Expected Correlation and Future Market Returns*

Adrian Buss **Lorenzo Schönleber** **Grigory Vilkov**

This version: December 18, 2018

Abstract

We document that information about the *comovement of individual stocks*, jointly extracted from index options and individual stock options, can be used to predict future *market* excess returns for horizons of up to 1 year, both in-sample and out-of-sample. The predictive power is incremental to that of risk measures exclusively based on the marginal distribution of the market, including (semi)variances and their risk premiums. We attribute this predictability to the ability of expected correlation to capture expected variations in idiosyncratic risk and in the cross-sectional dispersion in systematic risk. A novel extension of the contemporaneous-beta approach significantly improves out-of-sample predictability.

Keywords: expected (implied) correlation, correlation risk premium, return predictability, idiosyncratic risk, option-implied information, contemporaneous betas

JEL: G11, G12, G13, G17

*Adrian Buss is affiliated with INSEAD and CEPR, `adrian.buss@insead.edu`; Lorenzo Schönleber and Grigory Vilkov are affiliated with Frankfurt School of Finance & Management, `l.schoenleber@fs.de` and `vilkov@vilkov.net`. We received helpful comments and suggestions from Turan Bali, Bruno Biais, Nicole Branger (discussant), Agostino Capponi (discussant), Peter Carr, Mathijs Cosemans, Bernard Dumas, Thierry Foucault, Amit Goyal (discussant), Leonid Kogan, Petter Kolm, Hugues Langlois, Philippe Müller, Viktor Todorov, Igor Pozdeev (discussant), Fabio Trojani, Andrea Vedolin, and Hao Zhou. We also thank participants of the China International Conference in Finance 2017, the 2nd Annual Eastern Conference on Mathematical Finance (ECMF) 2017, the BofAML Quant and Risk Premia Conference 2017, the Frontiers of Factor Investing Conference 2018, the FMA Europe 2018, and the First Asset Pricing conference by LTI at the Collegio Carlo Alberto and seminar participants at Erasmus University (Rotterdam), ESSEC Business School, Frankfurt School of Finance & Management, HEC Paris, INSEAD, UNSW Sydney, the University of Sydney, the University of Technology Sydney, the University of Balearic Islands, and the University of Maastricht for useful comments. The authors gratefully acknowledge financial support from the Canadian Derivatives Institute (CDI).

Option prices, by construction, reflect investors' expectations about future price movements, and, hence, measures of risk extracted from index options are natural candidates for predicting future market returns. However, empirical evidence also suggests that information about the *comovement of individual stocks*, jointly extracted from index options and the cross-section of individual stock options, can be used to predict future *market* returns.[1]

The objective of this paper is to improve our understanding of the predictive power and the information content of options-based measures of stock comovement and to study the underlying economic sources explaining their return predictability. In particular, we carefully analyze the implied correlation and the correlation risk premium, and we contrast their ability to predict future market excess returns with that of variables exclusively based on the marginal distribution of the aggregate stock market (i.e., inferred solely from index options).

We find that both the implied correlation and the correlation risk premium significantly predict future market excess returns for horizons of up to 1 year and that this predictability mostly can be attributed to the forward-looking information encapsulated in option prices. Moreover, their information content is incremental to that of variables exclusively extracted from index options known to forecast future market returns.[2] Even out-of-sample, we document a significant predictive power for future market excess returns for horizons of up to 1 year; in particular so when using our novel extension of the contemporaneous-beta approach that combines high-frequency increments in option-implied variables with the respective risk premiums.

This strong and incremental return predictability is rather surprising. The implied correlation reflects the relative pricing of individual and index options and, hence, is a function of individual stocks' expected variances and the expected market variance. However, there exists hardly any empirical evidence supporting market excess return predictability by expected market variance.[3] Moreover, because the predictive power of implied correlation persists even after

[1]See, for example, the empirical evidence in Driessen, Maenhout, and Vilkov (2005, 2012), Cosemans (2011), and Faria, Kosowski, and Wang (2016).

[2]For instance, Bollerslev, Tauchen, and Zhou (2009), Drechsler and Yaron (2011), and Bollerslev, Marrone, Xu, and Zhou (2014) document market return predictability by the variance risk premium. Similarly, Feunou, Jahan-Parvar, and Okou (2017) and Kilic and Shaliastovich (2017) show that the downside ("bad") semivariance risk premium predicts future market excess returns.

[3]See Goyal and Welch (2008) and Bollerslev, Tauchen, and Zhou (2009), among others. An exception is the work by Ghysels, Santa-Clara, and Valkanov (2005), who document short horizon (1-month) market return predictability using a mixed data sampling approach.

controlling for index-options-based variables, the implied correlation does not simply act as a proxy for these variables. Hence, the predictability must be related to information contained in individual stocks' expected variances.

In particular, in contrast to measures of risk exclusively based on the marginal distribution of the market, the expected correlation is intimately linked to idiosyncratic risk and the cross-sectional dispersion in systematic risk. Consequently, we hypothesize that temporal variations in the amount of (priced) idiosyncratic risk and the cross-sectional dispersion in systematic risk are largely responsible for the incremental return predictability by the expected correlation. That is, if idiosyncratic risk is priced or the equity risk premium is affected by the cross-sectional dispersion in systematic risk,[4] the implied correlation will be able to capture fluctuations in future market excess returns resulting from variations in these variables, whereas purely market-based measures will not.

We provide substantial empirical evidence in favor of this hypothesis. First, we carefully document the in-sample predictive power of the implied correlation. In univariate regressions, we find that the implied correlation, extracted from options data, can predict future market excess returns for horizons of up to 1 year. Its regression coefficient is always highly significantly positive (with t-statistics consistently above 2), and its predictive power, measured in terms of R^2, is increasing from about 3.5% at the monthly horizon to around 5%–6% for 9- and 12-month horizons.[5] Indeed, excluding short horizons of up to one quarter, the implied correlation delivers the strongest predictive power; dominating index-based variables, such as the implied (semi)variances and their risk premiums. Moreover, even after controlling for these index-based variables in multivariate regressions, the implied correlation significantly predicts future market excess returns, with a consistently positive sign. We also relate the predictive power of the implied correlation to the forward-looking information encapsulated in option prices; that is, the predictive power of the historical realized correlation is much weaker.

[4]We discuss the empirical evidence and potential theoretical explanations for the pricing of idiosyncratic risk and the cross-sectional dispersion in systematic risk in the literature review.

[5]The R^2s are corrected for persistence in the predictors using the method suggested in Boudoukh, Richardson, and Whitelaw (2008). In the absence of the correction, the explained variability goes beyond 20% for an annual horizon.

We then provide a conceptual framework for interpreting the documented return predictability by implied correlation. Using a stylized model for individual stock returns, we study the determinants of the average expected correlation and illustrate its relation to idiosyncratic risk and the cross-sectional dispersion in systematic risk. Variations in the expected correlation can be summarized in terms of (1) the amount of aggregate risk, (2) the amount of idiosyncratic risk, and (3) the cross-sectional dispersion in systematic risk. Intuitively, an increase in aggregate risk implies stronger comovement among the individual stocks, because it strengthens the importance of aggregate shocks. Conversely, when idiosyncratic risk increases, the expected correlation declines, because stock-specific shocks are, to a larger extent, responsible for fluctuations in individual stock returns. Finally, we show that the expected correlation falls as the cross-sectional dispersion in systematic risk rises. We then empirically study the relation between the implied correlation and measures of the amount of future idiosyncratic risk and the future cross-sectional dispersion in systematic risk. Consistent with our stylized framework, we find that the implied correlation negatively predicts future idiosyncratic volatility and the future cross-sectional dispersion in market betas. Importantly, the empirical evidence indicates that these predictions are distinctly different from those of the implied (semi)variances.

In the next step, we present a variety of additional direct and indirect evidence for market excess return predictability by implied correlation. First, we document that, consistent with the intuition from the present-value relation (see, e.g., Campbell and Shiller (1988) and Campbell (1991)), revisions in expected market returns, as predicted by the implied correlation, are negatively correlated with contemporaneous market excess returns. Second, using a vector autoregression (VAR) model similar to that of Campbell (1991), we show that the implied correlation predicts future market excess returns, even if controlling for "traditional" variables, like dividend growth. This analysis also shows that, consistent with economic intuition, past dividend growth negatively affects the implied correlation; that is, the implied correlation increases in adverse economic conditions. Moreover, shocks to implied correlation propagate to some of the index-based variables. Using the same model, we find that, within the set of options-based variables, the implied correlation explains the largest fraction of the future market return variance. Third, we demonstrate that the correlation risk premium, defined as the difference between the expected correlation under the risk-neutral and the objective probability measures, can also significantly predict future market excess returns.

Lastly, we study out-of-sample return predictability. The implied correlation, the downside semivariance risk premium, and the correlation risk premium show similar out-of-sample predictability for short horizons of up to one quarter, but only the implied correlation can predict future market excess returns at longer horizons of up to 1 year. However, we also find that traditional out-of-sample regression techniques cannot fully exploit the predictive power of many variables, because they require a long historical estimation window for the regression coefficients. For example, in our applications, the out-of-sample predictability evaporates when the estimation window is shortened to 5 years or less. In lieu of this evidence, we propose a novel out-of-sample predictability approach that extends the contemporaneous-beta approach (see, e.g., Cutler, Poterba, and Summers (1989) and Roll (1988)). This approach combines high-frequency increments in option-implied variables with the variables' risk premiums to predict future market excess returns. For predictions based on correlation and downside semivariance risk, the approach leads to out-of-sample R^2s of around 8% for horizons of 3 to 6 months and of about 7% for 12 months. Consistent with our initial motivation, we link the better performance of our new approach to its stable and up-to-date regression coefficients.

Our work is related to several strands of the literature. First, it is related to the literature on market return predictability using information extracted from option markets.[6] Bollerslev, Tauchen, and Zhou (2009), Drechsler and Yaron (2011), and Bollerslev, Marrone, Xu, and Zhou (2014) show that the variance risk premium is a robust predictor of market returns for horizons of up to one quarter. Longer-term predictability by various components of the variance risk premium is presented in Fan, Xiao, and Zhou (2018), who rely on a decomposition into a premium for variance risk and one for higher-order risks, in Feunou, Jahan-Parvar, and Okou (2017) and Kilic and Shaliastovich (2017), who use decompositions into upside (good) and downside (bad) semivariance risk premiums, and in Bollerslev, Todorov, and Xu (2015), who focus on jump components. These papers exclusively focus on information extracted from the marginal distribution of the market, whereas we highlight the importance of information extracted from

[6] For an extensive literature survey, confer with Christoffersen, Jacobs, and Chang (2013). Goyal and Welch (2008) carefully analyze and discuss market return predictability more generally.

the joint distribution of the market and individual stocks. In particular, we demonstrate that information from the cross-section of individual stock options, encapsulated in the implied correlation, leads to incremental market return predictability. In that regard, we are related to the extant literature studying market return predictability by correlation-related variables, such as Driessen, Maenhout, and Vilkov (2005, 2012), Pollet and Wilson (2010), Cosemans (2011), and Faria, Kosowski, and Wang (2016). Unlike these papers, we study not only in-sample market return predictability but also out-of-sample return predictability. In addition, we carefully analyze the economic rationale behind the predictive power of the implied correlation, tracing it back to idiosyncratic risk and the cross-sectional dispersion in systematic risk. Methodologically, we also contribute to this literature by proposing a novel predictability technique that improves out-of-sample return forecasts, particularly for option-implied variables. Finally, our work is related to Martin and Wagner (Forthcoming), who document return predictability for individual stock returns using a combination of expected market variance and a stock's expected ("excess") variance, both of which are extracted from option prices.

Second, our work is related to the literature studying correlation risk. Buraschi, Porchia, and Trojani (2010) study optimal portfolios in the presence of correlation risk. Leippold and Trojani (2010) propose a multivariate framework in which stochastic volatilities, stochastic correlations, and jumps can be consistently modeled. In Buraschi, Trojani, and Vedolin (2014), correlation risk is endogenously priced due to differences-in-beliefs. Relatedly, in Piatti (2015), a correlation risk premium arises from agents' disagreement about the likelihood of systematic disasters. Empirically, Buraschi, Kosowski, and Trojani (2014) relate correlation risk to a "no-place-to-hide" state variable, Chang, Christoffersen, Jacobs, and Vainberg (2012) and Buss and Vilkov (2012) use option-implied correlations to measure systematic risk, and Mueller, Stathopoulos, and Vedolin (2017) study correlation risk in foreign exchange markets. Our contribution to this literature is twofold. First, we carefully document that variations in the implied correlation and the correlation risk premium can be used to predict future market excess returns and that they are not redundant relative to market variance components. Second, we empirically link variations in implied correlation to future variations in realized idiosyncratic risk and in the realized dispersion in market betas.

Accordingly, our analysis also relates to the literature on idiosyncratic risk. Ang, Hodrick, Xing, and Zhang (2006, 2009) document a negative relation between idiosyncratic risk and the cross-section of stock returns. Schneider, Wagner, and Zechner (2017) relate the pricing of idiosyncratic risk to (implied) skewness. Goyal and Santa-Clara (2003) document that average variance (their proxy for idiosyncratic risk) positively predicts future market returns, whereas Bali, Cakici, Yan, and Zhang (2005) cannot find a significant relationship between idiosyncratic volatility and future market returns. Guo and Savickas (2006) show that, when combined with aggregate stock market volatility, the value-weighted idiosyncratic volatility is significantly negatively related to future market returns. We contribute to this literature by linking the predictive power of the implied correlation to idiosyncratic risk. In particular, because the implied correlation positively predicts future market returns but is negatively related to idiosyncratic volatility, our evidence is suggestive of a negative price for idiosyncratic risk. Consistent with our conceptual framework for the expected correlation and our empirical findings, Kogan and Papanikolaou (2012, 2013) demonstrate that the idiosyncratic volatility and the dispersion in market betas are positively related to growth opportunities, which, in turn, are negatively related to the equity risk premium. Similarly, in the model of Santos and Veronesi (2004), the equity risk premium is low when the dispersion in systematic risk is high.

The rest of this paper is organized as follows: Section 1 documents the in-sample predictive power of the implied correlation. Section 2 discusses the determinants of the expected correlation, and Section 3 presents a variety of additional evidence for market return predictability by implied correlation. Finally, Section 4 studies out-of-sample predictability and introduces a novel forecasting approach based on contemporaneous betas. Section 5 concludes.

1 Implied Correlation: In-Sample Return Predictability

In this section, we first discuss the construction of the expected correlation using options data only. After a brief description of the data, we then document the ability of the implied correlation and other option-implied variables to predict future market excess returns in-sample.

