Network for Studies on Pensions, Aging and Retirement

Netspar DISCUSSION PAPERS

Joost Driessen, Pascal Maenhout and Grigory Vilkov

Option-Implied Correlations and the Price of Correlation Risk

DP 07/2013-061

Option-Implied Correlations and the Price of Correlation Risk*

Joost Driessen† Pascal Maenhout‡ Grigory Vilkov§

First version: September 2005

Current version: July 26, 2013

Abstract

Motivated by extensive evidence that stock-return correlations are stochastic, we analyze whether the risk of correlation changes (affecting diversification benefits) may be priced. We propose a direct and intuitive test by comparing option-implied correlations between stock returns (obtained by combining index option prices with prices of options on all index components) with realized correlations. Our parsimonious model shows that the substantial gap between average implied (39.5% for S&P500 and 46.0% for DJ30) and realized correlations (32.5% and 35.5%, respectively) is direct evidence of a large negative correlation risk premium. Empirical implementation of our model also indicates that the index variance risk premium can be attributed to the high price of correlation risk. Finally, we provide evidence that option-implied correlations have remarkable predictive power for future stock market returns, which also stays significant after controlling for a number of fundamental market return predictors.

*We would like to thank Yacine Aït-Sahalia, David Bates, Jonathan Berk, Oleg Bondarenko, Michael Brandt, Menachem Brenner, John Campbell, Mike Chernov, Greg Duffee, Darrell Duffie, Rob Engle, Jan Ericsson, Gerard Gennotte, Jens Jackwerth, Chris Jones, Frank de Jong, Hayne Leland, Toby Moskowitz, Anthony Neuberger, Pedro Santa-Clara, Josh Rosenberg, Mark Rubinstein, Ken Singleton, Otto van Hemert, Robert Whitelaw, Zhipeng Zhang and especially Bernard Dumas for comments and stimulating discussions. We received helpful comments from seminar participants at Berkeley Haas School of Business, BI Oslo, Cornell Johnson School, HEC Lausanne, INSEAD, LBS-LSE-Oxford Asset Pricing Workshop, MIT Sloan, NY Fed, NYU Stern, OptionMetrics Users Conference, Stanford GSB, Tilburg, University of Amsterdam, University of Bonn, University of Frankfurt, University of Rotterdam, Warwick Business School, Yale SOM, CEPR Summer Symposium Gerzensee, EFA, the Duke-UNC Asset Pricing Conference and WFA. We gratefully acknowledge the financial support of INSEAD R&D.

†Tilburg University, Faculty of Economics and Business, Department of Finance, P.O. Box 90153, 5000 LE, Tilburg, the Netherlands. Email: j.j.a.g.driessen@uvt.nl.

‡INSEAD, Finance Department, 1 Ayer Rajah Avenue, 138676 Singapore, Singapore. Email: pascal.maenhout@insead.edu.

§Goethe University Frankfurt, Finance Department, Grüneburgplatz 1, 60323 Frankfurt am Main, Germany. Email: vilkov@vilkov.net.

Option-Implied Correlations and the Price of Correlation Risk

Abstract

Motivated by extensive evidence that stock-return correlations are stochastic, we analyze whether the risk of correlation changes (affecting diversification benefits) may be priced. We propose a direct and intuitive test by comparing option-implied correlations between stock returns (obtained by combining index option prices with prices of options on all index components) with realized correlations. Our parsimonious model shows that the substantial gap between average implied (39.5% for S&P500 and 46.0% for DJ30) and realized correlations (32.5% and 35.5%, respectively) is direct evidence of a large negative correlation risk premium. Empirical implementation of our model also indicates that the index variance risk premium can be attributed to the high price of correlation risk. Finally, we provide evidence that option-implied correlations have remarkable predictive power for future stock market returns, which also stays significant after controlling for a number of fundamental market return predictors.

The correlation structure between assets constitutes one of the most fundamental concepts in financial economics and is essential for portfolio choice, risk management and asset pricing since it drives portfolio and market risk. While this correlation structure between assets is often assumed to be fixed, there is substantial evidence that correlations are stochastic and in fact subject to risk themselves. Moreover, as correlations tend to increase during market crashes[1], correlation risk negatively affects investor welfare by making diversification more difficult in expensive states of nature. It is therefore natural to ask whether market-wide correlation risk carries a risk premium in the sense that assets that pay off well when market-wide correlations are higher than expected (thus providing a hedge against correlation risk) earn lower returns than can be justified by their exposure to other priced risk factors. Index options are an obvious example of such assets and will appear expensive when correlation risk is priced, as shown by Driessen, Maenhout, and Vilkov (2009) (henceforth DMV). In this companion paper, we are the first to provide a direct and intuitive test of the existence and significance of the correlation risk premium, by analyzing option-implied correlations and by comparing option-implied with realized correlations.

We contribute to the existing research in three important ways. First, we develop a simple theoretical model of stochastic correlations and of priced correlation risk and show how to extract implied correlations from equity options. We prove that the difference between implied and realized correlations reflects the price of correlation risk. Second, using data on index and individual options for all index constituents, we estimate the times-series of implied correlations for S&P500 and DJ30 samples over 17 years and find a significantly positive difference between implied and realized correlations for both samples, which is equivalent to a significantly negative price of correlation risk. Moreover, because individual variance risk is not priced on average, the correlation risk premium drives the index variance risk premium. Third, motivated by the recent findings that the variance risk premium has substantial predictive power for future aggregate market returns (e.g., Bollerslev, Tauchen, and Zhou (2009)), and knowing that correlation risk significantly contributes to variance risk, we study the predictive power of implied correlations for future market returns. We find that implied correlations explain future aggregate returns across various horizons from one day to one year, and that the predictive power is very large at 6-month and 1-year horizon. Interestingly, implied correlations and the index variance risk premium exhibit large joint significance when explaining market excess returns, with $R^{2\prime}s$ between 12% and 15%, depending on the sample and

[1] See for example Roll (1988), Jorion (2000) and Longin and Solnik (2001).

return horizon.

We start our analysis by developing a simple option pricing model that incorporates stochastic correlations and a correlation risk premium. Each individual stock price follows a standard Ito process, and the correlation between stock returns is stochastic and mean-reverting. The model produces endogenous stochastic index-return variance, even with constant individual variances. Moreover, we allow for a negative risk premium on correlation risk. Formally, the negative correlation risk premium generates higher expected correlation paths under the risk-neutral measure than under the actual measure. We prove that this divergence in expected correlations under the two measures results in option-implied correlations exceeding average realized correlations between equity returns. Another important implication of this model is that the risk-neutral expected integrated variance implied by index options is higher than the realized index variance.

We use two samples, namely a broad stock market index (S&P500) from 1996 until 2012, and a narrower index (DJ30), from 10/1997 until 2012, therefore also including the recent financial crisis. Using data on index options and on individual options on all the index components, combined with prices of the underlying stocks, we analyze the empirical content of implied correlations in several ways. First, we calculate each day an option-implied correlation from 30-day index and individual options. The time-series for this implied correlation is closely related to realized equity return correlations. Moreover, we show that implied correlations predict realized correlations, which supports the interpretation of the implied correlation as a measure of risk-neutral expected future correlations. Most importantly, we find a systematic difference between implied and realized correlations. The average S&P500 implied correlation is 39.5%, while realized correlations are on average 32.6%. This provides direct and intuitive evidence for a large correlation risk premium. The results for the DJ30 sample are even more striking with an average implied correlation of 46.0% and an average realized correlation of 35.5%. This offers an interesting perspective on the "overpriced index option puzzle":[2] index options are priced as if correlations between stocks in the index are on average more than 20% higher than seems historically the case.

While our main focus is the pricing of correlation risk in options and using implied and realized correlations to obtain a direct estimate of the correlation premium, we also study whether market-wide correlations have predictive power for market returns. This analysis is motivated by

[2] The relevant literature includes Bakshi and Kapadia (2003), Bondarenko (2003a), Bondarenko (2003b), Buraschi and Jackwerth (2001), Carr and Wu (2009), Coval and Shumway (2001), Jackwerth and Rubinstein (1996), Jones (2006), and Pan (2002). Bates (2003) surveys earlier work. Gabaix (2012) provides a theoretical explanation based on rare disasters.

recent work of Bollerslev, Tauchen, and Zhou (2009), who show that the market variance risk premium, defined as the difference between the model-free implied variance and the lagged realized variance has substantial predictive power for future market excess returns, especially at the quarterly return horizon. Similarly, Cosemans (2011) analyzes the predictive power of the correlation risk premium. We replicate their results for our samples, but also show as a new finding that implied correlations have in fact even larger predictive power and continue to do so for longer return horizons. For the S&P500 sample, the R^2 for predictive regressions of future market excess returns on implied correlations is 14.0% for 6-month returns and 11% for annual returns. At the 3-month and 6-month horizons we also document strong joint significance for both predictors (the market variance risk premium of Bollerslev, Tauchen, and Zhou (2009) and our implied correlation) with an R^2 of 14.9% for quarterly, and 15.4% for semi-annual horizons, respectively, when we use overlapping regressions, and R^2's of 26.6% and 22.5% when end-of-month regressions are used. We also include into predictive regressions with the market variance risk premium and implied correlation a number of fundamental variables shown to predict market returns (Goyal and Welch (2008), Ferreira and Santa-Clara (2011), among others), and the improvements in model R^2's are typically only marginal. Most importantly, the implied correlation proves to be robust predictor for all return horizons under consideration (from one month to one year), while for semi-annual and annual horizons inclusion of some variables render variance risk premium insignificant.

Several articles have investigated the correlation structure of interest rates of different maturities. Longstaff, Santa-Clara, and Schwartz (2001), De Jong, Driessen, and Pelsser (2004) and Han (2007) provide evidence that interest rate correlations implied by cap and swaption prices differ from realized correlations. Collin-Dufresne and Goldstein (2001) propose a term structure model where bond return correlations are stochastic. Campa and Chang (1998) and Lopez and Walter (2000) study the predictive content of implied correlations obtained from foreign exchange options for future realized correlations between exchange rates. Skinzi and Refenes (2004) describe how index and individual stock options can be used to find implied equity correlations for the Dow Jones Industrial Average index. They study the statistical properties and the dynamics of the implied correlation measure with one year of data, but do not analyze the key implications for index option pricing. In fact, none of these articles investigates or estimates a risk premium on correlation risk.

Many recent papers use the idea of correlation risk to address a variety of interesting research questions. Buraschi, Porchia, and Trojani (2010) model the stochastic covariance matrix to obtain a closed-form portfolio choice solution when facing correlation risk. Buraschi, Kosowski, and Trojani (2012) show that hedge funds may have different styles, but most of them exhibit exposure to the same market-wide correlation risk, quantified by the correlation risk factor suggested in Driessen, Maenhout, and Vilkov (2009). Kelly, Lustig, and van Nieuwerburgh (2012) use options on sector indices and compare with options on individual stocks in the indices to analyze and to quantify the implicit 'put' or government bailout guarantee for the financial sector during the 2007-2009 crisis. Buraschi, Trojani, and Vedolin (2011) study an equilibrium model with heterogeneous agents and correlation risk. Mueller, Stathopoulos, and Vedolin (2012) provide evidence of priced correlation risk in foreign exchange markets using the concepts of implied and realized correlations.

The paper is organized as follows. Section 1 presents the theoretical model of priced correlation risk. The data is described in section 2. We analyze option-implied and realized correlations empirically in section 3 and present evidence on implied versus realized variances, for both index and individual options. Section 4 studies the predictive power of implied correlations and implied variance, and associated risk premiums, for future stock market excess returns. Section 5 examines the robustness of our results, before we conclude in section 6. Appendices A and B contain propositions and proofs omitted from the main text.

1 A Model of Priced Correlation Risk

This section describes the model of priced correlation risk. We start by specifying the stochastic process for each stock that is included in the stock market index.

1.1 Individual Stock Price Processes

The stock market index is composed of N stocks. Under the physical probability measure P, the price of stock i, S_i, is assumed to follow an Ito process with expected return μ_i and possibly stochastic diffusion $\phi_i(t)$:

$$dS_i = \mu_i S_i dt + \phi_i S_i dB_i \tag{1}$$

where B_i is a standard scalar Wiener process. We omit time as an argument for notational convenience throughout, except when placing particular emphasis. The special case where $\phi_i(t)$ is constant simplifies (1) to the standard Black-Scholes set-up. More generally, the instantaneous variance $\phi_i^2(t)$ is an Ito process, driven by a standard scalar Wiener process B_{ϕ_i}, which is taken to be uncorrelated with B_j for all j.[3] Importantly, we assume that $\phi_i^2(t)$ follows the same process under the physical probability measure P as under the risk-neutral probability measure Q, so that individual variance risk is not priced. This assumption is motivated by tractability considerations, as well as by the empirical findings in DMV. We discuss the empirical validity of this assumption for our samples below in Section 3.2.

The instantaneous correlation between Wiener processes B_i and B_j for $i \neq j$ is modeled as:

$$E_t^P[dB_i dB_j] = E_t^Q[dB_i dB_j] = \rho_{ij}(t)\,dt. \tag{2}$$

The stochastic nature of this instantaneous correlation $\rho_{ij}(t)$ and especially the associated risk premium constitute the key innovations of this paper. We assume that a single state variable $\rho(t)$ drives all pairwise correlations:

$$\rho_{ij}(t) = \overline{\rho_{ij}}\rho(t). \tag{3}$$

Although more elaborate correlation dynamics with multiple state variables are certainly interesting, assuming a single state variable is a natural starting point for the analysis, given our interest in priced correlation risk. If the risk of correlation changes carries a risk premium, we expect this to be compensation for the risk of economy-wide correlation changes. Furthermore, we make the following homogeneity assumption:

$$\overline{\rho_{ij}} = 1, \ \forall i \text{ and } j. \tag{4}$$

While this assumption may seem restrictive, it allows us to capture correlation risk in the most parsimonious way and can be extended as shown by Buss and Vilkov (2012).

We impose as initial condition $\rho(0) \in (0,1)$. Under measure P, the correlation state variable $\rho(t)$, which is also the instantaneous correlation because of (4), is assumed to follow a mean-reverting process with long-run mean $\overline{\rho}$, mean-reversion parameter λ and diffusion parameter σ_ρ:[4]

$$d\rho = \lambda(\overline{\rho} - \rho)\,dt + \sigma_\rho\sqrt{\rho(1-\rho)}dB_\rho. \tag{5}$$

Because of the $\sqrt{\rho(1-\rho)}$ factor, the process is of the Wright-Fisher type, used extensively in

[3] This means that there is no asymmetric volatility or 'leverage effect' (Black (1976)) at the level of individual stock returns. However, our correlation-risk model nonetheless generates an endogenous leverage effect for index returns. This is consistent with the empirical findings of Dennis, Mayhew, and Stivers (2006), who obtain a much smaller leverage effect for individual stock returns than for the index.

