UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

IMPLIED CORRELATION AND OPTION RETURNS

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN ECONOMÍA APLICADA,
MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

DANIEL EDUARDO SZMULEWICZ FIERRO

PROFESOR GUÍA:
MARCELA VALENZUELA BRAVO

MIEMBROS DE LA COMISIÓN:
ALEJANDRO BERNALES SILVA
PATRICIO VALENZUELA AROS

Este trabajo ha sido parcialmente financiado por Fondecyt 1190477

SANTIAGO DE CHILE
2020

RESUMEN DE LA MEMORIA PARA OPTAR AL TÍTULO DE
MASTER OF APPLIED ECONOMICS
POR: DANIEL EDUARDO SZMULEWICZ FIERRO
FECHA: 2020
PROF. GUÍA: MARCELA VALENZUELA BRAVO

IMPLIED CORRELATION AND OPTION RETURNS

La presente tesis estudia si la correlación es un factor de riesgo sistemático, analizando su efecto en el mercado de opciones. Para este propósito, calculamos la *correlación implícita*, como una medida de la correlación en todo el mercado, y probamos si hay un efecto significativo en el time-series y en el cross-section de los retornos de la opción.

Mostramos que los períodos de alta correlación implícita son seguidos por una disminución en el time-series de los retornos de opciones tipo Put. Al utilizar los retornos de la opción sobre el índice S&P100 como variable dependiente, mostramos que nuestro índice de correlación implícita tiene un gran poder predictivo tanto in-sample como out-of-sample del time-series.

Estudiamos el cross-section de los retornos de las opciones sobre acciones, clasificando estas últimas en función de la sensibilidad a las innovaciones en la correlación implícita. Encontramos que las acciones con más exposición tienen un rendimiento promedio alto de opciones tipo Put. Una estrategia que consiste en una posición larga (corta) en la cartera con opciones escritas en acciones más (menos) expuestas, produce un rendimiento mensual promedio económico y estadísticamente significativo.

Creamos una estrategia innovadora utilizando un pronóstico mensual de la correlación implícita y, por lo tanto, predecimos el comportamiento de los inversores en el mercado de opciones. Esta estrategia produce un rendimiento mensual significativo desde el punto de vista estadístico y económico, incluso incluyendo todo el bid-ask spread como costo de transacción.

Los resultados son sólidos para diferentes condiciones de mercado y diferentes períodos de muestra, y no se explican por los modelos habituales de factores de riesgo.

ABSTRACT OF THE REPORT TO QUALIFY TO THE DEGREE OF MASTER OF SCIENCE, MENTION ECONOMICS
BY: DANIEL EDUARDO SZMULEWICZ FIERRO
DATE: 2020
GUIDE: MARCELA VALENZUELA BRAVO

IMPLIED CORRELATION AND OPTION RETURNS

The current thesis studies whether the correlation is a systematic risk factor, analyzing its effect on the options market. For this purpose, we calculate the *implied correlation* of the market, as a measure of the marketwide correlation, and we test if there is a significant effect on the time-series and the cross-section of the option returns.

We show that periods of high implied correlation are followed by a decrease in the time-series of Put option returns. Using the option returns on the S&P100 index as a dependent variable, we show that our implied correlation index has an impressive predictive power both in-sample and out-of-sample in the time-series of returns.

We study the cross-section of stock option returns by sorting stocks on the sensitivities to innovations in implied correlation. We find that stocks with more exposure have high average Put options returns. A trading strategy that is long (short) in the portfolio with options written on stocks more (less) exposed, produces an economically and statistically significant average monthly return.

We create an innovative strategy using a monthly forecast of the implicit correlation, and thus predict the behavior of investors in the options market. This strategy produces a statistically and economically significant monthly return, even including the entire bid-ask spread as transaction cost.

The results are robust to different market conditions and different sub-sample periods, and are not explained by usual risk factor models.

To my parents, Claudio and Silvia.
Thank you for your endless love and support.

Contents

Introduction

Despite the growth in equity options research, the question of how marketwide correlation affects the cross-section of expected options returns has received less attention. Most research has focused on volatility as the main factor to explain the difference in options returns, while very few on the marketwide correlation.

Correlation is one of the key concepts in financial markets, and it has been proved that changes over time. When stock return correlations are high, the returns are low and the market volatility increase, so the benefits of the diversification decline precisely when needed. The investors would like to protect from this risk and keep a portfolio that could behave well in periods of high correlation. One asset that hedge against downside risk are Put options. Investors demand put options contracts to keep in their portfolios, and hence be prepared for periods with high marketwide correlation. The greater the exposure, the greater the demand for options contract on this stock. Thus, contracts with underlyings more sensitive to correlation risk are expensive and earn low returns, because they offer a hedge against a marketwide correlation increase, in contrast with contracts whose underlying is less sensitive to correlation risk.

We compute an Implied Correlation metric as proxy for the marketwide correlation, using index and individual options prices. Using this metric gives us some advantages: first, it has been shown that the predictive power of implied measures are better than historical ones, the former providing incremental information on future values not included in historical estimates. Second, using the prices of the options we obtain the expectations of the investors, their sentiment about the market and the months ahead; since we will build strategies based on investors' decisions that are more sentiment-driven, this is a convenient measure. Moreover, as Buss and Vilkov [2012] point out, option prices update faster in response to new market conditions, since historical data have some inertia incorporated in them. Finally, when estimating risk-neutral expectations of higher moments using options, we do not face the trade-off between using long time-series of data to obtain precise estimates and short windows to produce conditional instead of unconditional estimates. Hence, we expect that the implied correlation is a better predictor of future market returns compared to other indicators relying on historical data.