1.1 Implied Correlation

Computing the historical pairwise correlation among any two stocks is rather easy; however, computing an expected pairwise correlation from options data is, in practice, not possible. Doing so would require not only options on the two individual stocks but also options on a basket comprising the two stocks. However, such options are not traded in financial markets. Instead, we construct the *implied correlation* as the average expected correlation among the stocks within an index.

To identify the average expected correlation among individual stocks, we rely on the observation that the index variance can be computed in two ways: (1) directly from the index and (2) indirectly through the portfolio of its constituents. Formally, this yields the restriction that the time-t-expected variance of the index, $\sigma^2_{I,t}$, must be equal to the expected variance of the portfolio of its constituents:

$$\sigma^2_{I,t} \triangleq \sum_{i=1}^{N} \sum_{i'=1}^{N} w_{i,t}\, w_{i',t}\, \sigma_{i,t}\, \sigma_{i',t}\, \rho_{i,i',t},$$

where N denotes the number of stocks in the index, $\sigma^2_{i,t}$ denotes the expected return variance of stock $i \in \{1, \ldots, N\}$, $w_{i,t}$ denotes the stock's index weight, and $\rho_{i,i',t}$ denotes the correlation between the returns of stocks i and i'. Under the assumption that all pairwise correlations are the same ($\rho_{i,i',t} = \rho_{j,j',t}, \forall i, i', j, j'$), the "equi-correlation" $Corr_t$ can be expressed as

$$Corr_t = \frac{\sigma^2_{I,t} - \sum\limits_{i=1}^{N} w^2_{i,t}\, \sigma^2_{i,t}}{\sum\limits_{i=1}^{N} \sum_{i' \neq i} w_{i,t}\, w_{i',t}\, \sigma_{i,t}\, \sigma_{i',t}}. \tag{1}$$

The implied variances for the index and individual stocks can be directly obtained from index and stock options. Hence, the average expected correlation (1) under the risk-neutral measure can be easily computed from index options and the cross-section of individual stock options. We refer to it simply as the implied correlation, denoted by IC.

Intuitively, the implied correlation is driven by the price of index options relative to that of individual stock options. All else equal, an increase in the price of index options implies an increase in the implied index variance, which, in turn, raises the implied correlation (1). Conversely, an increase in the prices of individual stock options implies an increase in the implied stock variances and, hence, leads to a lower implied correlation.

1.2 Data Description

Our analysis focuses on the S&P 500 for a sample period from January 1996 to December 2017. We use daily data on index and individual stock options from the Surface File of Ivy DB OptionMetrics and data on the daily realized returns from CRSP.[7] For the S&P 500, we also obtain intraday returns from TickData. Data on the S&P 500 index composition and the stocks' index weights (i.e., relative market capitalizations) are obtained from Compustat and CRSP.[8]

To compute the day-t implied variances (i.e., the expected variance under the risk-neutral measure) for options with maturity T, $IV(t,T)$, we rely on simple variance swaps, like in Martin (2013, 2017).[9] Implied correlation, $IC(t,T)$ is then constructed directly from the implied variances of the index and the individual stocks, using expression (1). To build the risk-neutral expectations of upside and downside semivariances, $IV^u(t,T)$ and $IV^d(t,T)$, we follow the corridor variance methodology of Andersen and Bondarenko (2007) and Andersen, Bondarenko, and Gonzalez-Perez (2015). Under the physical measure, we construct realized variance and semivariances following Andersen, Bollerslev, Diebold, and Ebens (2001) and Feunou, Jahan-Parvar, and Okou (2017), respectively.[10] The *ex-ante* variance risk premium, $VRP(t,T)$, is computed as the difference between the day-t implied variance from options with maturity T and the realized

[7] We select options with 1 to 12 months to maturity and an (absolute) delta of less than or equal to 0.5. On average, options data are available for 491 of the 500 index constituents.

[8] We merge the two datasets through the CCM Linking Table using GVKEY and IID to link to PERMNO. The matching to options data is implemented through the historical CUSIP link, provided by OptionMetrics.

[9] Computing expected variances using log contracts, that is, model-free implied variances, like in Dumas (1995), Britten-Jones and Neuberger (2000), and Bakshi, Kapadia, and Madan (2003), does not affect our results. These results are available on request.

[10] Realized variance $RV(t,T)$ is computed as the sum of squared returns, whereas upside (downside) realized semivariance $RV^u(t,T)$ $(RV^d(t,T))$ is computed as the sum of the squared positive (negative) returns, plus, if positive (negative), the squared overnight return. For individual stocks, we rely on demeaned daily returns, whereas, for the S&P 500, we use 5-minute intraday returns (as suggested by Liu, Patton, and Sheppard (2015)).

	1			3			6			9			12		
	Mean	*Std*	*Skew*	*Mean*	*Std*	*Skew*	*Mean*	*Std*	*Skew*	*Mean*	*Std*	*Skew*	*Mean*	*Std*	*Skew*
IC	0.378	0.129	0.546	0.417	0.114	0.164	0.444	0.105	-0.108	0.453	0.102	-0.170	0.459	0.097	-0.201
$\sqrt{IV}$	0.204	0.077	1.823	0.205	0.066	1.501	0.207	0.057	1.267	0.209	0.053	1.162	0.211	0.050	1.085
$\sqrt{IV^u}$	0.130	0.047	1.563	0.129	0.040	1.141	0.131	0.034	0.852	0.132	0.032	0.742	0.133	0.031	0.674
$\sqrt{IV^d}$	0.164	0.067	2.130	0.169	0.061	1.837	0.172	0.056	1.563	0.174	0.055	1.381	0.175	0.055	1.254
VRP	0.006	0.028	-7.706	0.006	0.027	-5.737	0.008	0.020	-3.236	0.009	0.023	-3.509	0.009	0.025	-2.782
VRP^u	-0.001	0.017	-8.265	-0.001	0.016	-6.162	0.000	0.012	-4.159	0.000	0.013	-3.818	-0.000	0.014	-2.974
VRP^d	0.009	0.013	-0.653	0.011	0.013	0.204	0.012	0.012	2.347	0.012	0.014	1.493	0.013	0.015	1.486

Table 1: Summary statistics. This table reports summary statistics for the implied correlation (IC), the implied volatility ($\sqrt{IV}$), and the upside and downside semivolatilities ($\sqrt{IV^u}$ and $\sqrt{IV^d}$) and the risk premiums for the (semi)variances (VRP, VRP^u, and VRP^d). Statistics are reported for maturities of 1 to 12 months over a sample period from January 1996 to December 2017 and, where applicable, in annual terms.

variance for the period from $t - (T - t) + 1$ to t. The same procedure is used to obtain upside and downside semivariance risk premiums, $VRP^u(t, T)$ and $VRP^d(t, T)$.

Table 1 reports summary statistics for these variables. Focusing on the 1-month maturity, the average implied correlation is about 0.38. It is highly persistent (with a first-order autocorrelation of about 0.77)[11] but still considerably varies over time, with fluctuations being almost symmetric around the mean. The annualized S&P 500 option-implied index volatility is 20.4%, a majority of which can be attributed to downside semivariance (in volatility terms, 16.4% relative to 13.0% for upside semivariance).[12] The implied (semi)variances are also highly persistent (with first-order autocorrelations between 0.78 and 0.82), substantially fluctuate over time, and generally have positive skewness, with the largest skewness for implied downside semivariance. The S&P 500 variance risk premium and downside semivariance risk premium are both significantly positive, whereas the risk premium for upside semivariance is not statistically different from zero. All (semi)variance risk premiums are considerably less persistent (with the first-order autocorrelations between 0.27 and 0.48) and negatively skewed. Table 1 also illustrates that the implied correlation considerably increases with option maturity, from 0.38 at a 1-month maturity to 0.46 at a 12-month maturity. The implied (semi)variances are also monotonically increasing with option maturity, though the magnitudes are smaller. Similarly, the (semi)variance risk premiums rise as option maturity increases.

[11] Tables A1 and A2 in Appendix C report the autocorrelations and the time-series correlations across variables.

[12] Upside and downside semivariances sum to the total variance if computed using the same methodology. However, because we compute implied variances from simple variance swaps, whereas semivariances are based on log contracts, the sum of the two semivariances is slightly different from total variance, with a value of 20.93% in volatility terms.

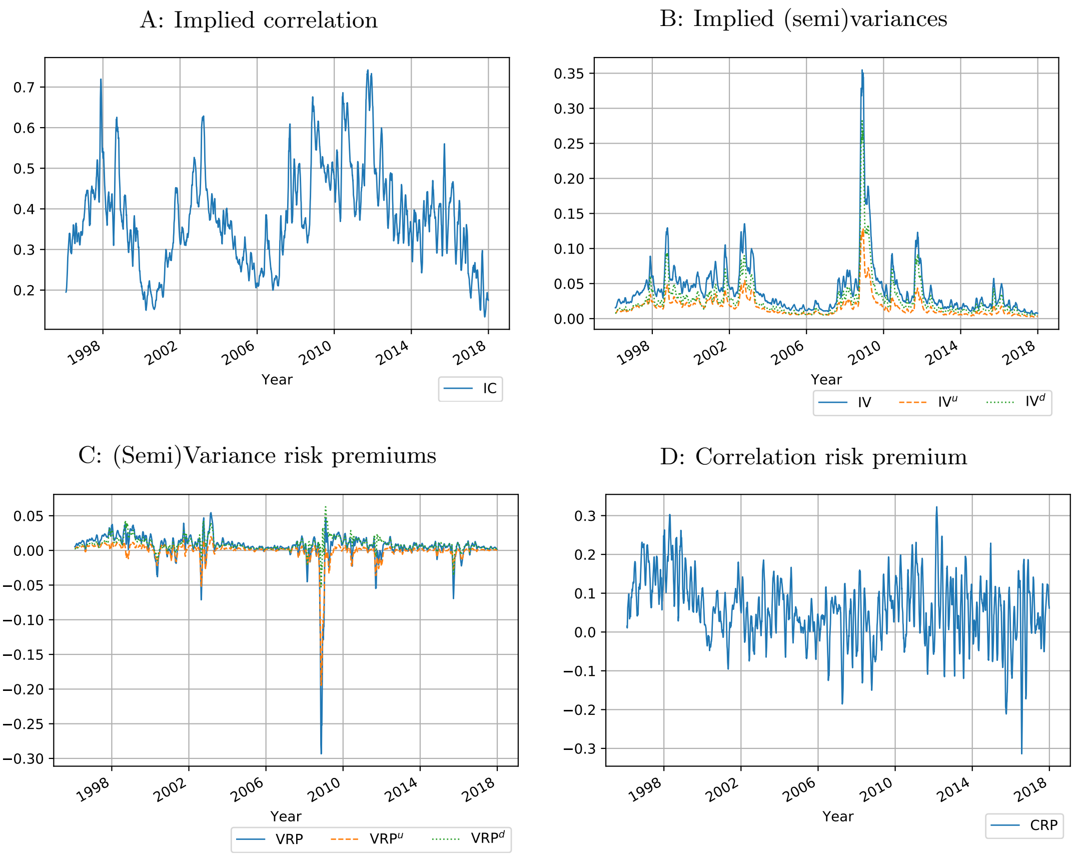


Figure 1: Time-series dynamics of option-implied variables. The figure shows the time series of the implied correlation (Panel A), the time series of the implied variance (IV) and the upside and downside implied semivariances (Panel B), the risk premiums for variance and semivariances (Panel C), and the correlation risk premium (Panel D). The sample period spans from January 1996 to December 2017, with variables being computed from 1-month options and sampled daily.

Panels A and B of Figure 1 depict the time-series dynamics of implied correlation, implied (semi)variances, and their risk premiums, computed from options with a 1-month maturity. The implied correlation, being a bounded variable, displays smaller fluctuations than do the implied (semi)variances. Moreover, the most pronounced spikes in the implied correlation do not necessarily coincide with those in the implied (semi)variances (which all behave in a very similar way).[13]. Consistent with this, the time-series correlations between the implied correlation and the implied (semi)variances are not that high; for example, the correlation with the implied variance is 0.58. Similarly, whereas the risk premiums on the (semi)variances display very similar time-series behaviors (Panel C), the fluctuations in the implied correlation are distinctly differ-

[13]The time-series correlations among the implied (semi)variances are all above 0.97 (see Table A2 in Appendix C)

ent, with time-series correlations between -0.15 (between IC and VRP^u) and 0.33 (between IC and VRP^d).

1.3 Return Predictability

To study the in-sample predictive power of the various variables, we rely on standard in-sample predictive regressions of the following form:

$$r_{t\to t+\tau_r} = \gamma + \sum_{k\in K} \beta_k \, PRED_k(t, t+\tau_r) + \varepsilon_t, \tag{2}$$

where $r_{t\to t+\tau_r}$ denotes the market excess return for a period from t to $t+\tau_r$ and $PRED_k(t, t+\tau_r)$, $k \in \{1, \ldots, K\}$ denotes a set of predictors (known at time t).[14] We use market excess returns from the end of each month in our sample period, and, when predicting returns for horizons longer than 1 month, we use Newey and West (1987) standard errors to correct for autocorrelation introduced by overlapping observations.

Figure 2 (graphically) reports the results. Panel A shows that, in univariate regressions, the implied correlation has the strongest predictive power (in terms of R^2) for future market excess returns, except for the 1-month horizon. Its predictive power increases from about 3.5% at the 1-month horizon to around 25% for the 9- and 12-month horizons. Moreover, its regression coefficient is always highly significantly positive, with t-statistics consistently above 2 (Panel B). Comparing the return predictability of the implied correlation to that of past realized correlation highlights the importance of using information implied by option prices.[15] That is, the predictability by the realized correlation is considerably weaker at all horizons, and it is, if at all, only borderline significant. Hence, the forward-looking information encapsulated in the option prices of the index and the individual stocks is critical for the ability of the expected correlation to predict future market excess returns.

[14] In the regressions, we always rely on variables extracted from options with a maturity matching the forecasting horizon.

[15] To be consistent, we also compute the expected correlation under the objective probability measure, RC, from (1) using realized stock and index variances.