[4] Collin-Dufresne and Goldstein (2001) discuss a similar model of correlation dynamics.

genetics (see e.g. Karlin and Taylor (1981)), and also in financial economics (Cochrane, Longstaff, and Santa-Clara (2003)). We multiply the diffusion parameter by $\sqrt{\rho\,(1-\rho)}$ to ensure that under certain parameter restrictions $\rho\,(t)$ remains between zero and one with probability 1, as is shown by lemma 1 in appendix B.[5] This guarantees the positive definiteness of the resulting correlation and variance-covariance matrices (proposition 1, appendix B).

1.2 Priced Correlation Risk

Priced correlation risk is introduced into the model by having the instantaneous correlation follow a similar mean-reverting process under the risk-neutral measure Q as in (5), but with a correlation risk premium proportional to κ subtracted from its drift:

$$d\rho = \left[\lambda\left(\overline{\rho}-\rho\right) - \kappa\sigma_\rho\sqrt{\rho\,(1-\rho)}\right]dt + \sigma_\rho\sqrt{\rho\,(1-\rho)}dB_\rho^Q. \tag{6}$$

A negative value for κ implies a negative correlation risk premium, so that the expected path of future correlations under Q exceeds the expected correlation path under P.[6] Assets that load positively on correlation risk earn a negative instantaneous excess return κ per unit of (standardized) correlation risk exposure. Intuitively, the negative excess return represents the insurance premium paid for assets that hedge against unexpected correlation increases.

If correlation risk is correlated with market risk, the total correlation risk premium κ includes compensation for market risk. A first component of κ therefore equals $\psi\eta$, where η denotes the Sharpe ratio for market risk and ψ is the correlation between the Brownian motion B_ρ driving $\rho\,(t)$ in (5) and the market risk factor in the pricing kernel. A negative value would be in line with empirical work that finds that correlations increase when prices decline.

DMV presents evidence that positive exposure to correlation risk is associated with negative excess returns relative to what is justified by exposure to market risk. Put differently, the empirical evidence suggests that the total correlation risk premium exceeds $\psi\eta$ or that the part of market-wide correlation risk that is orthogonal to stock market risk is also priced. In that case, correlation risk constitutes a second priced risk factor in the economy. Denoting the compensation for orthogonal correlation risk by κ_o, the total correlation risk premium κ can then be decomposed (through an

[5] We could have accommodated negative values for ρ, but since ρ is the average correlation between the stocks in the index, this generalization is not needed empirically.

[6] Lemma 2 (appendix B) states the conditions under which the process $\rho(t)$ remains within interval $(0,1)$ under measure Q. These conditions nest the ones listed in Lemma 1, so that, if satisfied (as we assume), measures P and Q are equivalent.

orthogonal decomposition of the Brownian motion dB_ρ) in a part driven by the equity risk premium and an orthogonal part as follows

$$\kappa = \psi\eta + \sqrt{1-\psi^2}\kappa_o. \tag{7}$$

While κ_o would be zero in a representative-agent model with i.i.d. consumption growth and CRRA preferences, a non-zero orthogonal correlation risk premium κ_o might arise in equilibrium in models with long-run risk for consumption growth, extending the equilibrium models of priced volatility risk of Bansal and Yaron (2004) and Bollerslev, Tauchen, and Zhou (2009) to a setting with multiple assets and stochastic correlations. Recent papers by Buraschi, Trojani, and Vedolin (2011) and Martin (2013) study correlation risk in equilibrium models.

1.3 Endogenous Stochastic Index Variance

Both the existence of correlation risk and of a correlation risk premium have fundamental implications for the pricing of index options, as they are driven by the variance of the stock market index. Even in a simplified version of our model with standard Black and Scholes (1973) dynamics for all individual stocks (the version of (1) with constant diffusion coefficients $\phi_i(t) = \phi_i$), we endogenously generate stochastic index variance. More generally, the variances of individual stock returns are stochastic and as described above. Denoting the drift of $\phi_i^2(t)$ by γ_i, and the diffusion scaling parameter (which will determine the "vol of vol") by ς_i, the instantaneous variance process of the individual stock return, under both P and Q, is

$$d\phi_i^2 = \gamma_i dt + \varsigma_i \phi_i dB_{\phi_i} \tag{8}$$

It is then possible to obtain the process for the variance of the stock market index endogenously generated by our model with correlation risk as presented in Proposition 2.

Proposition 2: *Define* $\nu_\phi \equiv \sum_{i=1}^{N} \sum_{j\neq i} w_i w_j \phi_i \phi_j$, $\iota_i \equiv \left(w_i^2 + \frac{1}{2}\rho \sum_{j\neq i} w_i w_j \frac{\phi_j}{\phi_i}\right)\phi_i \varsigma_i$ *and* $\delta_\phi \equiv \frac{1}{2}\sum_{i=1}^{N}\left[\left(2w_i^2 + \rho\sum_{j\neq i} w_i w_j \frac{\phi_j}{\phi_i}\right)\gamma_i - \frac{1}{4}\rho \sum_{j\neq i} w_i w_j \frac{\phi_j}{\phi_i}\varsigma_i^2\right]$. *Given a set of fixed index weights* $\{w_i\}$, *individual stock price processes given by (1), (2), (3), (4) and (8), and a correlation process given by (5) under* P *and by (6) under* Q, *the instantaneous index variance* ϕ_I^2 *under* P *follows*

$$d\phi_I^2 = \nu_\phi \left[\lambda\left(\bar{\rho} - \rho\right)dt + \sigma_\rho \sqrt{\rho\left(1-\rho\right)} dB_\rho\right] + \delta_\phi dt + \sum_{i=1}^{N} \iota_i dB_{\phi_i}. \tag{9}$$

Under Q, the instantaneous variance ϕ_I^2 of the stock market index follows

$$d\phi_I^2 = \nu_\phi \left[\left(\lambda \left(\bar{\rho} - \rho \right) - \kappa \sigma_\rho \sqrt{\rho \left(1 - \rho\right)} \right) dt + \sigma_\rho \sqrt{\rho \left(1 - \rho\right)} dB_\rho^Q \right] + \delta_\phi dt + \sum_{i=1}^{N} \iota_i dB_{\phi_i}. \tag{10}$$

The process for the index variance implied by our model inherits the $\sqrt{\rho\left(1-\rho\right)}$ factor in its diffusion term from the bounds imposed on the correlation process and is therefore also of the Wright-Fisher type. The terms δ_ϕ and ι_i stem from the non-priced individual volatility risk and vanish in the special case of constant ϕ_i's as in the Black-Scholes model. In subsequent sections we show empirically that correlation risk is indeed an important driver of stochastic index variance. In addition, because of our premise of non-priced individual variance risk, the sole difference between (9) and (10) is precisely the correlation risk premium. In other words, the model attributes the entire index variance risk premium to priced correlation risk.

1.4 Risk-Neutral Expected Average Correlation

We now introduce the concept of the risk-neutral expected average correlation. The idea is to infer the expected average future correlation between N stocks from the expected future variances of these N stocks and from the expected future variance of the stock index.

Consider the risk-neutral expected integrated variance of the return on asset a (where $a \in \{I, 1, ..., i, ..., N\}$) over a discrete interval of length τ starting at time t:

$$\sigma_a^2\left(t\right) = E_t^Q \left[\int_t^{t+\tau} \phi_a^2\left(s\right) ds \right] \tag{11}$$

From the definition of the market index and using the homogeneity assumption (4), this can be written out for $a = I$ as

$$\sigma_I^2\left(t\right) = E_t^Q \left[\int_t^{t+\tau} \sum_{i=1}^{N} w_i^2 \phi_i^2\left(s\right) ds \right] + E_t^Q \left[\int_t^{t+\tau} \sum_{i=1}^{N} \sum_{j \neq i} w_i w_j \phi_i\left(s\right) \phi_j\left(s\right) \rho\left(s\right) ds \right] \tag{12}$$

Rather than attempting to extract the entire path of future correlations $\rho\left(s\right)$ from (12), we aim to obtain a 'certainty equivalent' of the future stochastic correlations, which we call the *risk-neutral expected average correlation* ($RNEAC$) and which is constant over the $[t, t+\tau]$ time interval.

$RNEAC$ is defined as

$$RNEAC\left(t\right)\equiv\frac{E_{t}^{Q}\left[\int_{t}^{t+\tau}\phi_{I}^{2}\left(s\right)ds\right]-\sum\limits_{i=1}^{N}w_{i}^{2}E_{t}^{Q}\left[\int_{t}^{t+\tau}\phi_{i}^{2}\left(s\right)ds\right]}{\sum\limits_{i=1}^{N}\sum_{j\neq i}w_{i}w_{j}E_{t}^{Q}\left[\int_{t}^{t+\tau}\phi_{i}\left(s\right)\phi_{j}\left(s\right)ds\right]}\tag{13}$$

$RNEAC$ is a certainty equivalent to the correlation process $\rho(s)$ under the risk-neutral measure Q over the $[t, t+\tau]$ time interval in the sense that it yields the same risk-neutral expected integrated index variance as when the entire stochastic process for $\rho(s)$ is used.

Index and individual option prices contain information about the risk-neutral expected future variances $E_{t}^{Q}\left[\int_{t}^{t+\tau}\phi_{I}^{2}\left(s\right)ds\right]$ and $E_{t}^{Q}\left[\int_{t}^{t+\tau}\phi_{i}^{2}\left(s\right)ds\right]$, respectively. The risk-neutral expectation in the denominator, however, is not observed, since it requires the instantaneous volatilities $\phi_i(s)$ and $\phi_j(s)$. As an alternative to $RNEAC(t)$, we calculate instead the *implied correlation* $IC(t)$, defined as

$$IC\left(t\right)\equiv\frac{E_{t}^{Q}\left[\int_{t}^{t+\tau}\phi_{I}^{2}\left(s\right)ds\right]-\sum\limits_{i=1}^{N}w_{i}^{2}E_{t}^{Q}\left[\int_{t}^{t+\tau}\phi_{i}^{2}\left(s\right)ds\right]}{\sum\limits_{i=1}^{N}\sum_{j\neq i}w_{i}w_{j}\sqrt{E_{t}^{Q}\left[\int_{t}^{t+\tau}\phi_{i}^{2}\left(s\right)ds\right]}\sqrt{E_{t}^{Q}\left[\int_{t}^{t+\tau}\phi_{i}^{2}\left(s\right)ds\right]}}\tag{14}$$

This measure of the option-implied correlation is readily estimated and is useful for the following reasons. First, IC is closely related to $RNEAC$ and Lemma 3 (Appendix B) establishes that $IC(t) \leq RNEAC(t)$. Second, one of the main contributions of this paper is to provide direct evidence on the importance of a negative correlation risk premium. This is achieved by comparing estimates of $IC(t)$ with estimates of realized correlations. More precisely, we can take expectations under measure P rather than under Q in equation (14) so as to capture the actual (as opposed to risk-neutral) expected average correlation. Since this will later be estimated from the time-series average of the (cross-sectionally weighted average of) realized correlations, we call this in short the realized correlation $RC(t)$, defined as:

$$RC\left(t\right)\equiv\frac{E_{t}^{P}\left[\int_{t}^{t+\tau}\phi_{I}^{2}\left(s\right)ds\right]-\sum\limits_{i=1}^{N}w_{i}^{2}E_{t}^{P}\left[\int_{t}^{t+\tau}\phi_{i}^{2}\left(s\right)ds\right]}{\sum\limits_{i=1}^{N}\sum_{j\neq i}w_{i}w_{j}\sqrt{E_{t}^{P}\left[\int_{t}^{t+\tau}\phi_{i}^{2}\left(s\right)ds\right]}\sqrt{E_{t}^{P}\left[\int_{t}^{t+\tau}\phi_{i}^{2}\left(s\right)ds\right]}}\tag{15}$$

Proposition 3 states that the difference between IC and RC is crucially linked to the price of correlation risk.

Proposition 3: *Given a set of fixed index weights $\{w_i\}$, individual stock price processes given by (1), (2), (3) and (4) with non-priced individual variance risk, and a correlation process given by (5) under P and by (6) under Q, the difference between IC and RC can be written as*

$$IC(t) - RC(t) = \Xi(t)\left[\int_t^{t+\tau} \Psi(s)\left\{E_t^Q[\rho(s)] - E_t^P[\rho(s)]\right\} ds\right] \tag{16}$$

where $\Xi(t) > 0$ and $\Psi(s) > 0, \forall s$. $IC(t) - RC(t) = 0$ if and only if $\kappa = 0$ and is strictly increasing in -κ.

Testing whether correlation risk is priced can therefore be done by testing whether $IC = RC$. Note that the proposition requires that individual stock return variance risk is not priced. As can easily be seen from equations (14) and (15), $IC - RC$ is then proportional to $E_t^Q\left[\int_t^{t+\tau} \phi_I^2(s)\, ds\right] - E_t^P\left[\int_t^{t+\tau} \phi_I^2(s)\, ds\right]$, so that correlation risk is priced if and only if index variance is priced, in line of course with the result of proposition 2.

It follows from proposition 3 that when $\phi_i(t) = \phi_i$, $\forall i$, $IC(t) - RC(t)$ simplifies to

$$IC(t) - RC(t) = \frac{1}{\tau}\int_t^{t+\tau} \left\{E_t^Q[\rho(s)] - E_t^P[\rho(s)]\right\} ds \tag{17}$$

i.e., the difference in expected integrated correlation paths under Q and P. Interestingly, $IC(t)$ then gives the no arbitrage correlation swap rate, fixed at t and to be paid at $t + \tau$ in exchange for the realized integrated correlation path.

2 Data Description

Our sample period is from January 1996 until December 2012 and we use daily data on stock and index returns from the Center for Research in Security Prices (CRSP). For stock and index options we utilize IvyDB (OptionMetrics), which contains data on all U.S.-listed index and equity options. For the riskfree rate we use OptionMetrics data on the zero-coupon curve.

We consider two major US indices, namely the Standard & Poor's 500 (S&P500) and the Dow Jones 30 (DJ30) indices, as well as the underlying stocks making up these indices over our sample period (the DJ30 sample starts almost two years later, namely in October 1997). We obtain the daily composition of the S&P500 from the Compustat Index Constituents file, and the composition of DJ30 with all additions and deletions from the Dow Jones Indexes website. Using the previous day market capitalization of the stocks from CRSP we construct the value weights on each day for the S&P500 Index sample. We construct the price weights for the DJ30 Index sample from previous day CRSP stock prices.

