1.60)		(4.52)***	(1.97)*		(5.67)***	(2.16)**		(4.30)***	(1.98)**		(2.43)**	(1.75)*		(2.23)**	(1.62)	
	[2.19]**	[1.72]*		[2.84]***	[1.77]*		[3.47]***	[1.77]*		[3.30]***	[1.70]*		[2.86]***	[1.63]		[2.42]**	[1.45]	
SVAR	11.01	-6.11	6.13	12.41	-2.07	9.82	12.90	-1.74	15.23	9.03	2.18	11.48	6.50	3.10	9.05	5.11	2.86	7.51
	(2.69)***	(-1.25)		(4.04)***	(-0.78)		(5.01)***	(-1.03)		(5.54)***	(1.27)		(3.96)***	(1.87)*		(3.88)***	(1.80)*	
	[2.02]**	[-0.87]		[2.69]***	[-0.38]		[3.39]***	[-0.35]		[3.88]***	[0.48]		[3.50]***	[0.87]		[2.95]***	[0.98]	
TAIL	12.42	2.24	5.18	13.00	-1.24	9.67	13.41	-1.29	15.13	8.54	-1.05	10.97	5.63	1.21	7.41	4.23	2.30	6.84
	(3.16)***	(0.61)		(4.83)***	(-0.51)		(5.75)***	(-0.64)		(5.83)***	(-0.47)		(3.03)***	(0.47)		(2.73)***	(0.92)	
	[2.19]**	[0.51]		[2.87]***	[-0.33]		[3.52]***	[-0.38]		[3.46]***	[-0.36]		[3.02]***	[0.45]		[2.55]**	[0.91]	

This table reports the in-sample results for bivariate predictive regressions of the CRSP value-weighted index excess return on the forward skewness factor (FS) and each of the other predictors used in the study. The remainder of the forecasting variables are the implied variance (IV), implied skewness (IS), dividend-price ratio (d-p), dividend payout ratio (d-e), yield term spread (TERM), default spread (DEF), relative short-term risk-free rate (RREL), stock market variance (SVAR) and tail risk (TAIL). The sample period is 1996:01-2015:08. Reported coefficients indicate the percentage annualized excess return resulting from a one standard deviation increase in each predictor variable. Newey and West (1987) and Hodrick (1992) t-statistics with lag length equal to the forecasting horizon are reported in parentheses and square brackets respectively. ***, ** and * denote significance at the 1%, 5% and 10% levels.

Table 6: Out-of-sample predictability

Horizon	1-month	2-month	3-month	6-month	9-month	12-month
Panel A: Forward Moments						
FV	1.93	1.93	2.30	0.45	-1.00	-4.05
	(3.70)**	(3.66)**	(4.34)**	(0.81)*	(-1.69)	(-6.46)
	[5.64]**	[7.28]**	[9.84]**	[6.17]**	[3.40]**	[0.25]
FS	0.72	3.06	4.08	3.13	3.52	-0.02
	(1.37)*	(5.87)**	(7.83)**	(5.75)**	(6.28)**	(-0.03)
	[2.42]**	[5.87]**	[8.71]**	[5.97]**	[5.89]**	[2.05]*
FV & FS	4.53	3.66	5.28	7.12	1.51	-3.07
	(8.93)**	(7.06)**	(10.25)**	(13.65)**	(2.64)**	(-4.95)
	[7.70]**	[7.51]**	[10.84]**	[12.68]**	[5.22]**	[0.99]
(FV & FS) vs FV	2.61	1.73	2.98	6.67	2.51	0.98
	(5.13)**	(3.34)**	(5.78)**	(12.79)**	(4.38)**	(1.58)**
	[3.12]**	[2.61]**	[4.58]**	[7.80]**	[2.65]**	[1.14]
Panel B: Alternative Predictors						
IV	-3.83	-7.33	-12.74	-1.61	-1.79	-3.47
	(-6.93)	(-12.70)	(-20.80)	(-2.82)	(-3.03)	(-5.57)
	[-1.38]	[-2.32]	[-5.07]	[1.85]*	[-0.27]	[-1.87]
IS	-5.62	-14.38	-23.79	0.80	2.21	-0.08
	(-10.00)	(-23.39)	(-35.36)	(1.43)*	(3.88)**	(-0.13)
	[-2.57]	[-1.87]	[-7.10]	[8.97]**	[5.25]**	[2.57]**
d-p	-0.33	-1.40	-2.64	-0.74	2.46	5.06
	(-0.62)	(-2.58)	(-4.73)	(-1.31)	(4.33)**	(8.85)**
	[1.43]	[1.65]*	[1.95]*	[6.94]**	[15.90]**	[26.67]**
d-e	-7.23	-19.34	-34.49	-80.88	-111.43	-111.18
	(-12.68)	(-30.14)	(-47.19)	(-79.59)	(-90.65)	(-87.39)
	[-0.82]	[-5.27]	[-10.22]	[-21.14]	[-22.52]	[-13.99]
TERM	-2.50	-6.45	-11.64	-31.13	-68.70	-78.96
	(-4.58)	(-11.27)	(-19.18)	(-42.26)	(-70.04)	(-73.24)
	[-0.77]	[-3.03]	[-5.51]	[-10.44]	[-16.52]	[-14.09]
DEF	-4.53	-10.53	-23.53	-64.34	-66.00	-34.73
	(-8.14)	(-17.72)	(-35.05)	(-69.69)	(-68.38)	(-42.79)
	[2.83]**	[4.47]**	[0.71]	[-9.80]	[-13.29]	[-11.87]
RREL	-0.55	-1.33	-3.39	-13.10	-43.96	-83.24
	(-1.03)	(-2.44)	(-6.04)	(-20.62)	(-52.52)	(-75.41)
	[0.45]	[0.88]	[0.76]	[-1.75]	[-9.11]	[-16.87]
SVAR	-4.96	-8.58	-12.09	-6.99	-6.34	-7.67
	(-8.88)	(-14.70)	(-19.84)	(-11.63)	(-10.25)	(-11.82)
	[2.38]**	[-3.07]	[-6.41]	[-4.17]	[-4.16]	[-4.43]
TAIL	-1.46	-3.01	-5.44	-9.58	-10.40	-19.22
	(-2.71)	(-5.43)	(-9.50)	(-15.57)	(-16.20)	(-26.77)
	[-0.47]	[-2.05]	[-3.50]	[-5.57]	[-3.90]	[-7.34]