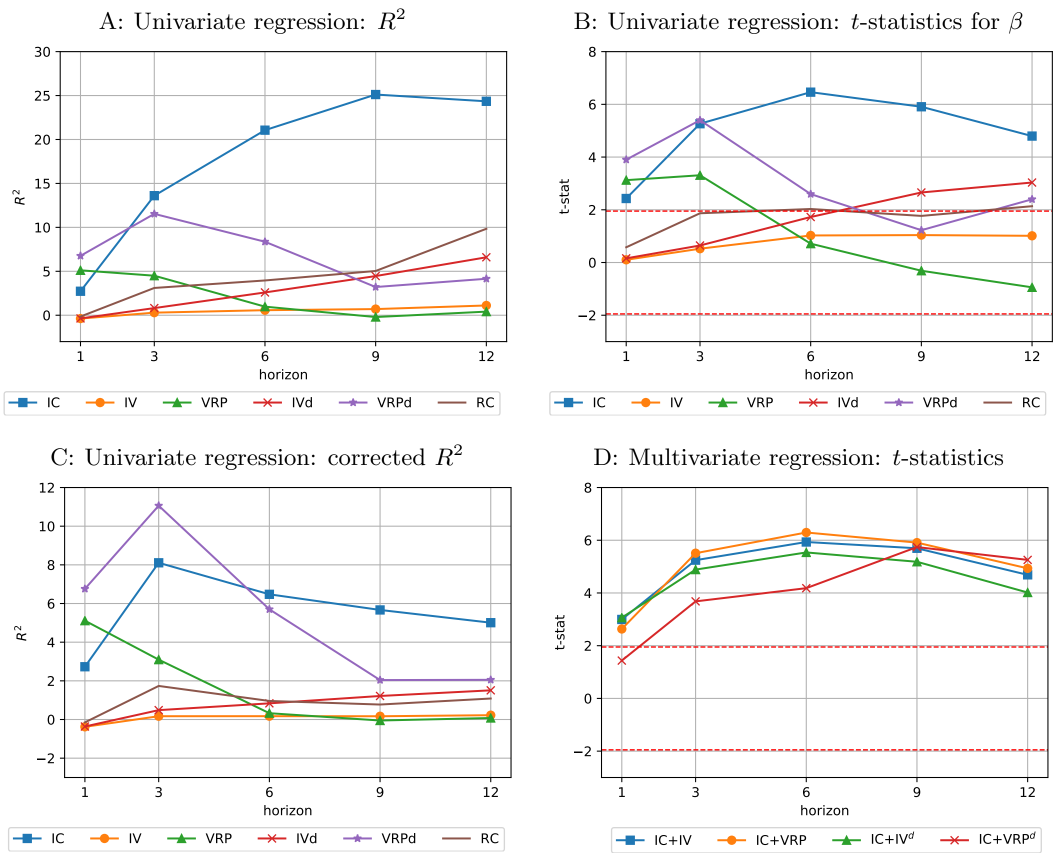


Figure 2: In-sample return predictability. The figure depicts the results of the in-sample market excess return predictability analysis for horizons from 1 to 12 months. Panels A and B show the adjusted R^2 and the regressors' t-statistics for univariate regressions of (2), respectively. Panel C shows the univariate R^2s corrected for the autocorrelation of the regressors. Finally, Panel D reports the t-statistic for the implied correlation in multivariate regressions. The standard errors are corrected for autocorrelation using Newey and West (1987), with the red dotted lines indicating 1.96 bounds around zero.

Panels A and B of Figure 2 also illustrate that other option-implied variables typically perform worse compared with the implied correlation in predicting future market excess returns. For example, the implied variance has no predictive power for future returns at any horizon, and the implied downside semivariance is only significant at the long horizons. Also, whereas the variance risk premium significantly predicts returns at short horizons with R^2s around 5%, its predictive power and statistical significance quickly decline as the horizon lengthens. The downside semivariance risk premium delivers a better predictability than does the total variance risk premium, with R^2s of 7%-11% at short horizons and around 5% at longer horizons (consistent with Feunou, Jahan-Parvar, and Okou (2017) and Kilic and Shaliastovich (2017)). Moreover, its regression coefficient is always significantly positive. Finally, similar to Kilic and

Shaliastovich (2017), the predictability by the upside semivariance risk premium is considerably weaker (data not shown).

Note that while the t-statistics are explicitly corrected for autocorrelation, one should treat the in-sample R^2s with caution. That is, as shown by Boudoukh, Richardson, and Whitelaw (2008), for highly persistent regressors, the in-sample R^2 mechanically increases with the predictability horizon. To understand the magnitude of the potential bias in our regressions, Panel C depicts the univariate R^2s adjusted for the autocorrelations of the predictors (using the procedure outlined in Boudoukh, Richardson, and Whitelaw (2008)). Notably, the implied correlation still delivers the strongest predictive power at longer horizons. To better understand the potential biases resulting from overlapping predictions, we have also performed additional tests based on the instrumental variable approach by Kostakis, Magdalinos, and Stamatogiannis (2015).[16] It turns out that, in most cases, our initial corrections for autocorrelation based on Newey-West standard errors are stricter than the ones from the instrumental variable approach. Consequently, the implied correlation and the down semivariance risk premium always remain significant.

Panel D reports the regression coefficients for the implied correlation (β_{IC}) in the multivariate versions of the predictive regression (2). As is apparent, the implied correlation remains a significant predictor of future market excess returns even after controlling for the index-based variables. In contrast, when used jointly with the implied correlation, the downside semivariance risk premium loses significance at the 6-month horizon, and, contrary to economic intuition, its regression coefficient becomes significantly negative for 9 and 12 months (data not shown).

In summary, in-sample, the implied correlation is a robust predictor for future market excess returns. In univariate regressions, it demonstrates a strong predictive power, dominating index-based variables, such as the implied (semi)variances and their risk premiums for almost all horizons. Moreover, even after controlling for these variables, the implied correlation significantly predicts future market excess returns, with a consistently positive sign.

[16] We are grateful to the authors for providing the code on their web-site.

2 Determinants of the Expected Correlation

To rationalize the predictive power of the implied correlation, we now present a stylized, conceptual framework that illustrates the determinants of the expected correlation. We will then empirically explore the implications of this framework.

2.1 Theoretical Framework

Our conceptual framework is based on that of Kelly, Lustig, and Van Nieuwerburgh (2016) and relies on a discrete-time, one-factor structure for stock returns. In particular, we consider a broad index (stock market) comprising a large number of stocks, denoted by $i \in \{1, \ldots, N\}$. The annual log-return of each individual stock, $r_{i,t+1}$, is described by the following dynamics:

$$r_{i,t+1} = \mu_{i,t} + \beta_{i,t}\, \varepsilon_{A,t+1} + \varepsilon_{i,I,t+1}, \tag{3}$$

where $\varepsilon_{A,t+1}$ denotes an aggregate shock common to all stocks, and $\varepsilon_{i,I,t+1}$ denotes an idiosyncratic (stock-specific) shock independent from the aggregate shock and from the idiosyncratic shocks of the other stocks. Both aggregate and idiosyncratic shocks might comprise several components, such as Gaussian shocks, jumps, or separate up and down jumps.[17] Importantly, the specific distribution of the aggregate and the idiosyncratic shocks is not important for our analysis; what is relevant is the total expected variance of each shock, denoted by $\sigma^2_{A,t} \equiv Var(\varepsilon_{A,t+1})$ and $\sigma^2_{I,t} \equiv Var(\varepsilon_{i,I,t+1})$. The exposures of the individual stocks to aggregate shocks ("market betas"), $\beta_{i,t}$, are assumed to be stochastic and distributed with time-varying variance $\sigma^2_{\beta,t}$ around the natural mean of 1.0.

[17]For instance, a Merton jump setting similar to that used by Kelly, Lustig, and Van Nieuwerburgh (2016) can be captured by decomposing the aggregate shock into a systematic Gaussian shock $\epsilon_{A,t+1}$ as well as a systematic jump $J_{A,t+1}$ and by decomposing the idiosyncratic shock into an idiosyncratic Gaussian shock $\epsilon_{i,I,t+1}$ as well as an idiosyncratic jump $J_{i,I,t+1}$. The stock return dynamics then would become

$$r_{i,t+1} = \mu_{i,t} + \beta_{i,t}\, \epsilon_{A,t+1} + \beta_{i,t}\, J_{A,t+1} + \epsilon_{i,I,t+1} + J_{i,I,t+1}.$$

Intuitively, the exposure with respect to the aggregate jump, $\beta_{i,t}$, captures the stronger reaction of the high beta stocks to systematic jumps and is equivalent to "scaling" the mean and the volatility of their jump size distributions.

The total expected return variance of stock i, $\sigma^2_{i,t} \equiv Var(r_{i,t+1})$, is equal to

$$\sigma^2_{i,t} = \beta^2_{i,t}\,\sigma^2_{A,t} + \sigma^2_{I,t}, \tag{4}$$

and its expected return is $\mu_{i,t}$, which is the sum of the expected return on the index and risk adjustments.

The log return on the index is given by $r_{t+1} \equiv \sum_{i=1}^{N} \omega_{i,t}\, r_{i,t+1}$, where $\omega_{i,t}$ denotes the index weight of stock i. As the number of stocks becomes large (i.e., $N \to \infty$), the idiosyncratic shocks are diversified. Moreover, the weighted average of the market betas is, by construction, 1.0. Hence, the conditional expected variance of the index return is equal to $\sigma^2_{A,t}$.

Within this stylized framework, we can now study the determinants of the expected correlation. For ease of exposition, we thereby focus on the expected correlation under the objective probability measure. As it follows from expression (4), the cross-sectional average expected stock variance is given by $(1+\sigma^2_{\beta,t})\sigma^2_{A,t} + \sigma^2_{I,t}$. Consequently, within our setting, the key determinants of the expected correlation are expected aggregate variance $\sigma^2_{A,t}$, expected idiosyncratic variance $\sigma^2_{I,t}$, and the cross-sectional dispersion in market betas $\sigma^2_{\beta,t}$.[18,19] To study the impact of these three variables on the expected correlation, we rely on a comparative statics analysis. For each combination of the parameters, we generate $1,000$ simulations of an index consisting of 500 individual stocks with equal weights, "drawing" their 500 betas from a normal distribution[20] with a mean of 1.0 and a variance of σ^2_{β}. Consistent with our empirical analysis, we then compute the average expected correlation using (1).

Figure 3 presents the results of the comparative statics analysis. Panel A illustrates that the expected correlation is monotonically increasing in the expected aggregate variance σ^2_A. Intuitively, an increase in the expected aggregate variance leads to a relatively stronger increase in index variance than in the individual stock variances (where aggregate variance is mixed with idiosyncratic variance). Consequently, the expected correlation rises, as is apparent from (1).

[18] Overall, variations in the index weights over short periods of time, such as 1 month, are quite limited and, hence, do not contribute much to variations in the expected correlation.

[19] Under the risk-neutral probability measure, risk premiums could also potentially matter, for example, in the case of jumps, by altering the likelihood of jumps and the jump size distribution.

[20] The results are robust to the use of other distributions for the betas of the individual stocks.

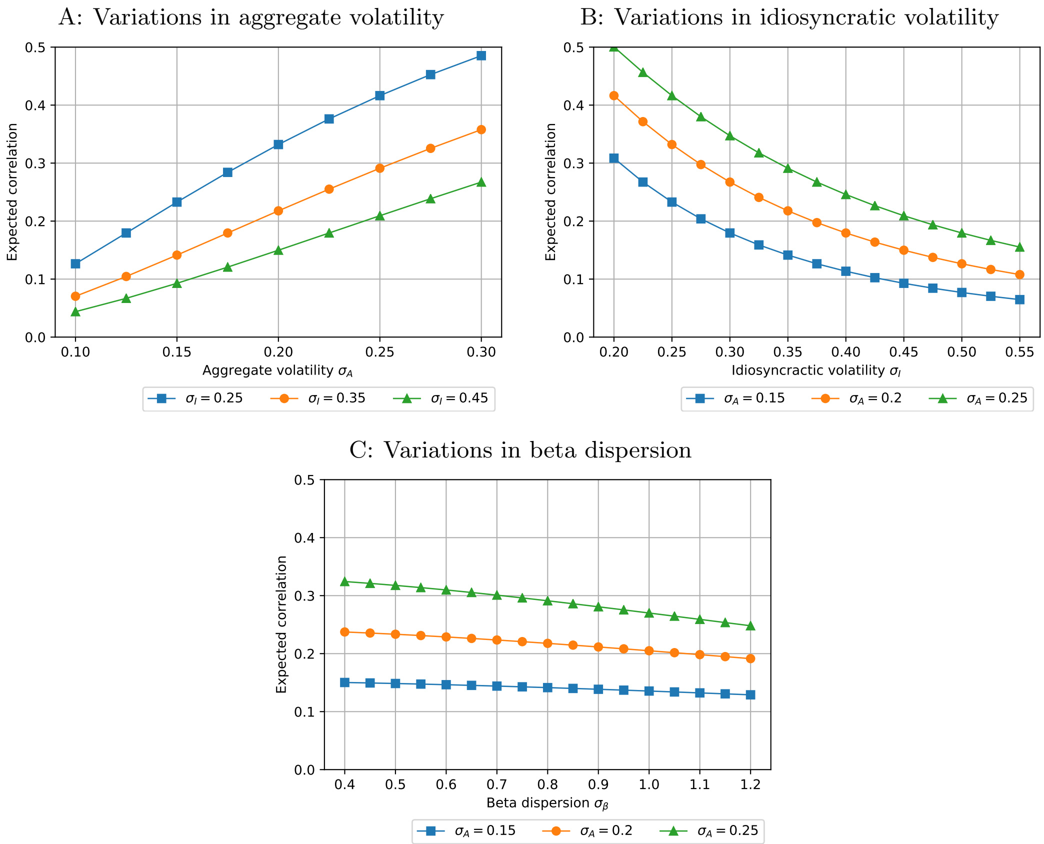


Figure 3: Determinants of the expected correlation. This figure illustrates the impact of the three key determinants of the expected correlation. Panel A shows the expected correlation as a function of expected aggregate volatility for various levels of expected idiosyncratic volatility and $\sigma_\beta = 0.8$. Panel B depicts the expected correlation as a function of expected idiosyncratic volatility for various levels of expected aggregate volatility and $\sigma_\beta = 0.8$. Panel C shows how the expected correlation varies with the cross-sectional dispersion in market betas for various levels of expected aggregate volatility and $\sigma_I = 0.3$. All graphs are based on an equal-weighted index with 500 stocks and 1,000 simulations.

For higher levels of expected idiosyncratic variance, the difference in the relative increase of index and individual stock variances is even more pronounced, and, hence, the expected correlation increases more strongly (in relative terms).

Conversely, the expected correlation is monotonically decreasing in the expected idiosyncratic variance, σ_I^2 (Panel B). Intuitively, an increase in the expected idiosyncratic variance leads to an increase in the expected variances of the individual stocks, whereas the expected variance of the index remains unchanged. Hence, the expected correlation falls. Moreover, for lower levels of expected aggregate variance, the relative increase in the individual stock variances is stronger, and, thus, the relative decline in the expected correlation is more pronounced.