We follow the methodology of Bakshi, Kapadia, and Madan (2003) (henceforth, BKM) to estimate the risk-neutral expected integrated variance $\sigma_a^2(t)$ defined in (11) from index options for $a = I$ and from individual options for $a = i$. Their procedure gives the correct estimate of the option-implied (i.e. risk-neutral) integrated variance over the life of the option contract when prices are continuous but volatility is stochastic, and is therefore labeled the *model-free implied variance* ($MFIV$). We use the standardized volatilities for maturities of 30, 60, and 91 days from OptionMetrics's Volatility Surface File, which contains a smoothed implied-volatility surface for a range of standard maturities and a set of option delta points. We select out-of-the-money implied volatilities for calls (delta smaller or equal to 0.5) and for puts (deltas larger than -0.5) to estimate on each day $MFIV$ for a given maturity using BKM's method.[7]

Given the time-series of model-free implied variances from index and individual options, as well as the index weights $\{w_i\}$, the implied correlation $IC(t)$ is calculated for each day t as

$$IC(t) = \frac{\sigma_{MF,I}^2(t) - \sum_{i=1}^{N} w_i^2 \sigma_{MF,i}^2(t)}{\sum_{i=1}^{N} \sum_{j \neq i} w_i w_j \sqrt{\sigma_{MF,i}^2(t)} \sqrt{\sigma_{MF,j}^2(t)}}. \tag{18}$$

Note that if there are missing data for variances on any given day, we simply rescale the weights of the available stocks to sum up to one. In the later years in our sample almost all the securities are always present.

To compute the realized moments of returns at day t we use daily returns for the indices and stocks from day $t+1$ to the end of the estimation window. For the realized variances we require at least 50% of the observations to be present for a given estimation window, i.e., for 30, 60, or 91 calendar days, and after computing the variance for the respective period we annualize it by linear rescaling.

[7] To compute the integrals that give the required values of the variance contracts precisely we need a continuum of option prices. We discretize the respective integrals and approximate them using the available options. We normally have 13 out-of-the-money call and put implied volatilities from the surface data for each maturity. Using cubic splines, we interpolate them inside the available moneyness range, and extrapolate using the last known (boundary for each side) value to fill in a total of 1001 grid points in the moneyness range from 1/3 to 3. Then we calculate the option prices from the interpolated volatilities using the known interest rate for a given maturity and use these prices to compute the model-free implied variance.

We obtain realized correlations RC between the stocks in the index from CRSP returns. Because standard realized correlations are more commonly used and because the difference between realized correlations calculated from (15) and from the standard definition is very small and not economically significant, we calculate standard realized correlations. For a given estimation window we also require that at least 50% of the observations be present for each stock; then we compute the correlations pairwise, i.e., for each pair of stocks we include the observations present for both stocks. After computing a correlation matrix for each day t, we aggregate the pairwise correlations into a cross-sectional weighted average across all pairs of stocks (using their index weights) to obtain $RC(t)$ for a given index and a given estimation window.

Following Carr and Wu (2009), we compute the Variance Risk Premium (VRP) for each day and each maturity by taking the difference between the respective implied and realized variances for each stock and index. Finally, following Bollerslev, Tauchen, and Zhou (2009) we also compute the value of the *ex ante* variance risk premium (VRP^{exa}) for the index proxy defined as the difference between the implied variance $MFIV$ observed on day t and the realized index variance over the last 30 calendar days.

3 Empirical Features of Correlations

We estimate implied correlations from option prices and show there is substantial correlation risk. We also investigate whether the implied correlation measure indeed captures the dynamics of historically observed correlations. We then test whether correlation risk is priced.

3.1 Implied versus Realized Correlations

Figure 1a plots the time-series of the 1-week moving averages of the implied correlation $IC(t)$ defined in (18) and of the weighted-average realized correlation for the S&P500. We use the S&P500 as the main sample, since it is a broad index that could be interpreted as a reasonable proxy for the stock market. The DJ30 index is narrower and less representative of the overall market, and therefore acts as a second sample to analyze for robustness reasons.[8] The implied correlation in Figure 1a is very volatile and peaks during financial crises. For instance the Asian crisis of 1997, the Russian crisis of 1998, the event of September 11, 2001, or the financial crisis of 2008 clearly show up as periods over which index options reflect high risk-neutral cross-stock

[8] The DJ30 index is more homogeneous and more likely to satisfy the homogeneity assumption in (4).

correlations. The fact that the 1-week moving average of the implied correlation fluctuates between 0.2 and 0.8 over our 17 year sample suggests that there is substantial correlation risk. Figure 1a shows that our implied correlation measure indeed captures the dynamics of the cross-sectional average of correlations between all stocks in the S&P500 index. The time-series correlation between the daily level of IC and of the realized correlation measure in Figure 1a is 59.65%. Figure 1b shows similar findings for the the DJ30 sample, with IC peaking at almost 0.9 during the Russian crisis. Implied and realized correlations are again highly correlated (time-series correlation of 55.85%), while implied correlations seem substantially higher in level than realized correlations, suggesting a large correlation risk premium.

Table 1 presents further information about the two series. Of particular interest is the extent to which the implied correlation IC exceeds the realized correlation RC, since the difference between IC and RC reflects the price of correlation risk κ, as shown in proposition 3. Table 1 shows that the time-series average for the realized 30-day correlation for S&P500 is 32.59%, while the time-series average for IC is 39.46%, which suggests a large risk premium for correlation risk. In other words, index options are on average priced as if the correlations between stocks in the index are more than 21% higher than seems historically the case.

For the DJ30 sample, correlations between stocks are on average higher, as would be expected for a narrower and more concentrated index. Also not surprisingly, the correlations, both implied and realized, are noisier and more volatile themselves. The correlation risk premium is now even larger; the average implied correlation of 46.03% and the average realized correlation of 35.53% generate an average "correlation risk premium" $IC - RC$ of more than 10% for the DJ30 sample.[9] The prices of DJ30 index and individual options reflect risk-neutral correlations that are on average about 30% higher than the realized correlations observed during our sample period.

The high time-series correlation between the implied correlation IC and the realized correlation (59.65% for S&P500 and 55.85% for DJ30) is reduced when considering quarterly changes rather than levels; the time-series correlation between IC and RC is then 36.6% for S&P500 and 35.0% for DJ30. To further investigate the relation between implied and realized correlations, we study whether IC can predict realized correlations. Table 2 reports the results from predictive regressions of cross-sectionally averaged realized correlations on lagged values of implied and of realized correlations. The predictive power of IC for realized correlation is quite high, especially when taking

[9]We follow the terminology of Bollerslev, Tauchen, and Zhou (2009) and others, who label $MFIV - RV$ the variance risk premium.

into account that realized correlations are themselves rather noisy estimates: IC explains 36% of the variation in future realized correlations for S&P500. Computing t-statistics based on Newey and West (1987) standard errors, the slope coefficient is statistically significant. Table 2 shows that lagged realized correlations also predict realized correlations with approximately the same R^2 of 37%. In the bivariate predictive regressions, the implied correlation continues to be a significant predictor. When using 60-day or 90-day predictors (IC or RC calculated over 60 or 90 days), the predictive power of implied correlations is somewhat smaller than for realized correlations, but continues to be economically and statistically significant. We obtain similar results for the DJ30 sample.

3.2 Implied versus Realized Variances

Our stylized model shows that priced correlation risk generates priced index variance risk, even when individual variance risk is not priced. We now shed light on this implication of the model. This section also serves to investigate the validity of our simplifying assumption of a zero risk premium for individual variance risk for our samples. Until recently, most papers on variance risk exclusively focused on index options and ignore individual options. Carr and Wu (2009) and DMV include individual options into the analysis and cannot reject the null hypothesis of a zero variance risk premium in individual options for the large majority of stocks considered. We repeat the analysis for our much larger and more recent sample. Recent papers by Goyal and Saretto (2009) and Schürhoff and Ziegler (2011)) study the risk premiums in the cross-section of individual options in greater detail and obtain interesting results that are complementary to our work since they zoom in on cross-sectional differences.

Figure 2 shows the time-series of (the square root of) the implied and realized variances for both the S&P500 index and the constituent stocks (weighted average of their variances), and DJ30 index and its constituents. Although all calculations are done for variances we take square roots of the computed variances for interpretation purposes. The negative index variance risk premium is apparent from the fact that implied index variance exceeds realized index variance almost throughout the sample. The key question is to what extent this is also true for individual options. Turning to Table 3 we find that the average (annualized) realized index volatility is 20.69% for the S&P500, while its $MFIV$ average is 23.09%. The null hypothesis that implied and realized variance are on average equal is strongly rejected, based on a t-test with Newey and West (1987) autocorrelation consistent standard errors for 21 lags. For individual options, the square root of realized variance (41.99%) is not statistically significantly different from the square root of the implied variance (43.38%). We can observe also that the longer the horizon is, the smaller is the point estimate of the average variance risk premium for individual stocks, and for 91-day maturity the average realized variance even exceeds the average implied variance, and the point estimate of the variance risk premium turns negative. For the index, however, the point estimate of the variance risk premium is very stable across different horizons and equals to around 1.05%. The null hypothesis that, on average across all stocks in the index, the implied and realized variance are equal, and that therefore the average variance risk premium is zero, cannot be rejected at any reasonable confidence level. More importantly, testing at the level of individual stocks in Table 4, the null of a zero variance risk premium for 30-day maturity is not rejected at the 5% confidence level for 503 stocks out of the 919 stocks that are at some point during the sample period part included in the S&P500 index. Although this means that there is some evidence against our assumption of a zero risk premium on individual variance risk, it can still be maintained as a first-order approximation that is a useful starting point for the analysis.

We find similar patterns for the DJ30 sample, with a statistically and economically significant index variance risk premium, but an insignificant average individual variance risk premium. Carrying out the test individually, the null of a zero individual variance risk premium is not rejected for 27 out of 43 stocks.

In summary, the underlying assumption of our stylized model of unpriced individual variance risk generally receives empirical support, and the support is stronger for longer horizons. The fact that we simultaneously obtain priced index variance risk points to priced correlation risk, consistent with the model of Section 1 and with the direct evidence of priced correlation risk in the previous subsection.

4 Predicting Stock Returns with Implied Correlations

Bollerslev, Tauchen, and Zhou (2009) (henceforth, BTZ) show that the (market) variance risk premium explains a substantial fraction of the time-series variation in aggregate stock market returns. The evidence of predictability is most pronounced at the quarterly return horizon and crucially relies on the simultaneous use of $MFIV$ and a high-frequency measure of RV to construct the variance risk premium. BTZ show that their empirical findings are consistent with a long-run risk equilibrium model that is extended to permit stochastic volatility of volatility for consumption growth (which equals dividend growth in the model). Their findings are interesting for several reasons. First, research on stock market predictability has generated a large and sometimes controversial literature, but recent work suggests that finding evidence for predictability is more difficult in the last two decades. Importantly, BTZ obtain their evidence of predictability for a sample spanning 1990 through 2007. Second, BTZ demonstrate that the predictability depends on the use of non-parametric option-implied variance, in particular the $MFIV$ used in the recent derivatives literature. Their work therefore presents an important link between this literature and the asset pricing research studying equity market predictability and time-series variation in expected equity premiums.

Because correlation risk is an important driver of variance risk, it is interesting to analyze whether implied correlations and correlation risk premiums might have any predictive power for aggregate equity market excess returns. This is a natural question, given the findings in DMV that the variance risk premium found in the cross-section of option returns is in fact attributable to priced correlation risk. It is therefore an important empirical question how the correlation risk premium fares in predictive regressions relative to the variance risk premium. Second, Pollet and Wilson (2010) find, partially motivated by DMV, that average realized correlations predict quarterly stock market excess returns. If realized correlations have predictive power for market returns, implied correlations would also (in fact *a fortiori*) be expected to contain relevant information. Our empirical analysis can provide direction for future research in equilibrium modeling as well. For example, the BTZ model could be extended to multiple Lucas trees with stochastic correlations, as in Buraschi, Trojani, and Vedolin (2011) or in Martin (2013). The correlation risk premium would then also drive dividend (and therefore consumption) growth volatility, as well as the volatility of dividend (and consumption) growth volatility, in the same way that the variance risk premium drives these quantities in the BTZ model. An alternative theoretical motivation for our analysis is the ICAPM of Merton (1973), since IC clearly matters to investors as a state variable, because its innovations drive future diversification benefits and market variance, i.e., future investment opportunities.

Like BTZ, we consider simple linear regressions of S&P500 excess returns on lagged predictor variables. As predictors we consider 1) IC, 2) $IC(t) - RC(t-1)$, which is analyzed in Cosemans (2011) and which can be viewed as the *ex ante* correlation risk premium, along the lines of the

VRP^{exa} as used by BTZ, 3) $MFIV$ and 4) $VRP^{exa} = MFIV(t) - RV(t-1)$. We consider predictive regressions for return horizons ranging from 1 day to 1 year and perform them in an overlapping fashion, i.e., we regress returns for a specified future horizon on a set of regressors observed every day in our sample.

Table 5 presents the results for the predictive regressions. At the 1-month return horizon, IC has predictive power and is statistically significant predictor for stock market returns, but with $R^2 = 2.53\%$ it explains slightly less variation than VRP^{exa} of BTZ with R^2 of 3.27%. All the coefficients in the regressions are positive, and high implied correlations are clearly associated with higher future returns. Turning to longer return horizons than 1 month, IC has striking predictive power and actually dominates the variance risk premium, despite the strong predictive power of the latter at the quarterly horizon, as found by BTZ and as also obtained here for our sample period. The R^2 of the univariate predictive regressions for the quarterly return horizon is 5.92% for the variance risk premium, but 9.58% for IC and grows to 13.97% for 6-month returns and is still 11.00% for annual returns. The variance risk premium on the other hand loses its significance at these longer horizons. While the correlation risk premium $IC - RC$ also demonstrates quite remarkable explanatory power for future market returns, the results are less strong than for IC and in fact less robust. In multivariate regressions combining IC and RC it is clear that IC is driving the predictive power, not RC (results not reported but available upon request). The univariate results are also easily seen in Figures 3a and 3b.

Interestingly, in bivariate regressions combining our predictor IC with BTZ's $VRP^{exa} = MFIV - RV$ we find very strong joint significance of the option-implied predictors; the point estimates and t-statistics are essentially unchanged relative to the univariate case and the R^2 is roughly the sum of the univariate $R^{2\prime}s$. This translates into almost 15% R^2 at the quarterly return horizon and 15.4% at the semi-annual return horizon, suggesting that implied correlations and the variance risk premium are complements in predicting market returns. From inspection of the univariate results for different horizons, one may conjecture that this could be due to the existence of multiple factors driving market-wide correlations and ultimately market variance and risk.