This table reports the results of out-of-sample predictability for the CRSP value-weighted index excess return. The total sample period is 1996:01-2015:08 and the forecasting period begins in 2000:01. The forecasting variables are the two forward moments factors (FV and FS), implied variance (IV), implied skewness (IS), dividend-price ratio (d-p), dividend payout ratio (d-e), yield term spread (TERM), default spread (DEF), relative short-term risk-free rate (RREL), stock market variance (SVAR) and tail risk (TAIL). Panel A shows the results for the forward moments factors and Panel B shows the results for the alternative predictors. The last part of Panel A compares a model with both forward variance and forward skewness factors to a model with just the forward variance factor. All other models are compared to the historical average model. For each forecasting model, the first row reports the out-of-sample coefficient of determination, the second row (in parentheses) reports the MSE F-statistic of McCracken (2007) and the third row (in square brackets) reports the encompassing ENC-NEW test of Clark and McCracken (2001). ** and * denote significance at the 1%, 5% and 10% levels.

Table 7: Market-timing strategy

	Mean (%)	**St. Dev. (%)**	**Sharpe**	**ΔCER (%)**
HAV	1.25	12.16	0.10	
FV	3.94	12.12	0.32	2.70
FS	2.83	13.67	0.21	0.99
FV & FS	4.94	12.18	0.41	3.68
IV	-1.14	14.77	-0.08	-3.45
IS	-0.94	13.25	-0.07	-2.61
d-p	7.69	17.92	0.43	3.84
d-e	3.56	10.68	0.33	2.81
TERM	1.37	11.67	0.12	0.30
DEF	1.49	11.17	0.13	0.59
RREL	3.91	11.33	0.35	2.95
SVAR	1.27	15.35	0.08	-1.30
TAIL	0.50	11.71	0.04	-0.59

This table reports the results of a market-timing strategy based on the 1-month ahead out-of-sample predictability for the CRSP value-weighted index excess return. The total sample period is 1996:01-2015:08 and the forecasting period begins in 2000:01. The forecasting variables are the two forward moments factors (FV and FS), implied variance (IV), implied skewness (IS), dividend-price ratio (d-p), dividend payout ratio (d-e), yield term spread (TERM), default spread (DEF), relative short-term risk-free rate (RREL), stock market variance (SVAR) and tail risk (TAIL). Mean denotes the average return, St. Dev. denotes the standard deviation of returns, Sharpe stands for the Sharpe ratio and ΔCER is the certainty equivalent return in excess of the historical average (HAV) strategy. All measures of performance are in annualized terms.

Table 8: Excess variance predictability

Horizon	1-month	2-month	3-month	6-month	9-month	12-month
Panel A: In-sample predictability						
FV	1.20	1.08	1.00	0.73	0.50	0.35
	(2.48)**	(3.09)***	(3.26)***	(2.36)**	(1.62)	(1.21)
	[2.15]**	[3.35]***	[4.24]***	[5.22]***	[4.72]***	[3.27]***
$\mathbf{R^2}$ **(%)**	8.39	6.83	6.07	3.75	1.94	1.01
FS	0.93	0.80	0.82	0.76	0.58	0.50
	(1.99)**	(2.02)**	(2.19)**	(2.41)**	(1.76)*	(1.58)
	[1.77]*	[2.01]**	[2.84]***	[5.85]***	[6.46]***	[6.35]***
$\mathbf{R^2}$ **(%)**	5.00	3.77	4.09	4.06	2.62	2.14
Panel B: Out-of-sample predictability						
FV	5.55	6.18	6.12	3.25	0.98	-0.18
	(11.04)**	(12.26)**	(11.99)**	(5.98)**	(1.71)**	(-0.30)
	[10.55]**	[10.58]**	[10.25]**	[5.92]**	[2.86]**	[1.25]
FS	3.12	4.74	5.85	3.19	2.52	1.64
	(6.05)**	(9.26)**	(11.44)**	(5.86)**	(4.45)**	(2.77)**
	[5.06]**	[6.37]**	[7.79]**	[4.60]**	[3.50]**	[2.18]*

This table reports the results for univariate predictive regressions of the excess variance on the forward variance factor (FV) and forward skewness factor (FS). Panel A shows in-sample predictability results and Panel B shows out-of-sample predictability results. The sample period is 1996:01-2015:08. In Panel A, reported coefficients indicate the percentage annualized excess return resulting from a one standard deviation increase in each predictor variable. Newey and West (1987) and Hodrick (1992) t-statistics with lag length equal to the forecasting horizon are reported in parentheses and square brackets respectively. In Panel B, the first row reports the out-of-sample coefficient of determination, the second row (in parentheses) reports the MSE F-statistic of McCracken (2007) and the third row (in square brackets) reports the encompassing ENC-NEW test of Clark and McCracken (2001). ** and * denote significance at the 1%, 5% and 10% levels.