Finally, as depicted in Panel C, the expected correlation is also decreasing in the cross-sectional dispersion in market betas σ_β^2. An increase in the cross-sectional beta dispersion implies an increase in the average squared beta of the individual stocks, which, in turn, increases the average stock return variance (4).[21] Because the expected index variance is unchanged, the expected correlation declines. For higher levels of expected aggregate variance, this effect is stronger, and, hence, the relative decline in the expected correlation strengthens.

Note that variations in expected aggregate variance—one of the key determinants of the expected correlation—(naturally) lead to changes in the expected (semi)variance of the index as well. However, most importantly, variations in the other two key determinants, namely, expected idiosyncratic variance and the cross-sectional beta dispersion, profoundly affect the expected correlation but leave expected index (semi)variances unchanged. Consequently, if idiosyncratic risk is priced or the equity risk premium is affected by the cross-sectional dispersion in market betas,[22] the expected correlation will be able to capture the resultant fluctuations in future market excess returns, whereas index-based variance measures will not. This could explain the incremental predictive power of the expected correlation for future market excess returns, as documented in the empirical analysis in Section 1.

For ease of exposition, we have made a variety of simplifying assumptions: (1) a one-factor structure for stock returns, (2) homogeneity in the variance of the idiosyncratic shocks, and (3)a limiting behavior in the return variance of the index. Robustness tests show that the main implications of our analysis are essentially unchanged when relaxing any of these assumptions.[23]

2.2 Empirical Analysis

For the implied correlation to predict future market excess returns through the channels outlined in the preceding subsection, the implied correlation must be able to predict idiosyncratic risk and the dispersion in market betas as *realized* in the future with the correct sign. Accordingly,

[21]A partially offsetting effect results from the fact that the *product* of the individual stock volatilities appears in the denominator of the expected correlation in (1). For example, in the two-stock case, the product of two stocks' betas $(1-\xi)(1+\xi)$ (which is part of the product of two stocks' expected volatilities) is decreasing in the "dispersion" ξ.

[22]Confer with the literature review for a discussion of the empirical and theoretical literature on the price of idiosyncratic risk and the impact of the cross-sectional dispersion in systematic risk.

[23]For brevity, we omit these experiments from the paper, but the results are available on request.

	1 month			3 months			6 months			9 months			12 months		
	β	p-val	R^2	β	p-val	R^2	β	p-val	R^2	β	p-val	R^2	β	p-val	R^2
A: Realized market variance															
IC	0.146	0.000	8.56	0.052	0.121	1.17	-0.028	0.326	0.38	-0.045	0.118	1.24	-0.031	0.242	0.62
IV	1.071	0.000	49.33	0.859	0.000	27.73	0.777	0.000	17.14	0.667	0.000	13.75	0.585	0.000	12.60
IV^u	2.870	0.000	49.39	2.417	0.000	28.79	2.195	0.000	18.34	1.927	0.000	15.37	1.788	0.000	15.33
IV^d	1.417	0.000	48.14	0.986	0.000	24.86	0.768	0.000	13.01	0.564	0.000	8.33	0.410	0.001	5.73
B: Idiosyncratic risk — market model															
IC	0.004	0.427	0.38	-0.025	0.304	1.26	-0.153	0.001	11.75	-0.26	0.00	15.69	-0.320	0.003	13.07
IV	0.138	0.000	51.39	0.455	0.000	38.38	1.024	0.000	29.11	1.49	0.00	25.89	1.870	0.000	24.49
IV^u	0.389	0.000	57.08	1.389	0.000	46.80	3.320	0.000	41.12	5.08	0.00	40.32	6.723	0.000	41.22
IV^d	0.173	0.000	45.14	0.480	0.000	29.03	0.831	0.000	14.89	0.91	0.00	8.17	0.820	0.034	4.36
C: Idiosyncratic risk – 4-factor model															
IC	0.003	0.461	0.32	-0.021	0.304	1.32	-0.136	0.001	12.40	-0.24	0.00	16.84	-0.294	0.003	14.07
IV	0.100	0.000	50.10	0.383	0.000	38.67	0.899	0.000	30.07	1.33	0.00	26.94	1.698	0.000	25.78
IVu	0.284	0.000	56.17	1.176	0.000	47.82	2.945	0.000	43.30	4.58	0.00	42.76	6.141	0.000	43.90
IVd	0.125	0.000	43.55	0.400	0.000	28.78	0.718	0.000	14.89	0.79	0.00	8.12	0.730	0.035	4.40
D: Dispersion of market betas – market model															
IC	-0.961	0.000	14.24	-0.499	0.000	12.19	-0.663	0.001	27.27	-0.656	0.004	30.57	-0.619	0.021	27.29
IV	-0.571	0.076	0.52	1.141	0.010	5.52	1.639	0.000	9.20	1.994	0.000	14.31	2.271	0.000	20.23
IV^u	-0.865	0.365	0.15	3.939	0.004	8.63	5.832	0.000	15.64	7.094	0.000	24.29	7.943	0.000	32.21
IV^d	-1.000	0.011	0.90	1.022	0.035	3.00	1.139	0.005	3.44	1.135	0.014	3.93	1.178	0.017	5.03
E: Dispersion of market betas – 4-factor model															
IC	-1.639	0.000	4.19	-0.567	0.000	9.79	-0.420	0.000	19.92	-0.316	0.000	22.04	-0.242	0.002	18.65
IV	-0.026	0.976	-0.02	0.691	0.201	1.25	1.008	0.084	6.31	0.977	0.070	10.65	0.925	0.057	14.96
IVu	3.221	0.278	0.22	3.806	0.048	5.01	4.753	0.009	18.87	4.447	0.004	29.61	4.088	0.002	38.06
IVd	-1.333	0.150	0.15	0.104	0.825	0.00	0.215	0.629	0.21	0.108	0.765	0.09	0.062	0.830	0.04

Table 2: Risk predictability. This table reports the regression coefficients (with corresponding p-values) and the R^2s from regressions of future realized measures of risk on lagged option-implied variables (5) for horizons of 1 to 12 months. Panel A reports the results for future realized market variance. Panels B and C report the results for future idiosyncratic risk, computed as the average sum of squared residuals from the market model and the 4-factor Carhart (1997) model, respectively. Panels D and E report the results for the future cross-sectional dispersion in market betas for the market model and the 4-factor model, respectively. The data are sampled daily, and we use Newey and West (1987) standard errors to adjust for autocorrelation.

we now study univariate predictive regressions of the form (2), but, instead of predicting future market excess returns, we focus on predictability for various future realized measures of risk:

$$\sigma_{\cdot,t\to t+\tau_r} = \gamma + \beta\, PRED(t, t+\tau_r) + \epsilon_t, \tag{5}$$

where $\sigma_{\cdot,t\to t+\tau_r}$ captures the various risk proxies, measured between date t and $t+\tau_r$.[24]

[24] Consistent with our return predictability analysis in Section 1, we always use options with maturity matching the forecasting horizon.

Not surprisingly, Panel A of Table 2 documents that implied (semi)variances, exclusively extracted from index options, have a strong predictive power for future realized market variance—at all horizons. In contrast, implied correlation, jointly extracted from individual stock and index options, predicts future realized market variance only at the 1-month horizon.

More importantly, Panels B and C show that, consistent with our conceptual framework, the implied correlation significantly negatively predicts future realized idiosyncratic risk, as measured by the cross-sectional average of the sum of the squared residuals, from a one-factor market model (Panel B) and a Carhart (1997) four-factor model (Panel C). The predictability is particularly strong at the 9- and 12-month horizons, where the return predictability by the implied correlation is also particularly strong. Implied (semi)variances also predict future realized idiosyncratic risk but with the *opposite sign*.

Also consistent with the stylized framework, the implied correlation is negatively related to the future realized dispersion in market betas. The relation is statistically significant at all horizons and for market betas computed from both a one-factor market model (Panel D) and the Carhart (1997) four-factor model (Panel E). Again, the predictability is particularly strong at long horizons. Similar to the analysis for idiosyncratic risk, implied (semi)variances predict, when statistically significant, the future dispersion in market betas with the *opposite sign*.

Overall, the empirical evidence indicates that the predictions of the implied correlation for future realized measures of risk are distinctly different from those of implied (semi)variances. That the regression coefficients for predicting both future realized idiosyncratic risk and the future dispersion in market betas are of opposite signs has important implications for their abilities to predict future market returns. In particular, in our sample period, realized market returns are, contemporaneously, *negatively* related to idiosyncratic risk (at all horizons) and the cross-sectional dispersion in market betas (for horizons of 3 months and longer). Consequently, the fact that the implied correlation predicts future realized idiosyncratic risk and the future cross-sectional dispersion in market betas with a negative sign is fully consistent with the positive return predictability documented in Section 1. We interpret this as support for our hypothesis that the return predictability of the implied correlation relates to temporal variations in idiosyncratic risk and the cross-sectional dispersion of market betas.

3 Additional Empirical Evidence

We now provide additional empirical evidence in favor of return predictability by option-implied correlation. First, we document that, consistent with economic intuition, the revisions in return forecasts based on implied correlation are negatively correlated with contemporaneous market excess returns. Second, we demonstrate market excess return predictability by implied correlation within a VAR model. Finally, we show that the correlation risk premium also predicts future market excess returns.

3.1 Return Forecasts and Contemporaneous Returns

Economic theory provides a tight link between revisions in expected market excess returns and contemporaneous market excess returns.[25] In particular, using the Campbell and Shiller (1988) log-linearization, one can relate contemporaneous, unexpected excess returns, $e_{t+1} - E_t[e_{t+1}]$ to revisions in expected dividend growth rates and revisions in expected future excess returns:

$$e_{t+1} - E_t[e_{t+1}] = (E_{t+1} - E_t) \sum_n \rho^{n-1} \Delta d_{t+n} - (E_{t+1} - E_t) \sum_n \rho^n r_{f,t+1+n} - (E_{t+1} - E_t) \sum_n \rho^n e_{t+1+n}, \quad (6)$$

where e_{t+1} denotes the time-$t+1$ excess return, Δd_t denotes log dividend-growth, $r_{f,t}$ denotes the time-t risk-free rate, and ρ denotes a parameter with a value slightly less than one.

All else equal, (6) implies that, because of the negative sign in front of revisions to future expected excess returns, an increase in the forecasts of future market excess returns, $(E_{t+1} - E_t)\, e_{t+1+n}$, causes a lower contemporaneous excess return, $e_{t+1} - E_t[e_{t+1}]$. Consequently, to be consistent with economic theory, revisions in the forecasts of future market excess returns from (2) (i.e., revisions in the "fitted values") must negatively correlate with contemporaneous excess returns.

[25] We are grateful to an anonymous referee for suggesting this line of reasoning.

Empirically, we find that the correlation between increments in the predicted market excess returns (as implied by univariate regressions (2) at the 1-month horizon) and contemporaneous daily market excess returns is negative for all options-based variables. For example, the correlation between daily excess returns and changes in forecasts based on the implied correlation is -0.61. Comparably, the respective correlation for the (semi)variance risk premiums is between -0.63 and -0.70. Moreover, even after controlling for total variance, upside and downside semivariances, their risk premiums or all variables together, the (partial) correlations for forecasts based on the implied correlation are all negative. Hence, there is strong and robust empirical evidence that predictions based on the implied correlation are consistent with the relation stipulated by economic theory in the form of the present value relation.

3.2 VAR Models

Studying the *joint* dynamics of the options-based variables and the market excess return can also provide additional information on in-sample return predictability. In that regard, we follow Campbell (1991) and estimate a variety of VAR models of the form

$$z_{t+1} = A\, z_t + \epsilon_{t+1}, \tag{7}$$

where z_t denotes a $k-$dimensional vector consisting of the market excess return ($MKTRF$) and the "predictors." Consistent with the literature, we always include the "standard" predictor—log dividend growth (DIV)—in addition to various combinations of options-based variables.

Table 3 summarizes the estimation results at the monthly frequency. The results are largely consistent with those obtained from the in-sample regression analysis in Section 1. That is, in most specifications only two options-based variables have significant predictive power: the implied correlation and the downside semivariance risk premium. In univariate settings, the coefficients on these variables are highly statistically significant, with R^2s in the range of 3%-4%. Overall, the best explanatory power for future market excess returns stems from a model that features implied correlation as well as the upside and downside semivariance risk premiums, together with dividend growth (leading to an R^2 of 7.2%). The corresponding R^2s for longer

A: VAR coefficients

Model	1	2	3	4	5	6	7	8
IC	0.157	-	0.145	0.114	0.153	-	-	-
	(0.013)	-	(0.030)	(0.093)	(0.058)	-	-	-
IV^u	-	-0.220	-	-	-	-	-	-
	-	(0.406)	-	-	-	-	-	-
IV^d	-	0.276	-	-	-	-	-	-
	-	(0.301)	-	-	-	-	-	-
CRP	-	-	0.041	-	-	-	0.027	0.027
	-	-	(0.534)	-	-	-	(0.697)	(0.698)
VRP	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-
VRP^u	-	-	-	-	0.093	-0.015	-	-0.015
	-	-	-	-	(0.374)	(0.865)	-	(0.868)
VRP^d	-	-	-	0.114	0.037	0.165	0.145	0.155
	-	-	-	(0.084)	(0.731)	(0.054)	(0.031)	(0.085)
DIV	-0.017	-0.003	-0.019	-0.013	-0.016	-0.003	-0.005	-0.005
	(0.781)	(0.959)	(0.761)	(0.833)	(0.796)	(0.961)	(0.939)	(0.938)
$MKTRF$	0.130	0.121	0.119	0.107	0.101	0.080	0.073	0.076
	(0.039)	(0.076)	(0.070)	(0.093)	(0.116)	(0.210)	(0.244)	(0.243)
R^2	3.952	1.024	4.536	6.376	7.177	5.544	5.802	5.643

B: Forecast-error variance decomposition

Model	1	2	3	4	5	6	7	8
IC	0.278	-	0.277	0.268	0.268	-	-	-
IVu	-	0.499	-	-	-	-	-	-
IV^d	-	0.002	-	-	-	-	-	-
CRP	-	-	0.058	-	-	-	0.038	0.038
VRP^u	-	-	-	-	0.009	0.067	-	0.044
VRP^d	-	-	-	0.040	0.041	0.055	0.016	0.079
DIV	0.006	0.006	0.011	0.006	0.004	0.004	0.013	0.005
$MKTRF$	0.716	0.493	0.654	0.686	0.678	0.874	0.933	0.835

Table 3: VARs. This table reports the results for various specifications of the VAR in (7). Panel A reports the coefficients, together with p-values and the R^2 for the market return equation. Panel B reports the forecast error variance decomposition for the market return equation. The estimation results are based on monthly data from January 1996 to December 2017.

horizons, depicted in Figure 4, confirm these conclusions. Quantitatively, models featuring implied correlation can deliver R^2s of up to 20% for 6 months and 12%-13% for 12 months.[26]

In contrast with the individual regressions (2), the VAR model also allows us also to understand how shocks to the various variables propagate through the system. Notably, in the

[26]The implied R^2s are computed by regressing realized returns on VAR-implied forecasts for the market excess return. Note that, as discussed above, R^2s are potentially biased at longer horizons because of autocorrelation in the predictive variables.