While BTZ demonstrate that the use of high-frequency data is important for the predictive power of the realized variance component of the variance risk premium, this turns out to be more challenging for correlations than for variance, since non-synchronicity can bias RC downwards severely, especially in the early years of our sample. This may explain why high-frequency data did not lead to improvements in our results and the lack of accurate high-frequency measures of realized correlations for large samples could be a potential reason why the correlation risk premium $IC - RC$ (as opposed to implied correlation IC) does not give us even stronger predictability results.

We also study predictability in the DJ30 sample. When interpreting these results, it is important to keep in mind that any differences in results with the S&P500 sample could have multiple sources. First, the DJ30 sample covers a slightly shorter time period, as the data starts in October 1997. Second, there is obviously the difference in the composition of the index and therefore also in the market excess return that is used on the left-hand side of the predictive regressions. We consider the excess return on the DJ30 index as the relevant aggregate stock market proxy when studying predictability for IC based on DJ30 index and individual options. Differences in our empirical results for predictability with S&P500 predictors versus with DJ30 predictors could therefore also be due to differences between the excess return on the DJ30 index and the S&P500 index.

Table 5 also presents the empirical results for the DJ30 sample. The variance risk premium has the highest explanatory power at the quarterly return horizon, as BZT found for the S&P500 sample, and the R^2's for variance risk premium and implied correlation are approximately the same and equal to 7.45%. However, as above we also see here that IC has even more explanatory power at the 6-month and at the 12-month return horizon, with $R^{2\prime}s$ of 12.1% and 12.2%.

Our finding above of very large joint significance for IC and the variance risk premium VRP^{exa} of BTZ in the S&P500 sample is also robust and continues to hold in the DJ30 sample. In bivariate predictive regressions the implied correlation and the variance risk premium jointly explain as much as 13.6% of the variation in quarterly returns and 14.9% for semi-annual returns. As we also obtained for the S&P500 sample, the variance risk premium is statistically and economically insignificant when predicting returns over a 1-year horizon though, and the R^2 of 12.49% is essentially the same as when only using IC as a predictor.

5 Robustness Checks

5.1 Different Sample Periods

To assess the robustness of our overall finding of a large correlation risk premium, we now repeat the analysis of Section 3 for 3 subsamples. This can also shed light on the extent to which the correlation risk premium driving a wedge between IC and RC can be seen as relatively constant, or instead subject to time-variation.

Table 6 reports the summary statistics for IC and RC for 3 subsamples of each 5 to 6 years (1996 to 2001, 2002 to 2007 and 2008 to 2012). The difference between IC and RC is extremely large in the first subsample, especially in the DJ sample where the implied correlation average of 43.41% is essentially twice the realized correlation (21.35%). For the S&P500 sample $IC - RC$ is also very large and economically significant. The difference between IC and RC is in fact always statistically significant, for any subsample, for any time-horizon and for both indices. The difference between implied and realized correlations becomes smaller when we go from the first sub period to the second, but grows again slightly when we move to the third sub period. While this is suggestive that the correlation risk premium may have become smaller over time, it remains nonetheless economically and statistically significant.

Turning to the results for implied and realized variances in the different subsamples, we see that the index variance risk premium is very large in the first subsample for both indices in Table 7, while most individual variance risk premiums are clearly insignificant in Table 8. Average variance risk premium for individual stocks in the S&P500 sample is positive and significant for 30-day horizon, but it loses its significance for 60-day horizon and even turns negative for 90 days. For the DJ30 sample we always get a negative average variance risk premium. This is of course consistent with the finding above of a very large correlation risk premium for the 1996 to 2001 subsample.

In the second subsample we still obtain a significant index variance risk premium in Table 9, but now there is some evidence against our simplifying assumption of no significant individual variance risk premiums in Table 10, though we still fail to reject the hypothesis that variance risk premium is not different from zero for about half of all the stocks in both indices. Finally, for the subsample including the recent financial crisis in Table 11, the index variance risk premium loses its significance due to the very high realized market variance; however, all point estimates for the average variance risk premium of individual stocks now turn negative for all horizons and for both indices.

Our results have interesting implications also for the term structure of correlation risk premiums, which may provide guidance when considering alternative theoretical specifications to model correlation risk and its dynamics. Throughout the paper we find that the difference between IC and RC grows with the maturity of the options used to calculate the implied correlations. This is the case in Table 1 for the overall sample, for both S&P500 and DJ30, and is also obtained in any of the subsamples as can be seen in Table 6.

5.2 Adding Fundamental Variables to Market Return Predictive Regressions

We also test how the option-based variables compare in predicting future market return with a number of fundamental variables, shown elsewhere to be successful predictors of market return. While there is a myriad of possible explanatory variables used in different studies (e.g., Goyal and Welch (2008), Ferreira and Santa-Clara (2011), among others), we limit our choice to only five of them, but so that they are largely non-redundant in terms of economic information they encompass. Specifically, we use the Earnings Price Ratio (EP), the Term Spread (TMS), the Default Yield Spread (DFY), the Book-to-Market Ratio (BTM), and the Net Equity Expansion ($NTIS$). We construct these variables from the data and following the procedures from the study of Goyal and Welch (2008).[10] EP is defined as the difference between the log of earnings and the log of prices; TMS is the difference between the long term yield on government bonds and the Treasury-bill; DFY is the difference between BAA and AAA-rated corporate bond yields; BTM is the ratio of book value to market value for the Dow Jones Industrial Average, and $NTIS$ is the ratio of 12-month moving sums of net issues by NYSE listed stocks divided by the total end-of-year market capitalization of NYSE stocks.

We perform a number predictive regressions, where we regress future market return, compounded for a given future period (one, three, six, or twelve months), on two option-based variables–implied correlation IC and variance risk premium $MFIV - RV$–and a number of other predictors specified above. Because the fundamental variables are available only at the end of each month, we sample the market returns and predictors once a month instead of doing it every day as in Section 4. From the results of the regressions in Table 13 we can see that implied correlation and variance risk premium do a very good job in explaining future returns, with a maximum R^2 of 26.6% achieved with the 3-month horizon returns. Both variables stay significant for all considered horizons. Adding fundamental variables does not have any effect on R^2 of monthly returns; moreover, neither coefficient nor significance of option-based variables change. For quarterly returns EP, TMS and DFY jointly add 2% to the R^2, though only the the Term Spread is significant at 10% level. Adding BTM and $NTIS$ (and removing DFY, because it is highly correlated with

[10] We are grateful to Amit Goyal for maintaining the updated data and making it available on his web-site `http://www.hec.unil.ch/agoyal/`.

$NTIS$ in our sample and causes mild collinearity problems) adds 5% more to the R^2, but has no effect on the significance of the option-based variables. For longer return horizons having four fundamental variables in regressions leads to variance risk premium first losing its significance and then even changing sign for an annual horizon. Overall we see that implied correlation proves to be the robust and significant predictor of future market returns, for the horizons from one month to one year.

6 Conclusion

We develop a parsimonious model of priced correlation risk for stock returns and implement it empirically using data on S&P500, as well as DJ30 index options and options on all index components. We demonstrate that correlation risk carries a large negative risk premium in our 1996-2012 sample, as evidenced by the substantial gap between average option-implied correlations (39.5% for S&P500 and 46.0% for DJ30) and average realized correlations (32.6% for S&P500 and 35.5% for DJ30). In our model this correlation risk premium drives the index variance risk premium that is widely documented in the literature.

Analyzing implied correlations gives a direct and intuitive test of whether correlation risk is priced, but it also offers an alternative interpretation of the "overpriced index option puzzle": index options seem indeed expensive in the sense that the risk-neutral expected average correlation embedded in their prices is substantially higher than average realized correlations, but this high price simply reflects the cost of insurance against the risk of unexpected correlation increases and the ensuing loss in diversification benefits.

As another important contribution we show that implied correlations have remarkable predictive power for future stock market excess returns, especially at the 6-month and 1-year horizons. Together with variance risk premiums, implied correlations explain as much as 15% of the observed variation in future aggregate stock returns.

Appendix A

This appendix contains the proof of Proposition 2.

Proof of Proposition 2: As individual variance risk is not priced, we have $E_t^Q\left[\int_t^{t+\tau}\phi_i^2(s)\,ds\right]=E_t^P\left[\int_t^{t+\tau}\phi_i^2(s)\,ds\right],\forall i$. To compare $IC(t)$ with $RC(t)$ it suffices to compare the part of the numerators of the definition:

$$
\begin{aligned}
IC(t)-RC(t) &= \Xi(t)\left(E_t^Q\left[\int_t^{t+\tau}\phi_I^2(s)\,ds\right]-E_t^P\left[\int_t^{t+\tau}\phi_I^2(s)\,ds\right]\right)\\
&= \Xi(t)\left[\int_t^{t+\tau}\sum_{i=1}^N\sum_{j\neq i}w_iw_j\left\{E_t^Q\left[\phi_i(s)\,\phi_j(s)\,\rho(s)\right]-E_t^P\left[\phi_i(s)\,\phi_j(s)\,\rho(s)\right]\right\}ds\right]\\
&= \Xi(t)\left[\int_t^{t+\tau}\Psi(s)\left\{E_t^Q\left[\rho(s)\right]-E_t^P\left[\rho(s)\right]\right\}ds\right]\\
\textit{where}\quad \Xi(t) &\equiv \frac{1}{\sum_{i=1}^N\sum_{j\neq i}w_iw_j\sqrt{E_t^P\left[\int_t^{t+\tau}\phi_i^2(s)\,ds\right]}\sqrt{E_t^P\left[\int_t^{t+\tau}\phi_i^2(s)\,ds\right]}}>0\\
\Psi(s) &\equiv \sum_{i=1}^N\sum_{j\neq i}w_iw_jE_t^P\left[\phi_i(s)\,\phi_j(s)\right]>0,\ \forall s.
\end{aligned}
$$

We use the Fubini Theorem to change the order of integration for the second equality and exploit that $\phi_i(s)\,\phi_j(s)$ is independent of $\rho(s)$ for the third equality.

If correlation risk is not priced ($\kappa=0$), then $E_t^Q[\rho(s)]=E_t^P[\rho(s)]$, so that $IC-RC=0$. To prove that $IC-RC$ is increasing in -κ, consider IC_{κ_1} and IC_{κ_2}, which are defined as in (14) for different equivalent martingale measures Q_1 and Q_2, each constructed from κ_1 and κ_2, respectively, where $\kappa_1<\kappa_2$. It suffices then to prove that $IC_{\kappa_1}>IC_{\kappa_2}$. Using the arguments above, we find:

$$
IC_{\kappa_1}(t)-IC_{\kappa_2}(t)=\Xi(t)\left[\int_t^{t+\tau}\Psi(s)\left\{E_t^{Q_1}\left[\rho(s)\right]-E_t^{Q_2}\left[\rho(s)\right]\right\}ds\right].
$$

The (normalized) kernel is $\xi_i(t)=e^{-\int_0^t\left(\eta_i dB_M+\kappa_{i,o}dB_{\rho,o}\right)-\frac{1}{2}\int_0^t\left(\eta_i^2+\kappa_{i,o}^2\right)ds}$, $i=1,2$, where dB_M and $dB_{\rho,o}$ are the independent Wiener processes driving stock market risk and orthogonal correlation risk, respectively. Define $f(t)\equiv\frac{\xi_2(t)}{\xi_1(t)}$. Under Q_1, f is a martingale and it is straightforward to show that $\frac{df}{f}=(\eta_1-\eta_2)\,dB_M^{Q_1}+(\kappa_{1,o}-\kappa_{2,o})\,dB_{\rho,o}^{Q_1}$. Then, using $\kappa_i=\psi\eta_i+\sqrt{1-\psi^2}\kappa_{i,o}$,

$$\begin{aligned}
E_t^{Q_1}[\rho(s)] - E_t^{Q_2}[\rho(s)] &= E_t^{Q_1}[\rho(s)] - f(t)^{-1} E_t^{Q_1}[f(s)\rho(s)] \\
&= E_t^{Q_1}[\rho(s)] - f(t)^{-1} E_t^{Q_1}[f(s)] E_t^{Q_1}[\rho(s)] - f(t)^{-1} cov_t^{Q_1}(f(s), \rho(s)) \\
&= -f(t)^{-1} \int_t^s E_t^{Q_1}[df(u)\, d\rho(u)]\, du \\
&= -f(t)^{-1} (\kappa_1 - \kappa_2)\, \sigma_\rho \int_t^s E_t^{Q_1}\left[f(u)\sqrt{\rho(u)(1-\rho(u))}\right] du \\
&= (\kappa_2 - \kappa_1)\, \sigma_\rho \int_t^s E_t^{Q_2}\left[\sqrt{\rho(u)(1-\rho(u))}\right] du > 0
\end{aligned}$$

Finally, since $\Xi(t) > 0$ and $\Psi(s) > 0$, $\forall s$, it follows that $IC_{\kappa_1} - IC_{\kappa_2} > 0$ if and only if $\kappa_1 < \kappa_2$, so that $IC - RC$ is strictly increasing in -κ. ■

Appendix B

This appendix elaborates some aspects of the analysis in section 1: lemmas 1 through 3 and proposition 1.[11]

Lemma 1: *The correlation state variable* $\rho(t)$ *following (5) with initial condition* $\rho(0) \in (0,1)$ *remains within interval* $(0,1)$ *with probability* 1 *if* $\lambda\overline{\rho} > \sigma_\rho^2/2$ *and* $\lambda(1-\overline{\rho}) > \sigma_\rho^2/2$, *and remains within interval* $[0,1)$ *with probability* 1 *if* $\lambda(1-\overline{\rho}) > \sigma_\rho^2/2$.

Proof of Lemma 1: Starting with the lower boundary, the solution to the Feller diffusion

$$d\rho = \lambda(\overline{\rho} - \rho)\, dt + \sigma_\rho \sqrt{\rho} dB \tag{19}$$

never reaches 0 from a strictly positive initial condition $\rho(0)$ with probability 1 under the Feller condition $\lambda\overline{\rho} > \sigma_\rho^2/2$. As the drift term $\lambda(\overline{\rho} - \rho)\, dt$ of (19) is strictly positive when $\rho \to 0$, it dominates the diffusion term $\sigma_\rho\sqrt{\rho}dB$ (in the sense that $|\lambda(\overline{\rho} - \rho)\, dt| \geq |\sigma_\rho\sqrt{\rho}dB|$ with probability 1) when $\rho \to 0$ and the Feller condition is satisfied. As $|\sigma_\rho\sqrt{\rho}\sqrt{1-\rho}dB| \leq |\sigma_\rho\sqrt{\rho}dB|$ for $\rho \to 0$, the diffusion term of (5) is also dominated by the drift, and hence the solution to (5) never reaches 0 from a strictly positive initial condition $\rho(0)$ with probability one if $\lambda\overline{\rho} > \sigma_\rho^2/2$.

To show that the solution to (5) never reaches the upper boundary we show that 0 is never

[11] We thank Zhipeng Zhang for suggestions on the proofs in Appendix B.

reached by $1-\rho$, which follows

$$d(1-\rho)=\lambda((1-\overline{\rho})-(1-\rho))\,dt-\sigma_{\rho}\sqrt{\rho}\sqrt{1-\rho}dB \tag{20}$$

For this SDE we can also look at the Feller diffusion for $1-\rho$:

$$d(1-\rho)=\lambda((1-\overline{\rho})-(1-\rho))\,dt-\sigma_{\rho}\sqrt{1-\rho}dB \tag{21}$$

This diffusion never reaches 0 from a strictly positive initial condition $1-\rho(0)$ with probability 1 under the Feller condition $\lambda(1-\overline{\rho})>\sigma_{\rho}^{2}/2$. Following the logic for the lower boundary case, we show that the solution to (20) is bounded away from 0 with probability one if $\lambda(1-\overline{\rho})>\sigma_{\rho}^{2}/2$. ■

Proposition 1: *The stochastic correlation matrix* $\Omega(t)=\left[\rho_{ij}(t)\right]$, *where* $i,j=1,...,N$, *with elements* $\rho_{ij}(t)=1$ *for* $i=j$, *and* $\rho_{ij}(t)=\rho(t)$ *for* $i\neq j$, *is positive definite for all* t *if* $\rho(t)\in[0,1)$. *The instantaneous variance-covariance matrix* $\Sigma(t)$ *of the* $N-$*dimensional stock price process with positive definite correlation matrix* $\Omega(t)$ *and instantaneous variance for any stock* $E[d\langle S_i\rangle]=\phi_i^2(t)\,dt$ *with* $\phi_i(t)\neq 0$, $\forall i$, *is positive definite for all* t.

Proof of Proposition 1: Take the correlation matrix $\Omega(t)=\left[\rho_{ij}(t)\right]$ at time t and omit the time argument. We can rewrite $\Omega=(1-\rho)\mathbf{I}+\rho\mathbf{11'}$, where $\mathbf{I}$ is the identity matrix and $\mathbf{1}$ is a column vector of ones. Let $X=(x_i)$ be a column vector with at least one element different from 0. Then the quadratic form created from this vector and the correlation matrix is always positive and the matrix Ω is positive definite by definition:

$$X'\Omega X=(1-\rho)X'\mathbf{I}X+\rho X'\mathbf{11'}X=(1-\rho)\sum_{i=1}^{n}x_i^2+\rho\left(\sum_{i=1}^{n}x_i\right)^2>0 \tag{22}$$

The variance matrix $\Phi=\left[\phi_i^2\right]$, where $i=1,...,N$, has positive diagonal elements $\phi_i^2(t)$ and zero off-diagonal elements. Next define A as the symmetric positive definite matrix with zero off-diagonal elements and ϕ_i on the diagonal, so that $\Phi=AA^T$. Then the instantaneous variance-covariance matrix $\Sigma=A\Omega A^T$. Pre- and post-multiplying this matrix by any $N\times 1$ vector $w\neq\mathbf{0}$ gives $w^TA\Omega A^Tw=w_1^T\Omega w_1>0$, where $w_1^T=w^TA=\left(A^Tw\right)^T$, as Ω is positive definite. ■

Lemma 2: *For* $\kappa<0$, *the correlation state variable* $\rho(t)$ *following (6) with initial condition* $\rho(0)\in(0,1)$ *remains within interval* $(0,1)$ *with probability* 1 *if* $\lambda\overline{\rho}>\sigma_{\rho}^{2}/2$ *and* $\lambda\left(1-\overline{\rho}+\frac{\kappa\sigma_{\rho}\sqrt{\varepsilon(1-\varepsilon)}}{\lambda}\right)>$

$\sigma_\rho^2/2$ *for some arbitrarily small positive* ε, *and remains within interval* $[0,1)$ *with probability* 1 *if* $\lambda\left(1-\overline{\rho}+\frac{\kappa\sigma_\rho\sqrt{\varepsilon(1-\varepsilon)}}{\lambda}\right) > \sigma_\rho^2/2$ *for some arbitrarily small positive* ε.

Proof of Lemma 2: The instantaneous correlation process under Q can be written as:

$$d\rho = \lambda\left(\overline{\rho} - \frac{\kappa\sigma_\rho\sqrt{\rho(1-\rho)}}{\lambda} - \rho\right)dt + \sigma_\rho\sqrt{\rho(1-\rho)}dB_\rho^Q \tag{23}$$

We see that $\overline{\rho}^* = \overline{\rho} - \frac{\kappa\sigma_\rho\sqrt{\rho(1-\rho)}}{\lambda}$ approaches the initial long-run mean $\overline{\rho}$ from above when ρ approaches 0 or 1, and the initial process (5) approaches the boundary 0 faster (informally) than the process under Q. Then 0 is a natural boundary for the process under Q if $\lambda\overline{\rho} > \sigma_\rho^2/2$.

For the upper boundary we modify Lemma 1's proof slightly. Approximate process (23) as

$$d\rho = \begin{cases} \lambda\left(\overline{\rho} - \frac{\kappa\sigma_\rho\sqrt{\rho(1-\rho)}}{\lambda} - \rho\right)dt + \sigma_\rho\sqrt{\rho(1-\rho)}dB_\rho^Q \text{ if } 1-\rho > \varepsilon \\ \lambda\left(\overline{\rho} - \frac{\kappa\sigma_\rho\sqrt{\varepsilon(1-\varepsilon)}}{\lambda} - \rho\right)dt + \sigma_\rho\sqrt{\rho(1-\rho)}dB_\rho^Q \text{ if } 1-\rho \leq \varepsilon \end{cases} \tag{24}$$

where ε is a very small positive number

The continuous process ρ, while approaching its boundary 1, passes through the point where the distance to the boundary is ε. At this point the nonlinearly increasing (for this range of ρ) part of the drift due to the change of measure $\frac{\kappa\sigma_\rho\sqrt{\rho(1-\rho)}}{\lambda}$ becomes a constant $\frac{\kappa\sigma_\rho\sqrt{\varepsilon(1-\varepsilon)}}{\lambda}$. This new process remains continuous.

For any ρ closer to the upper boundary 1 than ε the following is true:

$$\frac{\kappa\sigma_\rho\sqrt{\varepsilon(1-\varepsilon)}}{\lambda} \leq \frac{\kappa\sigma_\rho\sqrt{\rho(1-\rho)}}{\lambda} \text{ for } \frac{\kappa\sigma_\rho}{\lambda} < 0$$

Hence for $\rho \in [1-\varepsilon, 1)$ the 'long-run mean' of the approximating correlation process will be higher than any value of the long-run mean of the original process:

$$\overline{\rho} - \frac{\kappa\sigma_\rho\sqrt{\varepsilon(1-\varepsilon)}}{\lambda} \geq \overline{\rho} - \frac{\kappa\sigma_\rho\sqrt{\rho(1-\rho)}}{\lambda}$$

It follows that the approximating process is attracted more strongly (formally measured by the speed of convergence) to the upper boundary 1 than the original process.

The sufficient conditions under which the value 1 is a natural boundary[12] for the approximating

[12] Here we deviate slightly from the definition of natural (Feller) boundary given in Karlin and Taylor (1981). We do not show that the process cannot be started from the boundary, instead, we assume that.

process (24) are also sufficient for 1 being an unattainable boundary for the original process (5).

The sufficient conditions for the approximating process follow from the derivation of those for the original correlation process under P. We just need to replace $\overline{\rho}$ by a new long-run mean $\overline{\rho} - \frac{\kappa\sigma_\rho\sqrt{\varepsilon(1-\varepsilon)}}{\lambda}$ for a small positive number ε. Then the boundary 1 is unattainable if, for some arbitrarily small positive ε, $\lambda\left(1-\overline{\rho}+\frac{\kappa\sigma_\rho\sqrt{\varepsilon(1-\varepsilon)}}{\lambda}\right) > \sigma_\rho^2/2$. ■

Lemma 3: *Given a set of fixed index weights $\{w_i\}$ and strictly positive volatility processes $\phi_i(t)$, $\forall i$, the relationship between $RNEAC(t)$ defined in (13) and $IC(t)$ defined in (14) is given by:*

$$IC(t) = RNEAC(t)\,\frac{\sum\limits_{i=1}^{N}\sum_{j\neq i} w_i w_j E_t^Q\left[\int_t^{t+\tau}\phi_i(s)\,\phi_j(s)\,ds\right]}{\sum\limits_{i=1}^{N}\sum_{j\neq i} w_i w_j \sqrt{E_t^Q\left[\int_t^{t+\tau}\phi_i^2(s)\,ds\right]}\sqrt{E_t^Q\left[\int_t^{t+\tau}\phi_i^2(s)\,ds\right]}} \leq RNEAC(t)$$

Proof of Lemma 3: From the definitions,

$$IC(t) = RNEAC(t)\,\frac{\sum\limits_{i=1}^{N}\sum_{j\neq i} w_i w_j E_t^Q\left[\int_t^{t+\tau}\phi_i(s)\,\phi_j(s)\,ds\right]}{\sum\limits_{i=1}^{N}\sum_{j\neq i} w_i w_j \sqrt{E_t^Q\left[\int_t^{t+\tau}\phi_i^2(s)\,ds\right]}\sqrt{E_t^Q\left[\int_t^{t+\tau}\phi_i^2(s)\,ds\right]}}. \tag{25}$$

Therefore, it suffices to determine the following inequality:

$$E_t^Q\left[\int_t^{t+\tau}\phi_i(s)\,\phi_j(s)\,ds\right] \overset{?}{\leq} \sqrt{E_t^Q\left[\int_t^{t+\tau}\phi_i^2(s)\,ds\right]}\sqrt{E_t^Q\left[\int_t^{t+\tau}\phi_j^2(s)\,ds\right]}$$

We apply the Cauchy-Bunyakovsky-Schwarz inequality twice to obtain:

$$E_t^Q\left[\int_t^{t+\tau}\phi_i(s)\,\phi_j(s)\,ds\right] \leq E_t^Q\left[\sqrt{\int_t^{t+\tau}\phi_i^2(s)\,ds}\sqrt{\int_t^{t+\tau}\phi_j^2(s)\,ds}\right] \tag{26}$$

and

$$E_t^Q\left[\sqrt{\int_t^{t+\tau}\phi_i^2(s)\,ds}\sqrt{\int_t^{t+\tau}\phi_j^2(s)\,ds}\right] \leq \sqrt{E_t^Q\left[\int_t^{t+\tau}\phi_i^2(s)\,ds\right]}\sqrt{E_t^Q\left[\int_t^{t+\tau}\phi_j^2(s)\,ds\right]} \tag{27}$$

from which the result follows directly. ■

References

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

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, 101–143.

Bansal, R., and A. Yaron, 2004, "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *The Journal of Finance*, 59, 1481–1509.

Bates, D. S., 2003, "Empirical option pricing: a retrospection," *Journal of Econometrics*, 116, 387–404.

Black, F., 1976, "Studies of Stock Price Volatility Changes," *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economics Statistics Section*, pp. 177–181.

Black, F., and M. Scholes, 1973, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81, 637–654.

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

Bondarenko, O., 2003a, "Statistical Arbitrage and Securities Prices," *Review of Financial Studies*, 16, 875–919.

Bondarenko, O., 2003b, "Why are Puts so Expensive?," Working Paper, University of Illinois.

Buraschi, A., and J. Jackwerth, 2001, "The price of a smile: hedging and spanning in option markets," *Review of Financial Studies*, 14, 495–527.

Buraschi, A., R. Kosowski, and F. Trojani, 2012, "When There is No Place to Hide - Correlation Risk and the Cross-Section of Hedge Fund Returns," Working Paper.

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

Buraschi, A., F. Trojani, and A. Vedolin, 2011, "When Uncertainty Blows in the Orchard: Comovement and Equilibrium Volatility Risk premiums," Working Paper.

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

Campa, J. M., and P. H. K. Chang, 1998, "The forecasting ability of correlations implied in foreign exchange options," *Journal of International Money and Finance*, 17, 855 – 880.

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

Cochrane, J. H., F. A. Longstaff, and P. Santa-Clara, 2003, "Two Trees: Asset Price Dynamics Induced by Market Clearing," NBER Working Papers 10116, National Bureau of Economic Research, Inc.

Collin-Dufresne, P., and R. S. Goldstein, 2001, "Stochastic Correlation and the Relative Pricing of Caps and Swaptions in a Generalized-Affine Framework," Working Paper Carnegie Mellon University.

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

Coval, J. D., and T. Shumway, 2001, "Expected Option Returns," *The Journal of Finance*, 56, pp. 983–1009.

De Jong, F., J. Driessen, and A. Pelsser, 2004, "On the Information in the Interest Rate Term Structure and Option Prices," *Review of Derivatives Research*, 7, 99–127.

Dennis, P., S. Mayhew, and C. Stivers, 2006, "Stock Returns, Implied Volatility Innovations, and the Asymmetric Volatility Phenomenon," *Journal of Financial and Quantitative Analysis*, 41, 381–406.

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

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

Gabaix, X., 2012, "Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance," *The Quarterly Journal of Economics*, 127, 645–700.

Goyal, A., and A. Saretto, 2009, "Cross-Section of Option Returns and Volatility," *Journal of Financial Economics*, 94, 310–326.

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

Han, B., 2007, "Stochastic Volatilities and Correlations of Bond Yields," *The Journal of Finance*, 62, 1491–1524.

Jackwerth, J., and M. E. Rubinstein, 1996, "Recovering Probability Distributions from Option Prices," *The Journal of Finance*, 51, 1611–1631.

Jones, C. S., 2006, "A Nonlinear Factor Analysis of S&P 500 Index Option Returns," *Journal of Finance*, 61, 2325 – 2363.

Jorion, P., 2000, "Risk Management Lessons from Long-Term Capital Management," *European Financial Management*, pp. 277–300.

Karlin, S., and H. Taylor, 1981, *A Second Course in Stochastic Processes*, Academic Press.

Kelly, B., H. Lustig, and S. van Nieuwerburgh, 2012, "Too-Systemic-To-Fail: What Option Markets Imply About Sector-wide Government Guarantees," CEPR Discussion Papers 9023, C.E.P.R. Discussion Papers.

Longin, F., and B. Solnik, 2001, "Extreme Correlation of International Equity Markets," *The Journal of Finance*, 56, 649–676.

Longstaff, F. A., P. Santa-Clara, and E. S. Schwartz, 2001, "The Relative Valuation of Caps and Swaptions: Theory and Empirical Evidence," *The Journal of Finance*, 56, pp. 2067–2109.