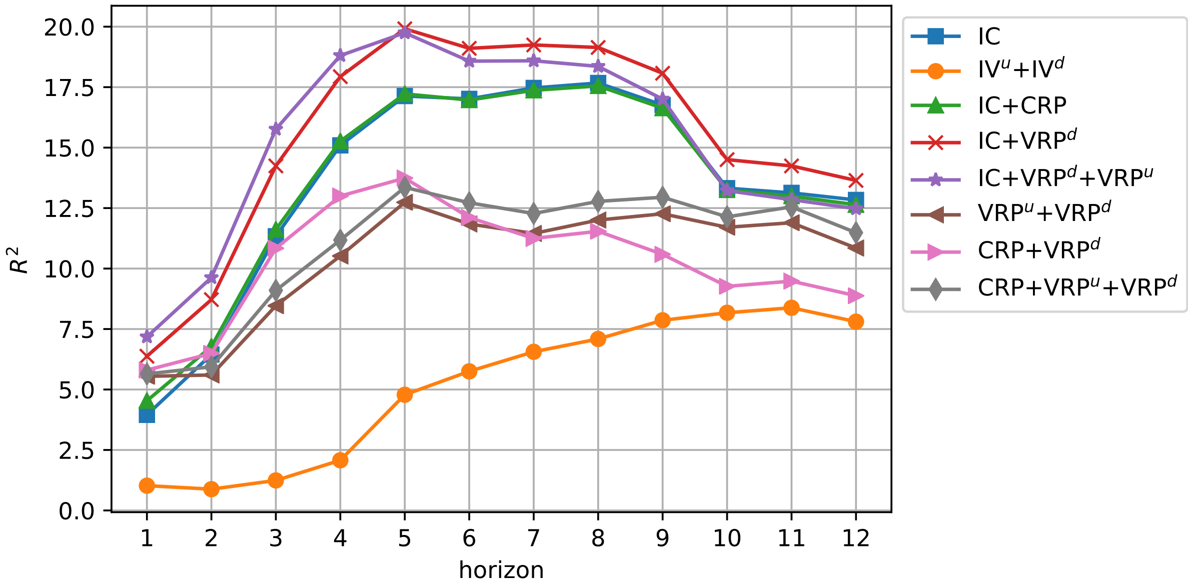


Figure 4: VAR-implied R^2 for market return forecasts. The figure shows the implied R^2 (expressed as a percentage) for predicting market excess returns based on various VAR(1) specifications of Equation (7) for forecasting horizons ranging from 1 to 12 months. All specifications include the market excess return and the log dividend-growth rate for the CRSP aggregate index. The implied R^2 is based on regressing realized monthly market excess returns on VAR-implied return forecasts for a sample period from January 1996 until December 2017.

model with the highest R^2 (as described above), implied correlation positively affects the future downside semivariance risk premium. Notably, implied correlation does not predict future dividend growth, but (only) past dividend growth affects implied correlation, with a negative sign.[27] That is, as intuition suggests, the implied correlation increases in adverse economic conditions.

Finally, Panel B of Table 3 reports variance decompositions of the forecast errors for the market excess return. As expected and consistent with Campbell (1991), the largest fraction of the market return variance can be attributed to news in expected excess returns (ranging from 49% to 93%). However, implied correlation also helps to explain a substantial fraction of the market return variance (up to 28%), whereas the contribution of any of the other variables is typically below 5%. Overall, the results from the VAR models lend further support to the in-sample predictability results.

[27] For the sake of brevity, we omit the impulse response functions here. Appendix B houses them.

3.3 Predictability by the Correlation Risk Premium

The expected correlation captures investors' expectations about stocks' future comovement. In contrast, the correlation risk premium, defined as the difference between the expected correlation under the risk-neutral and the objective probability measure, captures the *price* of correlation risk. Empirically, the correlation risk premium, $CRP(t,T)$, can be computed as the difference between the day-t implied correlation for options with maturity T, $IC(t,T)$, and the corresponding realized correlation for the period $t-(T-t)+1$ to t, $RC(t-(T-t)+1,t)$.[28]

To motivate the pricing of correlation risk, Panel A of Table 4 contrasts the pricing of variance risk in individual stocks and the index. Consistent with earlier studies, the average variance risk premium for individual stocks is not significantly different from zero. At the index level, however, the variance risk premium is significantly positive. Moreover, we find that both upside and downside semivariance risk premiums are significantly different from zero for individual stocks (with opposite signs); on the index level, however, the upside semivariance risk premium is not significant. Because systematic variance risk is directly propagated from individual stocks to the index level, the observed differences in the pricing of variance risk at the stock and index levels only can be explained by a component that contributes to the index variance and is absent from individual stock variances, for example, priced correlation risk, as discussed in detail in Driessen, Maenhout, and Vilkov (2005, 2009).[29]

Consistent with this reasoning, Panel B demonstrates that the correlation risk premium is significantly positive in the data. That is, the expected correlation under the risk-neutral measure is substantially higher than under the physical measure. Also, the correlation risk premium considerably increases with option maturity. Relative to the (semi)variance risk premiums, the correlation risk premium displays fewer abrupt jumps and has fewer "visual" outliers (compare Panels C and D of Figure 1). Moreover, their dynamics are quite distinct, with time-series

[28]Our definition of the correlation risk premium closely follows that used in the literature. But, technically, the definition of implied (realized) correlation does not allow for tradability using delta-hedged option payoffs. However, as shown by Driessen, Maenhout, and Vilkov (2009) and Buraschi, Kosowski, and Trojani (2014), it serves as a good proxy for a correlation risk premium computed from option strategies or correlation swaps.

[29]We remain agnostic about the source of priced correlation risk and merely document its properties. Confer with the literature review for a discussion of potential explanations.

A: Variance and variance risk premium

Months	$\sqrt{IV}$	$\sqrt{RV}$	VRP	p-val	$\sqrt{IV^u}$	$\sqrt{RV^u}$	VRP^u	p-val	$\sqrt{IV^d}$	$\sqrt{RV^d}$	VRP^d	p-val
Index												
1	0.204	0.189	0.006	0.000	0.130	0.134	-0.001	0.269	0.164	0.134	0.009	0.000
3	0.205	0.189	0.006	0.007	0.129	0.134	-0.001	0.458	0.169	0.134	0.011	0.000
6	0.207	0.189	0.007	0.030	0.131	0.134	-0.001	0.673	0.172	0.134	0.012	0.000
9	0.209	0.189	0.008	0.039	0.132	0.134	-0.001	0.826	0.174	0.134	0.012	0.000
12	0.211	0.189	0.009	0.041	0.133	0.134	-0.000	0.934	0.175	0.134	0.013	0.000
Index Constituents												
1	0.390	0.387	0.002	0.584	0.252	0.270	-0.009	0.000	0.306	0.277	0.017	0.000
3	0.372	0.386	-0.010	0.147	0.233	0.270	-0.019	0.000	0.304	0.276	0.016	0.000
6	0.364	0.384	-0.015	0.120	0.222	0.269	-0.023	0.000	0.302	0.275	0.016	0.000
9	0.361	0.384	-0.017	0.153	0.217	0.269	-0.026	0.000	0.301	0.274	0.016	0.000
12	0.358	0.383	-0.019	0.166	0.212	0.269	-0.028	0.000	0.300	0.274	0.015	0.002

B: Correlation and correlation risk premium

Months	IC	p-val	RC	p-val	CRP	p-val
1	0.378	0.000	0.319	0.000	0.059	0.000
3	0.417	0.000	0.318	0.000	0.099	0.000
6	0.444	0.000	0.321	0.000	0.126	0.000
9	0.453	0.000	0.322	0.000	0.134	0.000
12	0.459	0.000	0.324	0.000	0.136	0.000

Table 4: Correlation risk premium. This table reports summary statistics for implied variances and the correlation risk premium for various option maturities. Panel A reports implied and realized variances, semivariances, and their risk premiums (with p-values) separately for the index and for individual stocks (computed as a cross-sectional average). Panel B provides the realized and implied correlations and the correlation risk premium, together with p-values. Statistics are reported for a sample period from January 1996 to December 2017 and, where applicable, annualized. p-values are computed with Newey and West (1987) adjustments for autocorrelation and a number of lags equal to the number of overlaps.

correlations between the correlation risk premium and the (semi)variance risk premiums being quite low (around 0.2).

Using standard in-sample predictive regressions (2), we find that the correlation risk premium also predicts future market excess returns quite well. In univariate regressions, its regression coefficient is highly statistically significant at all horizons, with R^2s of up to 10%. At the 6- and 9-month horizon, it outperforms the other (noncorrelation) variables.[30] Jointly using the implied correlation and the correlation risk premium (slightly) improves the R^2 of the predictive regression at longer horizons. Moreover, consistent with their low time-series correlations, the

[30]Predictions for future market excess returns by the correlation risk premium are also consistent with the present value relation. That is, the correlation between daily market excess returns and increments in market excess return forecasts based on the correlation risk premium is −0.49.

information content of the correlation risk premium is incremental to that of the (semi)variance risk premiums.

In summary, the correlation risk premium is also a robust predictor of future market excess returns, providing nonredundant information relative to purely index-based predictors.

4 Out-of-Sample Predictability

Although many variables have strong predictive power in-sample, hardly any evidence supports *out-of-sample predictability*, as convincingly shown by Goyal and Welch (2008). Accordingly, we now concentrate on the out-of-sample performance of correlation and variance (risk premiums).

4.1 Methodology

To evaluate the different forecasting models, indexed by s, we compare their performance relative to a model based on the historical mean of the market excess return ($s = 0$). This forecast serves as a natural benchmark, because, as documented by Goyal and Welch (2008) and Campbell and Thompson (2008), almost all predictive variables fail to beat it out-of-sample.

We rely on two performance criteria. First, we consider the out-of-sample R^2 relative to the forecasts from the (benchmark) historical average return model:

$$R^2_{s,\tau_r} = 1 - \frac{MSE_{s,\tau_r}}{MSE_{0,\tau_r}}, \tag{8}$$

where $MSE_{s,\tau_r} = \frac{1}{N}\left(e^\top_{s,\tau_r} \times e_{s,\tau_r}\right)$ denotes the mean-squared error of model s. Second, we consider the Diebold and Mariano (1995) loss function, that is, the average squared-error loss relative to the predictions from the benchmark model:

$$\delta_{s,\tau_r} = MSE_{s,\tau_r} - MSE_{0,\tau_r}. \tag{9}$$

A particular model, $s \geq 1$, outperforms the benchmark model based on the average historical return if its out-of-sample R-squared, R^2_{s,τ_r}, is significantly positive and if the average squared-error loss, δ_{s,τ_r}, is significantly negative.

Because of the limited availability of options data, our sample period spans less than 20 years. As a consequence, asymptotic standard errors may not be accurate, so we resort to bootstrapping. Specifically, we use the moving-block bootstrap procedure by Künsch (1989),[31] to randomly resample with replacement from the time series of a model's forecasts to construct bootstrapped distributions for both performance measures.

4.2 Traditional Approach

Traditionally, out-of-sample predictions are based on rolling-window estimations of the predictive regression (2):

$$r_{t-\tau_r \to t} = \gamma_t + \sum_{k \in K} \beta_{k,t} \, PRED_k(t-\tau_r, t) + \varepsilon_t, \tag{10}$$

using, at date t, only observations from the past to avoid any look-ahead bias. The estimated coefficients $\beta_{k,t}$, together with the time-t value of the "predictors" $PRED_k(t, t+\tau_r)$, then form the out-of-sample return forecast $r_{t \to t+\tau}$.

For our empirical analysis, we estimate regressions (10) using a 10-year rolling window of historical data.[32] Panel A of Figure 5 depicts the results for univariate predictability regressions across various forecasting horizons. Although the out-of-sample R^2 for predictions based on the implied correlation is initially negative, it demonstrates a steady increase from 3.5% for 3 months to about 11% for 9- and 12-month horizons, delivering the highest predictive power among the predictors at horizons of 6 months or longer. These results are confirmed by the average squared-error loss (cf. Appendix B for all figures on the squared-error loss). The correlation risk premium

[31] The moving-block bootstrap is shown (see, e.g., Lahiri (1999)) to be comparable in performance to other widely used methods, for example, the stationary bootstrap by Politis and Romano (1994) and the circular block bootstrap of Politis and Romano (1992). However, constant block sizes lead to smaller mean-squared errors compared with random block sizes like in a stationary bootstrap. Specifically, we draw 10,000 random samples of 200 blocks, with blocks of 12 observations (i.e., 1-year blocks) to preserve the autocorrelation in the data.

[32] We follow Kilic and Shaliastovich (2017), who use a 10-year window to show that decomposing the variance risk premium into good and bad components improves the out-of-sample predictability of market returns.

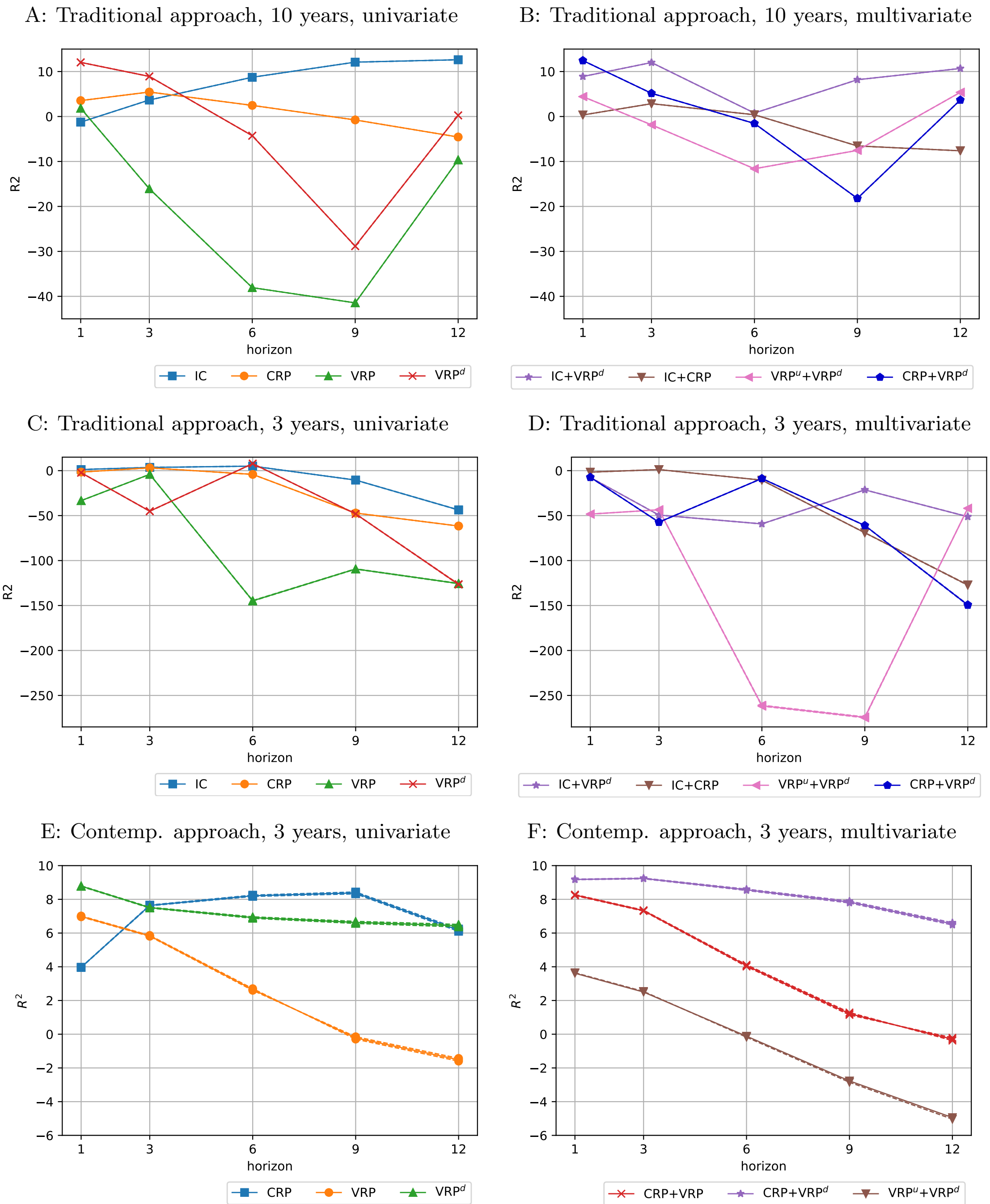


Figure 5: Out-of-sample return predictability. This figure shows the out-of-sample R^2 from (8) for various model specifications, forecasting horizons, and predictability approaches. The results depicted in Panels A and B are based on the traditional approach with a 10-year estimation window. The results shown in Panels C and D are based on the traditional approach with a 3-year estimation window. Finally, the results in Panels E and F are based on the contemporaneous-beta approach with a 3-year estimation window. The left panels show the statistics for the univariate specifications, and the right panels show the statistics for the multivariate ones. In all cases, predictions are made at a monthly frequency.

performs even better than does the implied correlation at short horizons (with R^2s of about 5%); however, its predictive power steadily declines from there onward. The performance of the total variance risk premium, excluding the 1-month horizon, is quite weak. The predictive power of the downside semivariance risk premium is strong at short horizons, delivering the best predictive power (with R^2s around 11%). However, the predictability is not robust in the long term; the out-of-sample R^2 is negative at medium horizons before turning positive again for the 1-year horizon. The results for multivariate regressions, shown in Panel B, are generally quite weak, probably because of overfitting. In particular, the short horizon experiences some improvement, but many out-of-sample R^2s are negative, and the average squared-error losses are positive.

However, we also find that traditional out-of-sample regression techniques cannot fully exploit the predictive power of many variables, because they require a long historical estimation window for the regression coefficients.[33] That is, in our applications, the out-of-sample predictability evaporates when the estimation window is shortened. For example, as shown in Panels C and D, the out-of-sample R^2s with a 3-year estimation window are either very close to zero or negative for both univariate and multivariate regressions.[34] The squared-error loss function mirrors these results. This has important implications: it implies that the approach is hardly applicable for option markets or for new instruments because of limited data availability.

In addition to requiring a relatively long estimation window to fit the parameters, the traditional approach has another drawback. That is, to avoid any look-ahead bias, when forecasting returns from t to $t + \tau_r$, the last observations of the predictors used in the estimation of the regression coefficients in (10) are from time $t - \tau_r$. For example, for an annual forecasting horizon, this implies that the predictors and the estimated predictive relation will be 1 year old and, in the case of changing economic conditions, severely outdated.

4.3 Contemporaneous-Beta Approach

We now propose a new approach for out-of-sample forecasts that relies on *contemporaneous betas* and significantly improves out-of-sample return predictability.

[33]When choosing the estimation horizon for the rolling-window regressions (10), one typically faces a trade-off between the precision of the coefficients $\beta_{k,t}$ and their ability to reflect the current economic conditions.

[34]We tried a variety of rolling-window horizons ranging from 1 to 5 years. The results are qualitatively the same.

4.3.1 Intuition

The approach effectively extends the contemporaneous-regression idea, introduced by Cutler, Poterba, and Summers (1989) and Roll (1988). Their approach traditionally has been used to study contemporaneous innovations in market excess returns and the respective variables (similar to our VAR analysis in Section 3.2). In contrast, our objective is to *combine* the coefficients obtained from the contemporaneous regressions with time-t risk premiums to efficiently forecast future market excess returns.

Intuitively, the idea can be motivated by a simple linear framework within the arbitrage pricing theory. For example, consider a setting in which the market excess return is described by a linear factor model with (time-varying) factor exposures:[35]

$$r_{t+1} = \sum_{k=1}^{K} \beta_{k,t}\, f_{k,t+1} + \epsilon_{t+1}, \tag{11}$$

where $f_{k,t+1}$ denotes the return on factor $k \in \{1, \ldots, K\}$, and ϵ_{t+1} denotes "noise." Then the equity risk premium is given by

$$ERP(t, t+1) = \sum_{k=1}^{K} \beta_{k,t}\, RP_{k,t+1} + p_t(\epsilon_{t+1}), \tag{12}$$

where $RP_{k,t+1}$ denotes the risk-premium on the kth factor, and $p_t(\epsilon_{t+1})$ denotes the pricing error, for example, because of omitted factors.[36] Consequently, in this setting, to arrive at a good estimate for the future market excess return, one needs precise estimates of the time-t factor exposures $\beta_{k,t}$ and the corresponding risk premiums $RP_{k,t+1}$.

Our approach relies on first estimating a regression of the type (11) using daily (or even intraday) market excess returns and, at the same frequency, shocks to the variables (e.g., the

[35] A similar structure endogenously arises in Bollerslev, Tauchen, and Zhou (2009), Bollerslev, Todorov, and Xu (2015), Kilic and Shaliastovich (2017), and Feunou, Jahan-Parvar, and Okou (2017).

[36] Although early applications of the arbitrage pricing theory (e.g., Ross (1976)) derived bounds on the pricing errors by assuming the correct specification of a linear factor model, recent studies (e.g., Raponi, Uppal, and Zaffaroni (2018)) show that the pricing errors are bounded even when the model is misspecified or when factors are omitted.

implied correlation or implied variance). This delivers accurate and up-to-date estimates of $\beta_{k,t}$. Intuitively, the higher frequency of the observations has two advantages: (1) it improves the properties of the regression coefficients (betas), which are essentially functions of second moments, and (2) it allows us to shorten the estimation window so that betas quickly adjust to new information. In the second step, one can then use the "pricing equation" (12) to predict future market excess returns, by combining the estimated coefficients with the time-t risk premiums on the respective variable (e.g., the correlation risk premium or the variance risk premium).

4.3.2 Implementation

For the estimation of regression (11), the contemporaneous-beta approach requires the estimation of shocks in the variables, that is, estimates of the "factors." In the case of option-implied variables, these shocks can be proxied for using shocks to integrated expected variables. For example, note that the time-t-expected integrated correlation obtained from options with maturity T can be decomposed as follows:

$$\begin{aligned} IC(t,T) &= E_t^{\mathbb{Q}}\left[E_{t+\Delta t}^{\mathbb{Q}}\left[\int_t^{t+\Delta t}\rho(s)ds + \int_{t+\Delta t}^{T}\rho(s)ds\right]\right] \\ &= E_t^{\mathbb{Q}}\left[\int_t^{t+\Delta t}\rho(s)ds\right] + E_t^{\mathbb{Q}}\left[IC(t+\Delta t,T)\right], \end{aligned}$$

where ρ denotes the stochastic process of correlation, and $\mathbb{Q}$ denotes the risk-neutral probability measure. Hence, increments to implied correlation are given by

$$\begin{aligned} \Delta IC(t+\Delta t,T) &= IC(t+\Delta t,T) - IC(t,T) \\ &= \underbrace{IC(t+\Delta t,T) - E_t^{\mathbb{Q}}\left[IC(t+\Delta t,T)\right]}_{\text{“true” shock to the state variable}} - E_t^{\mathbb{Q}}\left[\int_t^{t+\Delta t}\rho(s)ds\right]. \end{aligned} \tag{13}$$

If the last term in equation (13)—the expected integrated correlation over a short period of time Δt—is small, that is, if risk-neutral expected integrated correlation can be well approximated by a martingale,[37] short-interval increments are indeed proxies for random shocks to correlation. The same argument holds for shocks to (semi)variance.

Importantly, our approach allows us to use increments from any option maturity; in particular, one can use variables extracted from the most liquid options with short maturities. Doing so reduces the potential bias in betas arising from nonlinearities in response to random shocks in a variable.[38] One only needs to correct the resultant betas for the difference in variability of the regressors used for beta estimation (i.e., increments in variables) in (11) and the variability of the risk premiums for the pricing equation (12). Appendix A derives a simple procedure that does exactly that.

4.3.3 Results

To illustrate the advantages of the contemporaneous-beta approach relative to the traditional approach, we now compare its performance for the case of a 3-year historical rolling-window estimation period. Recall that for this case, the traditional approach essentially delivers zero out-of-sample predictability.

We use daily market excess returns and daily increments in the implied correlation and implied (semi)variances from options with 1-month maturity to estimate the contemporaneous betas.[39] To arrive at the out-of-sample predictions, we then combine these betas with the time-t-expected risk premium on the predictors, that is, the correlation and the downside semivariance risk premium, for an option maturity matching the forecasting horizon.

As shown in Panels E and F of Figure 5, the new approach leads to stable out-of-sample return predictability for most horizons, despite the short estimation window. For univariate predictions (Panel E), the correlation risk premium delivers the highest out-of-sample R^2 for 3- to 9-month horizons, whereas the downside semivariance risk premium performs best at the very short and

[37]Empirical evidence supports this approximation. For example, Filipović, Gourier, and Mancini (2016, p. 58) find that a "martingale model provides relatively accurate forecasts for the one-day horizon variance." Moreover, integrated expected variance and the integrated expected correlation are highly persistent, with first-order autocorrelations between 0.97 and 0.994 for variance and between 0.97 and 0.993 for correlation at various option maturities in our sample period.

[38]We thank an anonymous referee for pointing out the potential effect of option maturity on the bias in betas.

[39]These are then adjusted by the ratio of the 3-year historical volatilities of the implied variables and the respective risk premiums, as described in Appendix A.

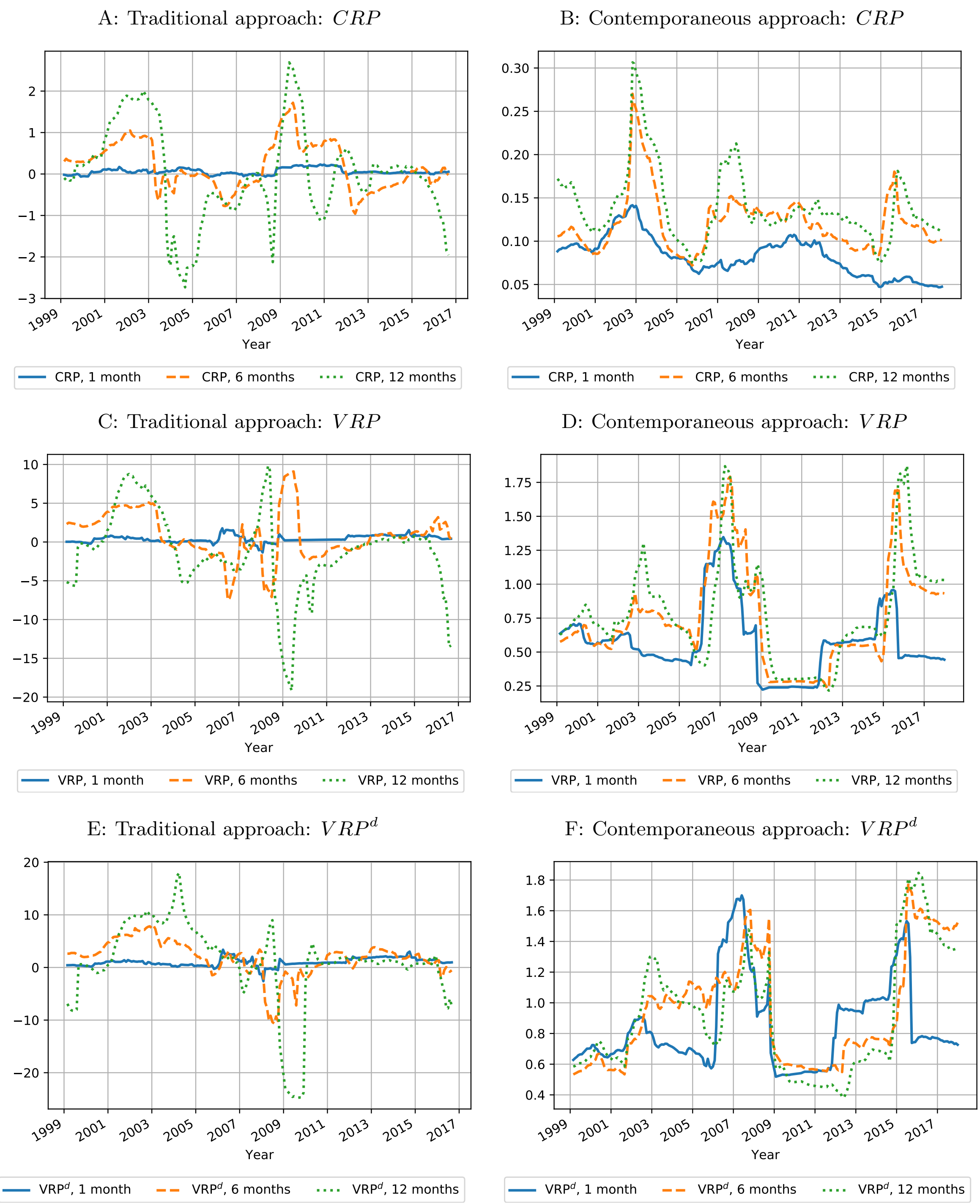


Figure 6: Dynamics of out-of-sample regression coefficients. This figure shows the time series of the regressions coefficients in univariate models for out-of-sample market return forecasts for forecasting horizons of 1, 6, and 12 months. The panels on the left show the regression coefficients from the traditional approach, and the ones on the right show the regression coefficients from the contemporaneous approach. For both approaches, we use a 3-year historical estimation window.

long horizons. Interestingly, even the variance risk premium displays a decent performance with an R^2 of 7% at a 1-month horizon, declining to 2.2% for 6 months. Combining the correlation and the downside semivariance risk premium leads to improvements at some horizons (Panel F). The results for the Diebold and Mariano (1995) loss function confirm these results.