Lopez, J. A., and C. Walter, 2000, "Is implied correlation worth calculating? Evidence from foreign exchange options and historical data.," *Journal of Derivatives*, 7, 65–82.

Martin, I., 2013, "The Lucas Orchard," *Econometrica*, 81, 55–111.

Merton, R. C., 1973, "An Intertemporal Capital Asset Pricing Model," *Econometrica*, 41, 867–888.

Mueller, P., A. Stathopoulos, and A. Vedolin, 2012, "International Correlation Risk," Working Paper, LSE.

Newey, W. K., and K. D. West, 1987, "A simple, positive-semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix," *Econometrica*, 55, 703–708.

Pan, J., 2002, "The jump-risk premiums implicit in options: evidence from an integrated time-series study," *Journal of Financial Economics*, 63, 3–50.

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

Roll, R., 1988, "The International Crash of October 1987," *Financial Analysts Journal*, 44, 19–35.

Schürhoff, N., and A. Ziegler, 2011, "Variance risk, financial intermediation, and the cross-section of expected option returns," CEPR Discussion Paper no. 8268.

Skinzi, V. D., and A.-P. N. Refenes, 2004, "Implied Correlation Index: A New Measure of Diversification," *Journal of Futures Markets*, 25, 171–197.

Table 1: Implied and Realized Correlations: Summary Statistics

The table reports summary statistics (time-series mean, median, 10^{th} and 90^{th} percentiles, and standard deviation) for the implied correlation (IC), realized correlation (RC), and for the difference between them (IC-RC), for two samples of stocks—components of S&P500, and DJ30 indices, for the sample period from 1996 to 12/2012, and from 10/1997 to 12/2012, respectively, and for three different maturities—30, 60, and 91 (calendar) days. $IC(t)$ is calculated from daily observations of model-free implied variances for the index and for all index components, using (14). $RC(t)$ is a cross-sectional weighted average (using the appropriate weights from the respective index) of all historical pairwise correlations at time t, each calculated over a 30-, 60-, or 91-day (calendar) window of daily stock returns.

	IC			RC			$IC - RC$		
	30	60	91	30	60	91	30	60	91
S&P500 Sample									
Mean	0.3946	0.4145	0.4326	0.3259	0.3254	0.3262	0.0687	0.0891	0.1064
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Median	0.3839	0.4150	0.4393	0.2967	0.2994	0.3051	0.0684	0.0911	0.1084
10^{th} Prctile	0.2331	0.2524	0.2710	0.1629	0.1806	0.1861	-0.0707	-0.0537	-0.0491
90^{th} Prctile	0.5675	0.5764	0.5846	0.5278	0.5040	0.4902	0.2179	0.2339	0.2548
StDev	0.1297	0.1239	0.1225	0.1436	0.1324	0.1270	0.1234	0.1257	0.1289
DJ30 Sample									
Mean	0.4603	0.4828	0.4977	0.3553	0.3544	0.3557	0.1050	0.1284	0.1420
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Median	0.4503	0.4820	0.5016	0.3340	0.3393	0.3416	0.1019	0.1201	0.1330
10^{th} Prctile	0.2643	0.2895	0.3125	0.1584	0.1649	0.1678	-0.0667	-0.0466	-0.0356
90^{th} Prctile	0.6600	0.6656	0.6674	0.5711	0.5396	0.5399	0.2809	0.3166	0.3387
StDev	0.1558	0.1562	0.1459	0.1601	0.1488	0.1440	0.1484	0.1622	0.1596

Table 2: Predicting Realized Correlations

The table shows the results of of predictive regressions of 30-day realized correlation on current implied correlation (IC) and on lagged realized correlation (RC), both univariately and bivariately, for two samples of stocks—components of S&P500, and DJ30 indices, for the sample period from 1996 to 12/2012, and from 10/1997 to 12/2012, respectively, and for predictors computer for three different horizons—30, 60, and 91 (calendar) days. $IC(t)$ is calculated from daily observations on model-free implied variances for the index and for all index components, using (14). $RC(t)$ is a cross-sectional weighted average (using the appropriate weights from the respective index) of all historical pairwise correlations at time t, each calculated over a 30-, 60-, or 91-day (calendar) window of daily stock returns. The p-values are based on Newey-West (1987) autocorrelation consistent standard errors with lags equal to the number of overlapping observations (21).

	30-day predictors			60-day predictors			91-day predictors		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
S&P500 Sample									
RC	0.6011	-	0.3665	0.6016	-	0.4250	0.6032	-	0.4649
	(0.00)		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)
IC	-	0.6604	0.3673	-	0.6439	0.3017	-	0.6136	0.2508
		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.01)
R^2	36.89	35.56	42.27	36.94	30.75	40.50	37.25	27.23	39.83
DJ30 Sample									
RC	0.6114	-	0.4310	0.6124	-	0.5041	0.6136	-	0.5145
	(0.00)		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)
IC	-	0.5739	0.2868	-	0.4939	0.1886	-	0.5134	0.1940
		(0.00)	(0.00)		(0.00)	(0.07)		(0.00)	(0.03)
R^2	37.47	31.17	41.98	37.60	23.18	39.79	37.83	21.78	39.94

Table 3: Individual and Index Variances, and Variance Risk Premiums

The table reports the time-series averages of realized ($\sqrt{RV}$) and model-free implied variances ($\sqrt{MFIV}$), expressed in volatility terms, and the difference between them ($MFIV - RV$), expressed as a difference in variances, for two samples of stocks—components of S&P500, and DJ30 indices, for the sample period from 1996 to 12/2012, and from 10/1997 to 12/2012, respectively, and for three different maturities—30, 60, and 91 (calendar) days. For individual stocks the variances are equally-weighted cross-sectional averages across all constituent stocks. Model-free implied variance ($MFIV$) is computed on each day using out-of-the money options with the respective maturity, and realized variance RV is calculated on each day from daily returns over a respective window, corresponding to the maturity of $MFIV$. All variances (volatilities) are expressed in annual terms. The p-value is for the null hypothesis that implied and realized variance are on average equal.

	Individual Stocks			Index		
Days	$\sqrt{RV}$	$\sqrt{MFIV}$	$MFIV - RV$	$\sqrt{RV}$	$\sqrt{MFIV}$	$MFIV - RV$
S&P500 Sample						
30	0.4199	0.4338	0.0117	0.2069	0.2309	0.0105
	-	-	(0.08)	-	-	(0.00)
60	0.4213	0.4243	0.0023	0.2066	0.2305	0.0105
	-	-	(0.82)	-	-	(0.02)
91	0.4227	0.4147	-0.0068	0.2066	0.2308	0.0106
	-	-	(0.60)	-	-	(0.05)
DJ30 Sample						
30	0.3600	0.3661	0.0044	0.1994	0.2278	0.0121
	-	-	(0.48)	-	-	(0.00)
60	0.3611	0.3596	-0.0011	0.1987	0.2286	0.0128
	-	-	(0.91)	-	-	(0.00)
91	0.3623	0.3508	-0.0083	0.1986	0.2286	0.0128
	-	-	(0.47)	-	-	(0.01)

Table 4: Individual Variance Risk Premiums

The table reports the results of individual tests of variance risk premiums, for two samples of stocks—components of S&P500, and DJ30 indices, for the sample period from 1996 to 12/2012, and from 10/1997 to 12/2012, respectively, and for three different maturities—30, 60, and 91 (calendar) days. The table shows the number of stocks, for which the respective hypothesis is either rejected ($MFIV - RV > 0$, and $MFIV - RV < 0$), or failed to be rejected ($MFIV = RV$). Model-free implied variance ($MFIV$) is computed on each day using out-of-the money options with the respective maturity, and realized variance RV is calculated on each day from daily returns over a respective window, corresponding to the maturity of $MFIV$. The test statistics are based on Newey-West (1987) autocorrelation consistent standard errors with lags equal to the number of overlapping observations (21, 41, or 62, respectively).

Days	Individual hypothesis	Number of Stocks
S&P500 Sample		
	$MFIV - RV > 0$ rejected	6
30	$MFIV - RV = 0$ not rejected	503
	$MFIV - RV < 0$ rejected	410
	$MFIV - RV > 0$ rejected	9
60	$MFIV - RV = 0$ not rejected	635
	$MFIV - RV < 0$ rejected	275
	$MFIV - RV > 0$ rejected	20
91	$MFIV - RV = 0$ not rejected	725
	$MFIV - RV < 0$ rejected	172
DJ30 Sample		
	$MFIV - RV > 0$ rejected	0
30	$MFIV - RV = 0$ not rejected	27
	$MFIV - RV < 0$ rejected	16
	$MFIV - RV > 0$ rejected	0
60	$MFIV - RV = 0$ not rejected	33
	$MFIV - RV < 0$ rejected	10
	$MFIV - RV > 0$ rejected	0
91	$MFIV - RV = 0$ not rejected	37
	$MFIV - RV < 0$ rejected	6

Table 5: Market Return Predictability: Option-Based Variables

The table shows significance of the coefficients (in terms of t-statistic) and the R^2 of the market return predictive regressions, for two samples of stocks—components of S&P500, and DJ30 indices, and for the sample period from 1996 to 12/2012, and from 10/1997 to 12/2012, respectively. We regress overlapping market returns (i.e., respective index returns) compounded over a specified horizon (one, three, six, and twelve months) on a constant and a given set of explanatory variables, which are implied correlation (IC) for 30 calendar days, the difference between implied correlation and lagged realized correlation also computed over 30 calendar days ($IC - RC$), model-free implied variance ($MFIV$) for 30 calendar days, and the variance risk premium equal to the difference between model-free implied variance and lagged realized variance also computed over 30 days ($MFIV - RV$). The t-statistics for the null hypothesis that the coefficients are equal zero are computed using Newey-West (1987) standard errors.

	Return Predictability Horizon			
Regressors	1 month	3 months	6 months	12 months
S&P500 Sample				
IC	3.1597	4.0676	4.6065	2.8201
R^2	2.53	9.58	13.97	11.00
$IC - RC$	2.0299	2.6706	2.2373	1.8686
R^2	1.22	3.49	3.25	2.75
$MFIV$	0.5530	0.3254	1.2630	2.4525
R^2	0.19	0.20	2.18	3.62
$MFIV - RV$	4.1059	7.2872	2.9988	0.5220
R^2	3.27	5.92	1.84	0.08
IC	3.0675	4.5760	4.5408	2.8046
$MFIV - RV$	3.9815	6.7654	3.9615	0.5883
R^2	5.58	14.90	15.42	11.02
DJ30 Sample				
IC	2.6506	3.5876	4.0268	2.9376
R^2	2.15	7.46	12.10	12.18
$IC - RC$	1.6495	2.5614	2.1214	1.6106
R^2	0.89	3.86	4.57	2.75
$MFIV$	1.0359	0.3935	1.1569	3.3233
R^2	0.66	0.38	2.84	8.40
$MFIV - RV$	3.2722	6.2312	3.1854	1.1164
R^2	2.31	7.42	3.98	0.76
IC	2.5725	3.9696	4.1118	2.8857
$MFIV - RV$	3.1741	6.1595	4.4368	1.5881
R^2	4.09	13.64	14.93	12.49

Table 6: Implied and Realized Correlations, Subperiods

The table reports the mean values for the implied correlation (IC), realized correlation (RC), and for the difference between them (IC-RC), for three sample periods (subsamples of the main sample period from 01/1996 to 12/2012 for S&P500 and from 10/1997 to 12/2012 for DJ30), for two samples of stocks—components of S&P500, and DJ30 indices, and for three different maturities—30, 60, and 91 (calendar) days. $IC(t)$ is calculated from daily observations of model-free implied variances for the index and for all index components, using (14). $RC(t)$ is a cross-sectional weighted average (using the appropriate weights from the respective index) of all realized pairwise correlations at time t, each calculated over a 30-, 60-, or 91-day (calendar) window of daily stock returns.

	1996-2001			2002-2007			2008-2012		
Days	RC	IC	$IC-RC$	RC	IC	$IC-RC$	RC	IC	$IC-RC$
S&P500 Sample									
30	0.2347	0.3599	0.1252	0.3299	0.3615	0.0316	0.4325	0.4773	0.0448
	-	-	(0.00)	-	-	(0.00)	-	-	(0.00)
60	0.2383	0.3804	0.1421	0.3240	0.3801	0.0561	0.4351	0.4993	0.0641
	-	-	(0.00)	-	-	(0.00)	-	-	(0.00)
91	0.2397	0.4012	0.1615	0.3222	0.3970	0.0748	0.4405	0.5170	0.0765
	-	-	(0.00)	-	-	(0.00)	-	-	(0.01)
DJ30 Sample									
30	0.2135	0.4341	0.2205	0.3398	0.4056	0.0659	0.4713	0.5552	0.0839
	-	-	(0.00)	-	-	(0.00)	-	-	(0.00)
60	0.2094	0.4777	0.2683	0.3375	0.4217	0.0843	0.4760	0.5741	0.0981
	-	-	(0.00)	-	-	(0.00)	-	-	(0.00)
91	0.2082	0.5028	0.2946	0.3372	0.4358	0.0986	0.4835	0.5848	0.1013
	-	-	(0.00)	-	-	(0.00)	-	-	(0.00)

Table 7: Individual and Index Variances, and Variance Risk Premiums, 1996–2001

The table reports the time-series averages of realized ($\sqrt{RV}$) and model-free implied variances ($\sqrt{MFIV}$), expressed in volatility terms, and the difference between them ($MFIV - RV$), expressed as a difference in variances, for two indices and respective samples of stocks—components of S&P500, and DJ30, for the sample period from 01/1996 to 12/2001 for S&P500, and from 10/1997 to 12/2001 for DJ30, and for three different maturities—30, 60, and 91 (calendar) days. For individual stocks the variances are equally-weighted cross-sectional averages across all constituent stocks. Model-free implied variance ($MFIV$) is computed on each day using out-of-the money options with the respective maturity, and realized variance RV is calculated on each day from daily returns over a respective window, corresponding to the maturity of $MFIV$. All variances (volatilities) are expressed in annual terms. The p-value is for the null hypothesis that implied and realized variance are on average equal.