To understand what drives the differences in the approaches' predictive power, recall that the pricing equation is the same for the traditional and the contemporaneous-beta approaches. However, the two approaches considerably differ in the estimation of the regression coefficients $\beta_{k,t}$. Figure 6 depicts the estimated betas for the respective variables across various forecasting horizons. As is apparent, the resultant betas show very different dynamics. That is, the traditional betas are far more volatile (pay attention to the y-axis scales) and exhibit more abrupt jumps in reaction to "extreme" observations. Also, whereas the traditional betas often change sign, the contemporaneous betas, though substantially fluctuating over time, always stay positive. Moreover, the traditional betas fluctuate much more with the forecasting horizon, with standard deviations across horizons that are 5 to 25 times larger than those of the contemporaneous approach.

In summary, the contemporaneous-beta approach delivers promising results for out-of-sample market return predictability. Notably, even for an extremely short estimation window of 1 month, the predictive performance is impressive. Moreover, because of its straightforward implementation, the contemporaneous beta approach naturally lends itself to many other settings and subsequent work.

5 Conclusion

In this paper, we further explore the empirical evidence that information about the comovement of individual stocks, extracted from option prices, can predict future market returns. In particular, the objective is to improve our understanding of the predictive power as well as the information content of options-based measures of stock comovement and to illustrate the underlying economic sources explaining the return predictability.

We document that both the implied correlation and the correlation risk premium, jointly extracted from index options and the cross-section of individual stock options, have strong predictive power for future market excess returns. For example, in-sample, the implied correlation predicts future market excess returns with impressive R^2s for horizons of up to 12 months and, except for the 1-month horizon, outperforms other option-implied predictors, such as implied (semi)variances and their risk premiums. Moreover, the predictability is not subsumed by information extracted from the risk-neutral marginal distribution of the aggregate stock market and can be mostly attributed to the forward-looking nature of options. The predictability is even robust *out-of-sample*, with the implied correlation and the correlation risk premium delivering positive out-of-sample R^2s at long horizons.

We provide evidence that temporal variations in idiosyncratic risk and the cross-sectional dispersion in systematic risk are largely responsible for this predictability. In particular, in contrast to measures of risk exclusively based on the marginal distribution of the market, the expected correlation is intimately linked to these variables. Consequently, if idiosyncratic risk is priced or the equity risk premium is affected by the cross-sectional beta dispersion, the implied correlation will be able to capture the resulting fluctuations in future market returns, whereas index-based variance measures will not. Consistent with this line of reasoning, we empirically find that the implied correlation negatively predicts both future idiosyncratic volatility and the future cross-sectional dispersion in market betas.

By highlighting the importance of information not spanned by the marginal distribution of the market, our results have important implications for theoretical work trying built pricing models from individual stock dynamics (see also Leippold and Trojani (2010)). Moreover, by providing complementary evidence on the pricing of idiosyncratic risk at the aggregate level, our work connects to the ongoing debate about the importance and pricing of idiosyncratic risk. Although we exclusively focus on future market returns, option-implied information also should be helpful in better understanding the price of idiosyncratic risk in the cross-section of expected stock returns. Moreover, the use of (newly available) options data on sector indices could help relax the assumption of homogeneity in the correlations among all stocks and, hence, further improve our understanding of the information content of the implied correlation.[40]

Methodologically, our extension of the contemporaneous-beta approach, which combines high-frequency increments in option-implied variables with the respective risk premiums, can help to improve out-of-sample return forecasts in many applications. In our case, the approach generates considerably higher market return predictability than does the traditional approach (with out-of-sample R^2s of around 8% for horizons of 3 to 6 months and 7% for 12 months).

[40]For example, Kelly, Lustig, and Van Nieuwerburgh (2016) relate implied correlation extracted from the financial sector index to the pricing of crash risk during the recent financial crises.

References

Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens, 2001, "The Distribution of Stock Return Volatility," *Journal of Financial Economics*, 61(1), 43–76.

Andersen, T. G., and O. Bondarenko, 2007, "Construction and Interpretation of Model-Free Implied Volatility," Working Paper 13449, National Bureau of Economic Research.

Andersen, T. G., O. Bondarenko, and M. T. Gonzalez-Perez, 2015, "Exploring Return Dynamics via Corridor Implied Volatility," *Review of Financial Studies*.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang, 2006, "The cross-section of volatility and expected returns," *Journal of Finance*, 61(1), 259–299.

——— , 2009, "High idiosyncratic volatility and low returns: international and further US evidence," *Journal of Financial Economics*, 91(1), 1–23.

Bakshi, G. S., N. Kapadia, and D. B. Madan, 2003, "Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options," *The Review of Financial Studies*, 16(1), 101–143.

Bali, T. G., N. Cakici, X. S. Yan, and Z. Zhang, 2005, "Does Idiosyncratic Risk Really Matter?," *The Journal of Finance*, 60(2), 905–929.

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(03), 633–661.

Bollerslev, T., G. Tauchen, and H. Zhou, 2009, "Expected Stock Returns and Variance Risk Premia," *Review of Financial Studies*, 22(11), 4463–4492.

Bollerslev, T., V. Todorov, and L. Xu, 2015, "Tail risk premia and return predictability," *Journal of Financial Economics*, 118(1), 113–134.

Boudoukh, J., M. Richardson, and R. F. Whitelaw, 2008, "The Myth of Long-Horizon Predictability," *Review of Financial Studies*, 21(4), 1577–1605.

Britten-Jones, M., and A. Neuberger, 2000, "Option Prices, Implied Price Processes, and Stochastic Volatility," *Journal of Finance*, 55, 839–866.

Buraschi, A., R. Kosowski, and F. Trojani, 2014, "When There Is No Place to Hide: Correlation Risk and the Cross-Section of Hedge Fund Returns," *Review of Financial Studies*, 27(2), 581–616.

Buraschi, A., P. Porchia, and F. Trojani, 2010, "Correlation Risk and Optimal Portfolio Choice," *Journal of Finance*, 65(1), 393–420.

Buraschi, A., F. Trojani, and A. Vedolin, 2014, "When Uncertainty Blows in the Orchard: Comovement and Equilibrium Volatility Risk Premia," *Journal of Finance*, 69(1), 101–137.

Buss, A., and G. Vilkov, 2012, "Measuring Equity Risk with Option-implied Correlations," *Review of Financial Studies*, 25(10), 3113–3140.

Campbell, J. Y., 1991, "A Variance Decomposition for Stock Returns," *The Economic Journal*, 101(405), 157–179.

Campbell, J. Y., and R. J. Shiller, 1988, "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1, 195–228.

Campbell, J. Y., and S. B. Thompson, 2008, "Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?," *The Review of Financial Studies*, 21(4), 1509–1531.

Carhart, M. M., 1997, "On Persistence in Mutual Fund Performance," *Journal of Finance*, 52(1), 57–82.

Chang, B.-Y., P. Christoffersen, K. Jacobs, and G. Vainberg, 2012, "Option-Implied Measures of Equity Risk," *Review of Finance*, 16(2), 385–428.

Christoffersen, P., K. Jacobs, and B. Y. Chang, 2013, "Forecasting with Option-Implied Information," in *Handbook of Economic Forecasting*. Elsevier, vol. 2, chap. 10, pp. 581–656.

Cosemans, M., 2011, "The Pricing of Long and Short Run Variance and Correlation Risk in Stock Returns," Working Paper.

Cutler, D. M., J. M. Poterba, and L. H. Summers, 1989, "What moves stock prices?," *The Journal of Portfolio Management*, 15(3), 4–12.

Diebold, F., and R. Mariano, 1995, "Comparing Predictive Accuracy," *Journal of Business & Economic Statistics*, 13(3), 253–263.

Drechsler, I., and A. Yaron, 2011, "What's Vol Got to Do with It," *Review of Financial Studies*, 24(1), 1–45.

Driessen, J., P. Maenhout, and G. Vilkov, 2005, "Option-Implied Correlations and the Price of Correlation Risk," Working paper, INSEAD.

——— , 2012, "Option-Implied Correlations and the Price of Correlation Risk," Working paper, INSEAD.

Driessen, J., P. J. Maenhout, and G. Vilkov, 2009, "The Price of Correlation Risk: Evidence from Equity Options," *Journal of Finance*, 64(3), 1377–1406.

Dumas, B., 1995, "The Meaning of the Implicit Volatility Function in Case of Stochastic Volatility," *HEC Working Paper*.

Fan, Z., X. Xiao, and H. Zhou, 2018, "Variance Risk Premium, Higher Order Risk Premium, and Expected Stock Returns," Working Paper.

Faria, G., R. Kosowski, and T. Wang, 2016, "The Correlation Risk Premium: International Evidence," Working Paper Imperial College Business School.

Feunou, B., M. R. Jahan-Parvar, and C. Okou, 2017, "Downside Variance Risk Premium," *Journal of Financial Econometrics*.

Filipović, D., E. Gourier, and L. Mancini, 2016, "Quadratic variance swap models," *Journal of Financial Economics*, 119(1), 44–68.

Ghysels, E., P. Santa-Clara, and R. Valkanov, 2005, "There Is a Risk-Return Trade-off After All," *Journal of Financial Economics*, 76(3), 509–548.

Goyal, A., and P. Santa-Clara, 2003, "Idiosyncratic Risk Matters!," *Journal of Finance*, 58(3), 975–1008.

Goyal, A., and I. Welch, 2008, "A Comprehensive Look at The Empirical Performance of Equity Premium Prediction," *Review of Financial Studies*, 21(4), 1455–1508.

Guo, H., and R. Savickas, 2006, "Idiosyncratic Volatility, Stock Market Volatility, and Expected Stock Returns," *Journal of Business and Economic Statistics*, 24(1), 43–56.

Kelly, B., H. Lustig, and S. Van Nieuwerburgh, 2016, "Too-Systemic-to-Fail: What Option Markets Imply about Sector-Wide Government Guarantees," *American Economic Review*, 106(6), 1278–1319.

Kilic, M., and I. Shaliastovich, 2017, "Good and Bad Variance Premia and Expected Returns," *Management Science*, forthcoming.

Kogan, L., and D. Papanikolaou, 2012, "A Theory of Firm Characteristics and Stock Returns: The Role of Investment-Specific Shocks," Working Paper 17975, National Bureau of Economic Research.

——— , 2013, "Firm Characteristics and Stock Returns: The Role of Investment-Specific Shocks," *The Review of Financial Studies*, 26(11), 2718–2759.

Kostakis, A., T. Magdalinos, and M. P. Stamatogiannis, 2015, "Robust Econometric Inference for Stock Return Predictability," *The Review of Financial Studies*, 28(5), 1506–1553.

Künsch, H. R., 1989, "The Jackknife and the Bootstrap for General Stationary Observations," *Ann. Statist.*, 17(3), 1217–1241.

Lahiri, S. N., 1999, "Theoretical comparisons of block bootstrap methods," *Ann. Statist.*, 27(1), 386–404.

Leippold, M., and F. Trojani, 2010, "Asset pricing with Matrix Affine Jump Diffusions," Working Paper.

Liu, L., A. Patton, and K. Sheppard, 2015, "Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes," *Journal of Econometrics*, 187(1), 293–311.

Martin, I., 2013, "Simple Variance Swaps," NBER Working Paper 16884.

——— , 2017, "What is the Expected Return on the Market?," *The Quarterly Journal of Economics*, 132(1), 367–433.

Martin, I., and C. Wagner, Forthcoming, "What is the Expected Return on a Stock?," *Journal of Finance*.

Mueller, P., A. Stathopoulos, and A. Vedolin, 2017, "International Correlation Risk," *Journal of Financial Economics*, 126(2), 270–299.

Newey, W. K., and K. D. West, 1987, "A Simple, Positive-semidefinite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55(3), 703–708.

Piatti, I., 2015, "Heterogeneous beliefs about rare event risk in the lucas orchard," Working Paper, Said Business School.

Politis, D., and J. Romano, 1994, "The Stationary Bootstrap," *Journal of the American Statistical Association*, 89, 1303–1313.

Politis, D. N., and J. P. Romano, 1992, "A General Resampling Scheme for Triangular Arrays of α-Mixing Random Variables with Application to the Problem of Spectral Density Estimation," *Ann. Statist.*, 20(4), 1985–2007.

Pollet, J. M., and M. Wilson, 2010, "Average correlation and stock market returns," *Journal of Financial Economics*, 96(3), 364–380.

Raponi, V., R. Uppal, and P. Zaffaroni, 2018, "Portfolio Choice with Model Misspecification: A Foundation for Alpha and Beta Portfolios," Working Paper EDHEC.

Roll, R., 1988, "R2," *The Journal of Finance*, 43(3), 541–566.

Ross, S., 1976, "The Arbitrage Theory of Capital Asset Pricing," *Journal of Economic Theory*, 13(3), 341–360.

Santos, T., and P. Veronesi, 2004, "Conditional Betas," Working Paper 10413, National Bureau of Economic Research.

Schneider, P., C. Wagner, and J. Zechner, 2017, "Low Risk Anomalies?," Working Paper.

Appendix

A "Normalizing" Variance and Correlation Betas

The variance and correlation betas, $\beta_{\Delta IV,t}$ and $\beta_{\Delta IC,t}$ estimated by regressing market excess returns on increments in option-implied variables in (11), differ from the exposures $\beta_{V,t}$ and $\beta_{\rho,t}$ in the pricing equation (12). In particular, they need to be adjusted for the difference in the variability of the regressors used for beta estimation (i.e., increments in risk-neutral expected variance and correlation) and the variability of the predictors in the pricing equation (i.e., the variance and correlation risk premium).

Intuitively, the variance beta in the pricing equation, $\beta_{V,t}$, can be decomposed into (i) the correlation between the market excess return and the *variance risk premium*, and (ii) the ratio of their volatilities. Similarly, the variance beta in the estimation equation, $\beta_{\Delta IV,t}$, can be decomposed into (i) the correlation between the market excess return and the *increments in implied variance*, and (ii) the ratio of their volatilities.