	Individual Stocks			Index		
Days	$\sqrt{RV}$	$\sqrt{MFIV}$	$MFIV - RV$	$\sqrt{RV}$	$\sqrt{MFIV}$	$MFIV - RV$
S&P500 Sample						
30	0.4320	0.4481	0.0140	0.1904	0.2297	0.0165
	-	-	(0.02)	-	-	(0.00)
60	0.4329	0.4335	0.0003	0.1905	0.2293	0.0163
	-	-	(0.96)	-	-	(0.00)
91	0.4338	0.4200	-0.0118	0.1905	0.2296	0.0164
	-	-	(0.18)	-	-	(0.00)
DJ30 Sample						
30	0.3988	0.3842	-0.0114	0.1946	0.2480	0.0236
	-	-	(0.09)	-	-	(0.00)
60	0.3979	0.3753	-0.0175	0.1913	0.2540	0.0279
	-	-	(0.05)	-	-	(0.00)
91	0.3987	0.3665	-0.0246	0.1910	0.2531	0.0276
	-	-	(0.02)	-	-	(0.00)

Table 8: Individual Variance Risk Premiums, 1996–2001

The table reports the results of individual tests of variance risk premiums, for two samples of stocks—components of S&P500, and DJ30 indices, for the sample period from 01/1996 to 12/2001 for S&P500 and from 10/1997 to 12/2001 for DJ30, and for three different maturities—30, 60, and 91 (calendar) days. The table shows the number of stocks, for which the respective hypothesis is either rejected ($MFIV - RV > 0$, and $MFIV - RV < 0$), or failed to be rejected ($MFIV = RV$). Model-free implied variance ($MFIV$) is computed on each day using out-of-the money options with the respective maturity, and realized variance RV is calculated on each day from daily returns over a respective window, corresponding to the maturity of $MFIV$. The test statistics are based on Newey-West (1987) autocorrelation consistent standard errors with lags equal to the number of overlapping observations (21, 41, or 62, respectively).

Days	Individual hypothesis	Number of Stocks
S&P500 Sample		
	$MFIV - RV > 0$ rejected	15
30	$MFIV - RV = 0$ not rejected	389
	$MFIV - RV < 0$ rejected	239
	$MFIV - RV > 0$ rejected	23
60	$MFIV - RV = 0$ not rejected	460
	$MFIV - RV < 0$ rejected	160
	$MFIV - RV > 0$ rejected	48
91	$MFIV - RV = 0$ not rejected	490
	$MFIV - RV < 0$ rejected	105
DJ30 Sample		
	$MFIV - RV > 0$ rejected	1
30	$MFIV - RV = 0$ not rejected	32
	$MFIV - RV < 0$ rejected	1
	$MFIV - RV > 0$ rejected	6
60	$MFIV - RV = 0$ not rejected	28
	$MFIV - RV < 0$ rejected	0
	$MFIV - RV > 0$ rejected	11
91	$MFIV - RV = 0$ not rejected	23
	$MFIV - RV < 0$ rejected	0

Table 9: Individual and Index Variance, and Variance Risk Premiums, 2002–2007

The table reports the time-series averages of realized ($\sqrt{RV}$) and model-free implied variances ($\sqrt{MFIV}$), expressed in volatility terms, and the difference between them ($MFIV - RV$), expressed as a difference in variances, for the sample period from 01/2002 to 12/2007, for two indices and respective samples of stocks—components of S&P500, and DJ30, and for three different maturities—30, 60, and 91 (calendar) days. For individual stocks the variances are equally-weighted cross-sectional averages across all constituent stocks. Model-free implied variance ($MFIV$) is computed on each day using out-of-the money options with the respective maturity, and realized variance RV is calculated on each day from daily returns over a respective window, corresponding to the maturity of $MFIV$. All variances (volatilities) are expressed in annual terms. The p-value is for the null hypothesis that implied and realized variance are on average equal.

	Individual Stocks			Index		
Days	$\sqrt{RV}$	$\sqrt{MFIV}$	$MFIV - RV$	$\sqrt{RV}$	$\sqrt{MFIV}$	$MFIV - RV$
S&P500 Sample						
30	0.3407	0.3698	0.0205	0.1624	0.1898	0.0096
	-	-	(0.00)	-	-	(0.00)
60	0.3428	0.3639	0.0148	0.1625	0.1894	0.0095
	-	-	(0.04)	-	-	(0.00)
91	0.3440	0.3578	0.0096	0.1628	0.1893	0.0093
	-	-	(0.32)	-	-	(0.01)
DJ30 Sample						
30	0.2861	0.3130	0.0162	0.1657	0.1962	0.0110
	-	-	(0.00)	-	-	(0.00)
60	0.2857	0.3094	0.0141	0.1662	0.1954	0.0106
	-	-	(0.00)	-	-	(0.00)
91	0.2850	0.3053	0.0120	0.1661	0.1951	0.0105
	-	-	(0.03)	-	-	(0.00)

Table 10: Individual Variance Risk Premiums 2002–2007

The table reports the results of individual tests of variance risk premiums, for the sample period from 01/2002 to 12/2007, for two samples of stocks—components of S&P500, and DJ30 indices, and for three different maturities—30, 60, and 91 (calendar) days. The table shows the number of stocks, for which the respective hypothesis is either rejected ($MFIV - RV > 0$, and $MFIV - RV < 0$), or failed to be rejected ($MFIV = RV$). Model-free implied variance ($MFIV$) is computed on each day using out-of-the money options with the respective maturity, and realized variance RV is calculated on each day from daily returns over a respective window, corresponding to the maturity of $MFIV$. The test statistics are based on Newey-West (1987) autocorrelation consistent standard errors with lags equal to the number of overlapping observations (21, 41, or 62, respectively).

Days	Individual hypothesis	Number of Stocks
S&P500 Sample		
	$MFIV - RV > 0$ rejected	7
30	$MFIV - RV = 0$ not rejected	238
	$MFIV - RV < 0$ rejected	378
	$MFIV - RV > 0$ rejected	14
60	$MFIV - RV = 0$ not rejected	331
	$MFIV - RV < 0$ rejected	278
	$MFIV - RV > 0$ rejected	25
91	$MFIV - RV = 0$ not rejected	419
	$MFIV - RV < 0$ rejected	179
DJ30 Sample		
	$MFIV - RV > 0$ rejected	0
30	$MFIV - RV = 0$ not rejected	9
	$MFIV - RV < 0$ rejected	24
	$MFIV - RV > 0$ rejected	0
60	$MFIV - RV = 0$ not rejected	14
	$MFIV - RV < 0$ rejected	19
	$MFIV - RV > 0$ rejected	0
91	$MFIV - RV = 0$ not rejected	18
	$MFIV - RV < 0$ rejected	15

Table 11: Individual and Index Variance, and Variance Risk Premiums, 2008–2012

The table reports the time-series averages of realized ($\sqrt{RV}$) and model-free implied variances ($\sqrt{MFIV}$), expressed in volatility terms, and the difference between them ($MFIV - RV$), expressed as a difference in variances, for the sample period from 01/2008 to 12/2012, for two indices and respective samples of stocks—components of S&P500, and DJ30, and for three different maturities—30, 60, and 91 (calendar) days. For individual stocks the variances are equally-weighted cross-sectional averages across all constituent stocks. Model-free implied variance ($MFIV$) is computed on each day using out-of-the money options with the respective maturity, and realized variance RV is calculated on each day from daily returns over a respective window, corresponding to the maturity of $MFIV$. All variances (volatilities) are expressed in annual terms. The p-value is for the null hypothesis that implied and realized variance are on average equal.

	Individual Stocks			Index		
Days	$\sqrt{RV}$	$\sqrt{MFIV}$	$MFIV - RV$	$\sqrt{RV}$	$\sqrt{MFIV}$	$MFIV - RV$
S&P500 Sample						
30	0.4868	0.4852	-0.0018	0.2663	0.2742	0.0043
	-	-	(0.93)	-	-	(0.66)
60	0.4894	0.4787	-0.0106	0.2659	0.2741	0.0044
	-	-	(0.75)	-	-	(0.76)
91	0.4927	0.4710	-0.0211	0.2666	0.2756	0.0049
	-	-	(0.62)	-	-	(0.78)
DJ30 Sample						
30	0.4218	0.4198	-0.0018	0.2419	0.2540	0.0060
	-	-	(0.92)	-	-	(0.45)
60	0.4265	0.4125	-0.0119	0.2419	0.2537	0.0059
	-	-	(0.65)	-	-	(0.63)
91	0.4309	0.3988	-0.0267	0.2425	0.2550	0.0062
	-	-	(0.42)	-	-	(0.67)

Table 12: Individual Variance Risk Premiums 2008–2012

The table reports the results of individual tests of variance risk premiums, for the sample period from 01/2008 to 12/2012, for two samples of stocks—components of S&P500, and DJ30 indices, and for three different maturities—30, 60, and 91 (calendar) days. The table shows the number of stocks, for which the respective hypothesis is either rejected ($MFIV - RV > 0$, and $MFIV - RV < 0$), or failed to be rejected ($MFIV = RV$). Model-free implied variance ($MFIV$) is computed on each day using out-of-the money options with the respective maturity, and realized variance RV is calculated on each day from daily returns over a respective window, corresponding to the maturity of $MFIV$. The test statistics are based on Newey-West (1987) autocorrelation consistent standard errors with lags equal to the number of overlapping observations (21, 41, or 62, respectively).

Days	Individual hypothesis	Number of Stocks
S&P500 Sample		
	$MFIV - RV > 0$ rejected	5
30	$MFIV - RV = 0$ not rejected	369
	$MFIV - RV < 0$ rejected	221
	$MFIV - RV > 0$ rejected	6
60	$MFIV - RV = 0$ not rejected	448
	$MFIV - RV < 0$ rejected	141
	$MFIV - RV > 0$ rejected	15
91	$MFIV - RV = 0$ not rejected	481
	$MFIV - RV < 0$ rejected	97
DJ30 Sample		
	$MFIV - RV > 0$ rejected	1
30	$MFIV - RV = 0$ not rejected	22
	$MFIV - RV < 0$ rejected	13
	$MFIV - RV > 0$ rejected	0
60	$MFIV - RV = 0$ not rejected	29
	$MFIV - RV < 0$ rejected	7
	$MFIV - RV > 0$ rejected	0
91	$MFIV - RV = 0$ not rejected	30
	$MFIV - RV < 0$ rejected	6

Table 13: Market Return Predictability: Adding Fundamentals

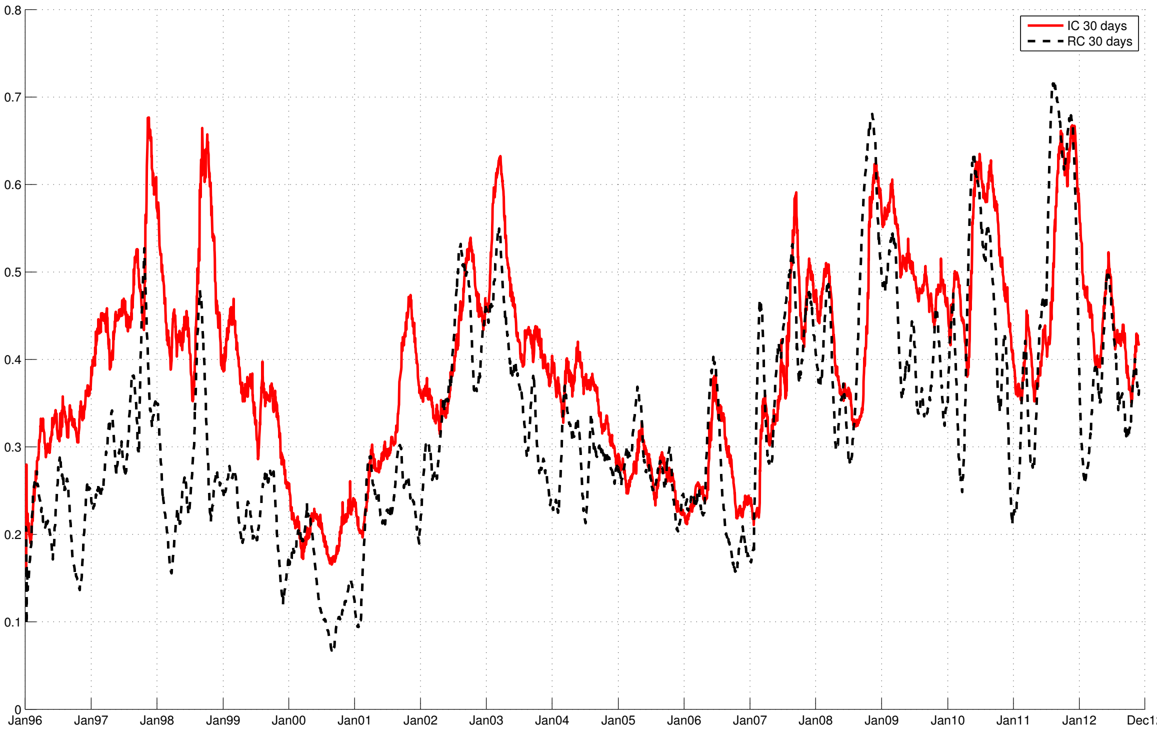
The table shows the results of predictive regressions, where we regress market return (S&P500 index return) on a number of predictors, always including in the list of predictors two option-based variables–the implied correlation (IC) for 30 calendar days,), and the variance risk premium equal to the difference between model-free implied variance and lagged realized variance also computed over 30 days ($MFIV - RV$). In addition to the option-based variables we include a number of fundamentals commonly shown to predict future returns. Due to limits in availability of data on fundamentals we perform the regressions sampling the returns and predictors at monthly frequency, and we predict market returns for four different horizons (one month, one and two quarters, and one year). The description of fundamental variables is given in Section 5.2. The p-values for the coefficients are given in parentheses, and they are based on Newey and West (1987) standard errors with a number of lags equal to the horizon minus one. The R^2's are the adjusted ones.

Horizon	Regression	IC	$MFIV - RV$	EP	TMS	DFY	BTM	$NTIS$	R^2
	(1)	0.0510	0.3982	-	-	-	-	-	11.84
		(0.04)	(0.00)						
	(2)	0.0544	0.4039	0.0088	-	-	-	-	11.95
1 month		(0.03)	(0.00)	(0.27)					
	(3)	0.0668	0.3863	0.0052	-0.2505	-0.2623	-	-	11.54
		(0.02)	(0.00)	(0.60)	(0.35)	(0.80)			
	(4)	0.0670	0.3674	0.0059	-0.2821	-	-0.0084	0.1478	11.39
		(0.02)	(0.00)	(0.56)	(0.34)		(0.89)	(0.45)	
	(1)	0.2373	0.8115	-	-	-	-	-	26.57
		(0.00)	(0.00)						
	(2)	0.2422	0.8196	0.0125	-	-	-	-	26.53
3 months		(0.00)	(0.00)	(0.54)					
	(3)	0.2938	0.7461	-0.0023	-1.0440	-1.0992	-	-	28.52
		(0.00)	(0.00)	(0.91)	(0.07)	(0.68)			
	(4)	0.2659	0.5967	-0.0272	-1.8407	-	0.2358	1.2125	33.98
		(0.00)	(0.00)	(0.32)	(0.00)		(0.07)	(0.03)	
	(1)	0.3946	0.7833	-	-	-	-	-	22.52
		(0.00)	(0.00)						
	(2)	0.3986	0.7900	0.0103	-	-	-	-	22.21
6 months		(0.00)	(0.00)	(0.81)					
	(3)	0.4780	0.7022	-0.0091	-1.8373	-0.8454	-	-	24.90
		(0.00)	(0.00)	(0.82)	(0.09)	(0.86)			
	(4)	0.4030	0.2224	-0.1008	-4.1635	-	0.7794	3.3834	47.88
		(0.00)	(0.40)	(0.04)	(0.00)		(0.00)	(0.00)	
	(1)	0.5155	0.5408	-	-	-	-	-	13.91
		(0.00)	(0.00)						
	(2)	0.5289	0.5621	0.0322	-	-	-	-	13.94
12 months		(0.00)	(0.00)	(0.66)					
	(3)	0.5215	0.6184	0.0406	-0.2550	1.6609	-	-	13.10
		(0.00)	(0.02)	(0.61)	(0.90)	(0.85)			
	(4)	0.4212	-0.3779	-0.1360	-4.0532	-	1.3063	6.0899	46.31
		(0.00)	(0.35)	(0.09)	(0.02)		(0.00)	(0.00)	

Figure 1: Implied versus Realized Correlations

The figure shows 1-week moving averages of the implied correlation and the realized correlation for S&P500 and DJ30 for the period from 1996 to 12/2012, and from 10/1997 to 12/2012, respectively. The implied correlation is calculated from daily observations on model-free implied variances for the index and for all index components, using (18). Each model-free implied variance is calculated from 30-day options. The realized correlation at time t is a cross-sectional weighted average (using the appropriate weights from the index) of all realized pairwise correlations at time t, each calculated over a 30-day window of daily stock returns.

(a) S&P500 Sample



(b) DJ30 Sample

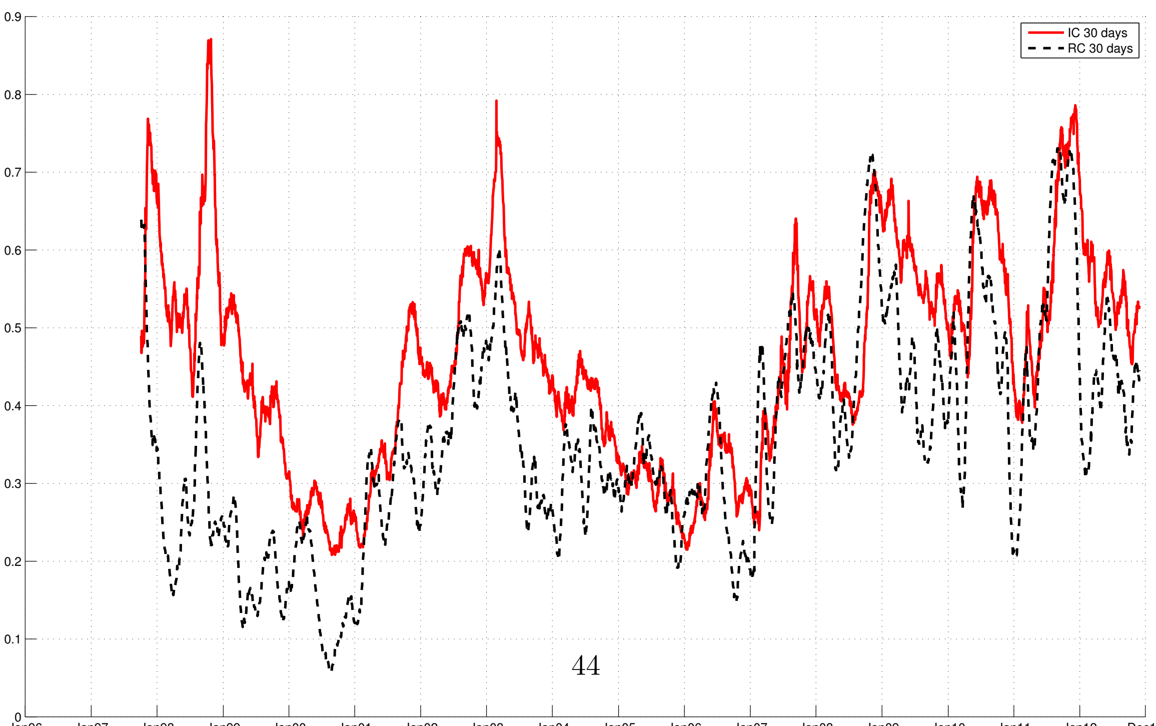
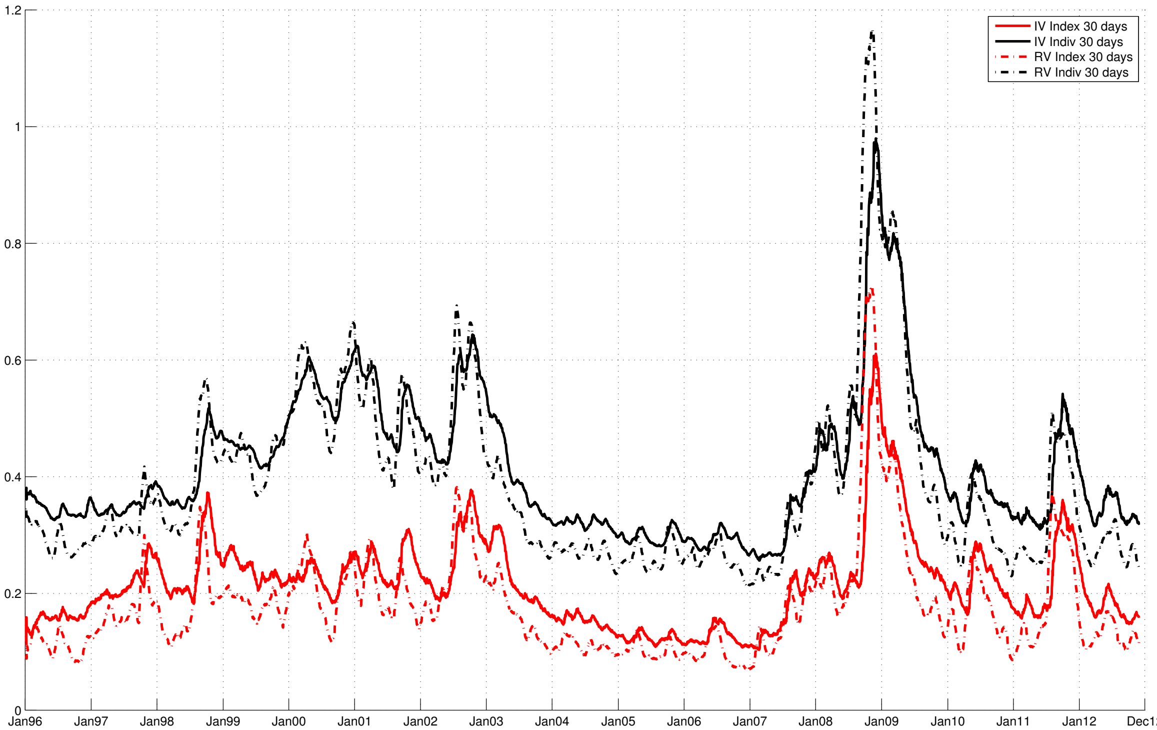


Figure 2: Implied versus Realized Variance

The figure presents the time-series of the square root of the index model-free implied and realized variances, as well as of the square root of the cross-sectional average of model-free implied and realized variances of individual stocks for our S&P500 and DJ30 samples for the period from 1996 to 12/2012, and from 10/1997 to 12/2012, respectively. Variances are computed over the period of 30 (calendar) days and are expressed in annual terms.

(a) S&P500 Sample



(b) DJ30 Sample

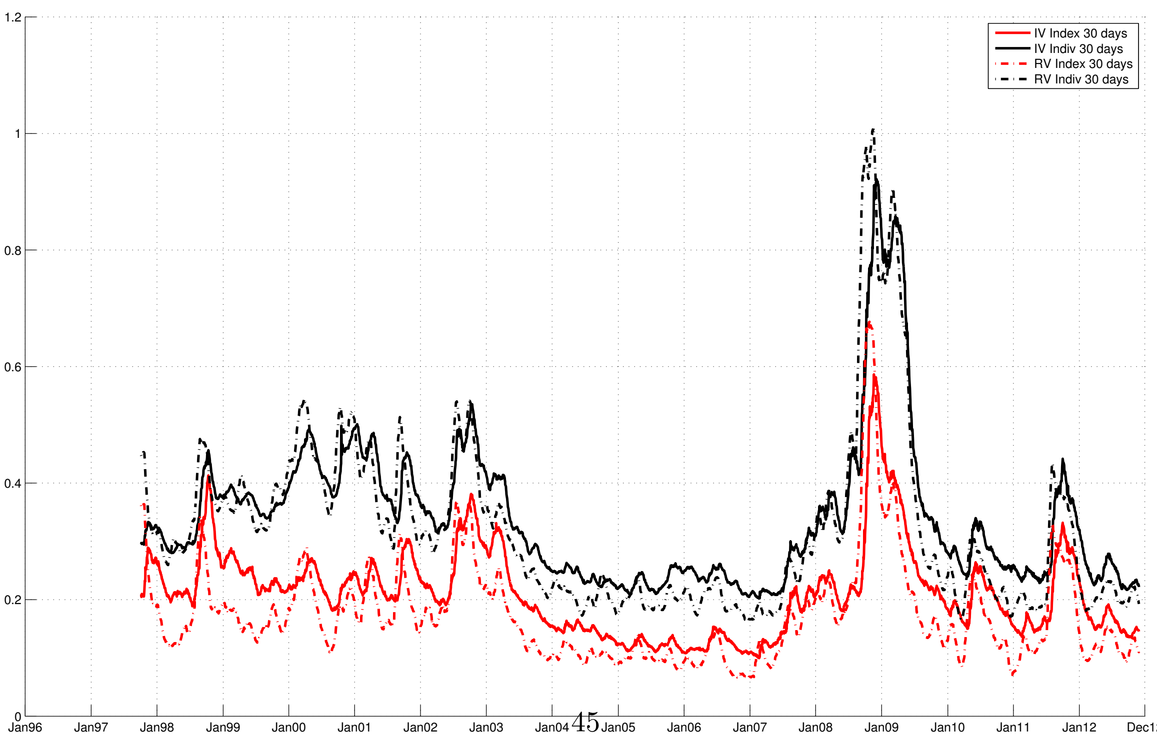
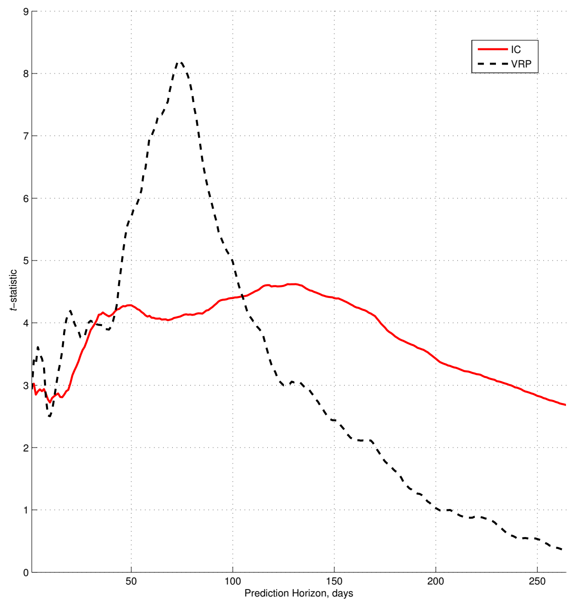


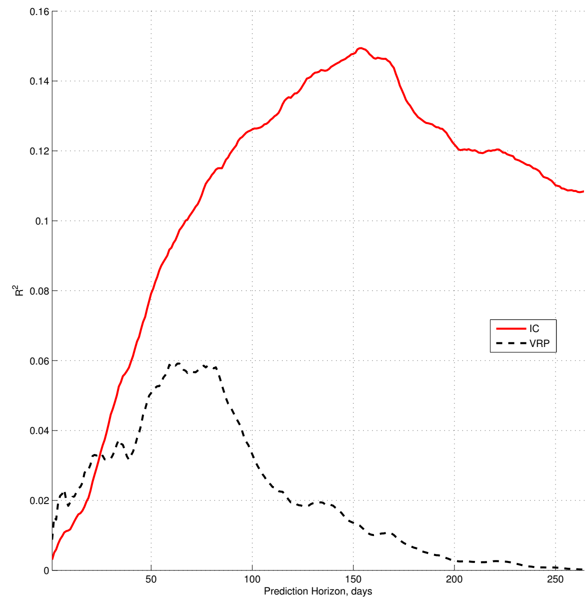
Figure 3: Implied versus Realized Variance

The figure presents the t-statistics and R^2 for the predictive regressions of aggregate market return, compounded over a horizon from one day to one year, on a constant and either implied correlation (IC) computed for 30-day maturity options, or the (market) variance risk premium (VRP), which is equal to the difference between 30-day model-free implied variance and lagged realized variance, also computed over 30 calendar days. We run the predictive regressions for S&P500 and DJ30 returns over the period from 1996 to 12/2012, and from 10/1997 to 12/2012, respectively.

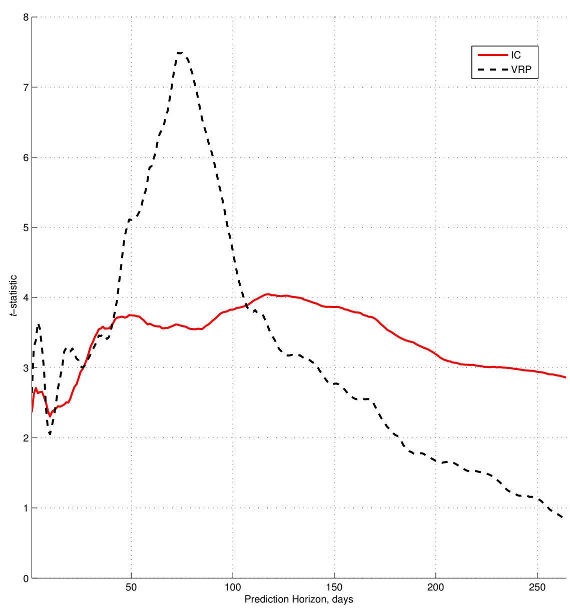
(a) t-statistics: S&P500 Sample



(b) R^2: S&P500 Sample



(c) t-statistics: DJ30 Sample



(d) R^2: DJ30 Sample