Consequently, if we assume that the correlation between returns and increments in implied variance equals the correlation between returns and the variance risk premium, that is, $\mathrm{Corr}\left(r_{t+\tau}, VRP(t,t+\tau)\right) = \mathrm{Corr}\left(r_{t+\tau}, \Delta IV(t,t+\tau)\right)$, one gets:

$$\begin{aligned}
\beta_{V,t} &= \mathrm{Corr}\left(r_{t+\tau}, VRP(t,t+\tau)\right) \times \frac{\mathrm{Vol}\left(r_{t+\tau}\right)}{\mathrm{Vol}\left(VRP(t,t+\tau)\right)} \\
&= \mathrm{Corr}\left(r_{t+\tau}, \Delta IV(t,t+\tau)\right) \times \frac{\mathrm{Vol}\left(r_{t+\tau}\right)}{\mathrm{Vol}\left(\Delta IV(t,t+\tau)\right)} \times \frac{\mathrm{Vol}\left(\Delta IV(t,t+\tau)\right)}{\mathrm{Vol}\left(VRP(t,t+\tau)\right)} \\
&= \beta_{DeltaIV,t} \times \frac{\mathrm{Vol}\left(\Delta IV(t,t+\tau)\right)}{\mathrm{Vol}\left(VRP(t,t+\tau)\right)}.
\end{aligned}$$

Accordingly, one can simply adjust the variance beta, $\beta_{\Delta IV,t}$, by the ratio of the volatility of the increments in implied variance and the volatility of the variance risk premium. Both variables are observable, so that the ratio can easily be estimated from the data. The same principle works for up and down semivariance betas, and for correlation betas. For example, computations for the correlation risk premium lead to a comparable adjustment:

$$\beta_{\rho,t} = \beta_{\Delta IC,t} \times \frac{\mathrm{Vol}\left(\Delta IC(t,t+\tau)\right)}{\mathrm{Vol}\left(CRP(t,t+\tau)\right)}.$$

B Additional Figures

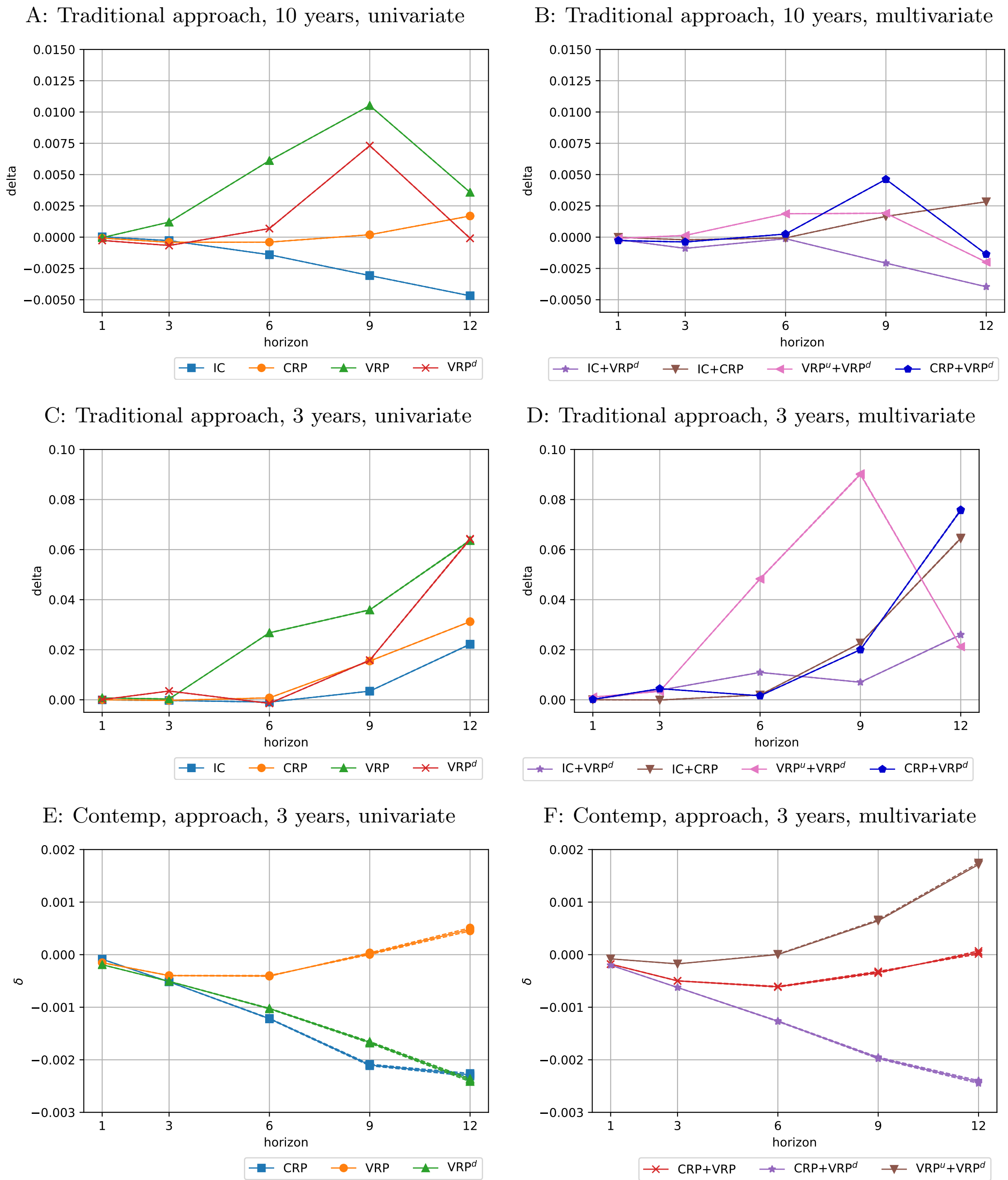
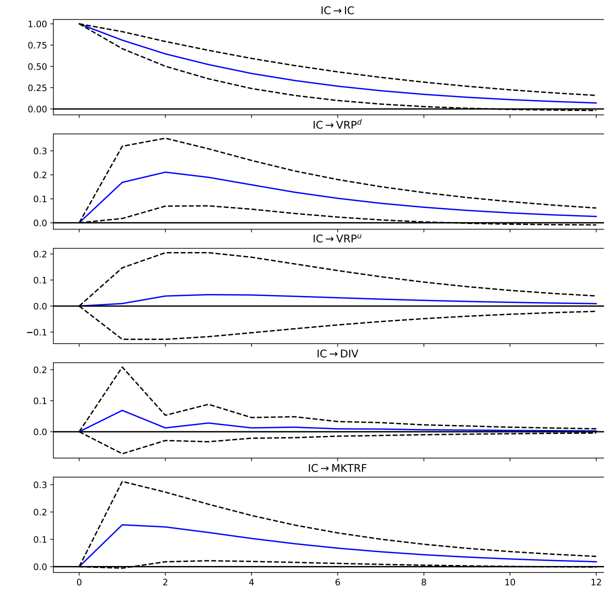
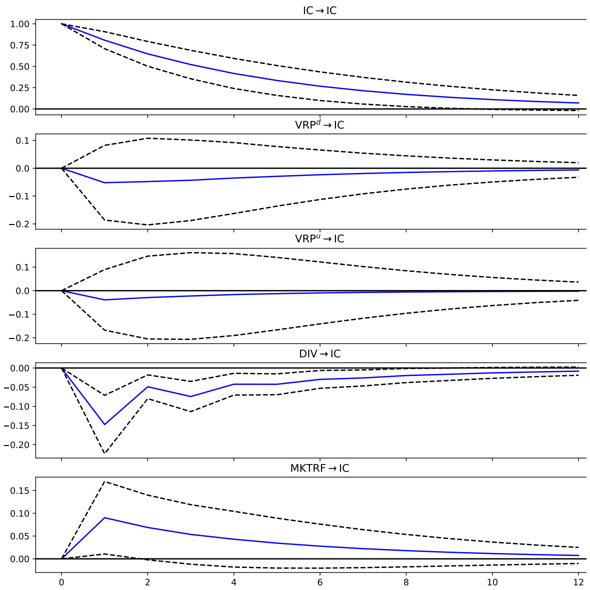


Figure A1: Out-of-sample return predictability. The figure shows the Diebold and Mariano (1995) squared-loss function (9); for various model specifications, forecasting horizons and predictability approaches. The results depicted in Panels A and B are based on the traditional approach with a 10-year estimation window. The results shown in Panels C and D are based on the traditional approach with a 3-year estimation window. Finally, the results in Panels E and F are based on the contemporaneous-beta approach with a 3-year estimation window. The left panels show statistics for univariate specifications and the right ones for multivariate specifications. In all cases, predictions are made at a monthly frequency.

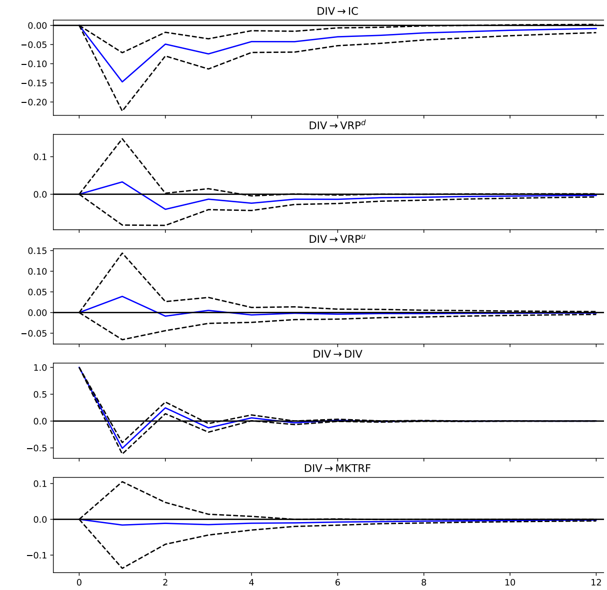
A: IC to other variables.



B: Other variables to IC.



C: DIV to other variables.



D: MKTRF to other variables.

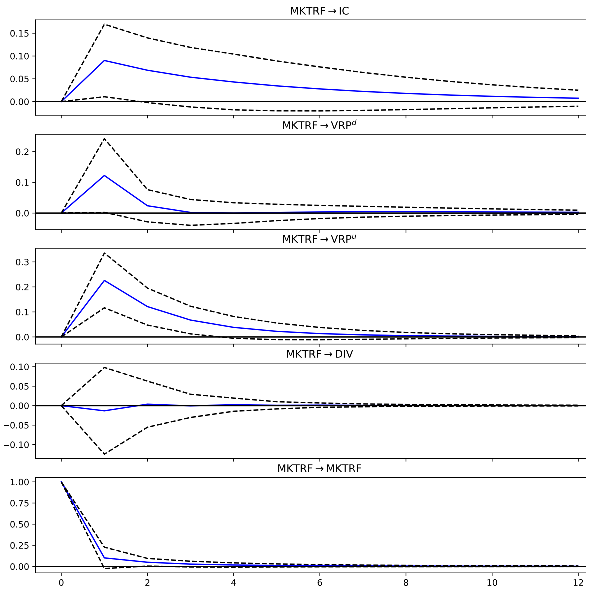


Figure A2: VAR impulse response functions. The figure shows a set of selected impulse response functions for horizons of up to 12 months; based on a VAR(1) model featuring the market excess return, log dividend-growth, the implied correlation and the up and down semi-variance risk premiums. The VAR(1) is estimated using data from January 1996 to December 2017. The dotted lines indicate confidence bounds.

C Additional Tables

	1	3	6	9	12
IC	0.766	0.832	0.892	0.913	0.914
IV	0.808	0.857	0.898	0.904	0.917
IV^u	0.826	0.876	0.918	0.927	0.938
IV^d	0.782	0.831	0.880	0.884	0.902
CRP	0.152	0.648	0.835	0.867	0.881
VRP	0.378	0.705	0.871	0.897	0.924
$VRPu$	0.480	0.792	0.921	0.946	0.963
VRP^d	0.266	0.444	0.687	0.745	0.808

Table A1: Autocorrelations. The table reports autocorrelations for options-based variables, computed for several maturities (from one to 12 months) and sampled at the end of each month—for a sample period from January 1996 to December 2017. The variables are the implied correlation (IC), implied variance (IV), implied upside and downside semi-variances (IV^u and IV^d), and the risk premiums for correlation (CRP), variance (VRP) and two semivariances (VRP^u and VRP^d).

A: Levels

	IC	IV	IV^u	IV^d	CRP	VRP	VRP^u	VRP^d
IC	1.000	0.579	0.536	0.596	0.253	-0.011	-0.157	0.333
IV	0.579	1.000	0.992	0.994	-0.013	-0.377	-0.551	0.130
IV^u	0.536	0.992	1.000	0.972	-0.015	-0.350	-0.518	0.134
IV^d	0.596	0.994	0.972	1.000	-0.014	-0.412	-0.588	0.109
CRP	0.253	-0.013	-0.015	-0.014	1.000	0.362	0.308	0.408
VRP	-0.011	-0.377	-0.350	-0.412	0.362	1.000	0.971	0.844
VRP^u	-0.157	-0.551	-0.518	-0.588	0.308	0.971	1.000	0.695
VRP^d	0.333	0.130	0.134	0.109	0.408	0.844	0.695	1.000

B: Differences

	IC	IV	IV^u	IV^d	CRP	VRP	VRP^u	VRP^d
IC	1.000	0.636	0.625	0.615	0.233	0.064	0.001	0.220
IV	0.636	1.000	0.987	0.991	0.050	-0.172	-0.274	0.109
IV^u	0.625	0.987	1.000	0.958	0.031	-0.129	-0.223	0.131
IV^d	0.615	0.991	0.958	1.000	0.064	-0.223	-0.329	0.071
CRP	0.233	0.050	0.031	0.064	1.000	0.202	0.182	0.235
VRP	0.064	-0.172	-0.129	-0.223	0.202	1.000	0.981	0.938
VRP^u	0.001	-0.274	-0.223	-0.329	0.182	0.981	1.000	0.858
VRP^d	0.220	0.109	0.131	0.071	0.235	0.938	0.858	1.000

Table A2: Time-series correlations. The table reports time-series correlations between option-implied variables and the respective risk premiums. Panel A shows correlations in levels and Panel B in the first differences. The variables include the implied correlation (IC), the implied variance (IV), the implied upside and downside semi-variances (IV^u and IV^d), and the risk premiums for correlation (CRP), variance (VRP) and two semivariances (VRP^u and VRP^d). All variables are computed from one-month options and are sampled daily.