Our empirical methodology splits in two: first we study if our metric of implied correlation has a predictive power in the time-series of the option market returns, using S&P100 index option contracts as dependent variable, and checking the in-sample and out-of-sample results; once the predictive power is established, we look for a possible pricing correlation risk premium in the cross-section of the option market, using the individual option contracts on the S&P100 constituents.

We show that aggregate implied correlation is significantly associated with a decrease in subsequent option returns. The predictive power is stronger for closer periods, and robust to the inclusion of predictors such as aggregate volatility factors, market-wide liquidity, variance and correlation risk premiums. The economic importance of implied correlation is high: a one standard deviation increase in implied correlation in a given month translates into a 22% decrease in the subsequent quarterly cumulative option return. We also verify that the implied correlation has good out-of-sample performance. Specifically, using an out-of-sample R^2 and a F-statistic with critical values from a bootstrap procedure, we show that predictive regressions relying on implied correlation deliver better out-of-sample forecasts than those employing only the historical average of option returns.

Further, using the cross-section of stocks returns, we create portfolios of put options based on the underlying sensitivity to innovations in implied correlation. Options written on underlyings that behave poorly in periods of high correlation earn higher returns than options contracts on shares that behave well. We show a significative difference between pricing of extreme portfolios, that could be attributed to correlation risk. Further, using the autoregressive feature of the implied correlation, we implement a trading strategy that consists in taking long/short positions in the first/fifth portfolio, depending on a forecast of the next month IC. If the prediction supposes an increase in implied correlation, we expect a growth in the demand of options written on poorly behaved stocks, so we take a long position in those options, taking advantage of the low price, and a short position in Put option contracts on well-behaved stocks, waiting for a price rise due to lower demand. This strategy provides a 10% monthly return.

In order to check if the returns of the strategies could be associated with any known risk factor, we apply a series of control variables to the returns in a factor-model regression. We test various specifications for the post-formation control regressions, and concluded that, for both strategies, the return remains significant and cannot be attributed to these risk premiums. As these are two investment strategies, we analyze the inclusion of transaction costs in the purchase and sale of the option contracts, and find that the returns remains at very reasonable estimates of effective spreads; in particular, our second strategy provides significant returns even at considering the entire bid-ask effective spread. We also check the robustness of the sample period, dividing our data in five, and examine the results for each one. Our findings remain the same for every expansion sub-sample, and the returns of our second strategy become negative in periods of contractions.

This thesis is part of a vast literature on the explanation of option returns. Previous studies have shown that Put options have been historically too expensive. Bondarenko [2014] exhibits that Put prices are too high to be compatible with canonical models. In the same way, Coval and Shumway [2001] found that both Call and Put contracts earn exceedingly low returns. To analyze this problem most work focuses on Black and Scholes [1973] and volatility related option mispricing, such as Coval and Shumway [2001] using trading in index options, and Goyal and Saretto [2009] using individual options. One of the first papers to analyze the cross-sectional differences in option expected returns due to marketwide correlation risk was

Driessen et al. [2009] through a trading strategy that involves selling index straddles and buying individual straddles and stocks. Our approach is much simpler, only using individual Put options, we analyze the most sensitive underlying in the stock market, and thus create a difference in the options cross-section. This approach is very similar to that used by Ang et al. [2006] who examine the pricing of aggregate volatility risk in the cross-section of stock returns.

Finally, it is interesting to note that some research have focus on trading strategies involving short or long positions in index and in individual options simultaneously, to exploit some risk premium. Our first strategy involves only individual option contracts, similarly to Goyal and Saretto [2009], using the cross-section in the options market. And the second strategy requires a prediction of our measure of marketwide correlation one month ahead, and depending on whether the correlation will rise or fall, take a certain position in option contracts. To the best of our knowledge, we are the first to implement such a strategy in options.

The rest of the paper proceeds as follows. Section 1 describe the data and the main estimations. Section 2 presents the results for the time series estimation and prediction power of the Implied Correlation in the option market. Section 3 presents the results of the portfolios strategies and regressions, taking into account transaction costs and a check of robustness of the sample. The final Section concludes.

Chapter 1

Data

1.1 Implied Correlation

The instantaneous variance of the index at a given time t, σ^2_{It}, is a function of the instantaneous variances of individual constituents, $\sigma^2_{\mathrm{i}t}$, and the correlations between pairs of stock returns, $\rho_{\mathrm{i}jt}$,

$$\sigma^2_{It} = \sum_{\mathrm{i}=1}^{N} w_\mathrm{i}^2 \sigma^2_{\mathrm{i}t} + \sum_{\mathrm{i}=1}^{N} \sum_{\mathrm{i} \neq j} w_\mathrm{i} w_j \sigma_{\mathrm{i}t} \sigma_{jt} \rho_{\mathrm{i}jt} \tag{1.1}$$

where $w_{\mathrm{i}t}$ denotes the market weight of the ith component. From this equation, we can obtain an expression for the expected integrated variance under the risk-neutral probability measure $\mathcal{Q}$ over an interval of length $T - t$,

$$E_t^{\mathcal{Q}} \left[\int_t^T \sigma^2_{I\tau} \mathrm{d}\tau \right] = \sum_{\mathrm{i}=1}^{N} w_\mathrm{i}^2 E_t^{\mathcal{Q}} \left[\int_t^T \sigma^2_{\mathrm{i}\tau} \mathrm{d}\tau \right] + \sum_{\mathrm{i}=1}^{N} \sum_{\mathrm{i} \neq j} w_\mathrm{i} w_j E_t^{\mathcal{Q}} \left[\int_t^T \sigma_{\mathrm{i}\tau} \sigma_{j\tau} \rho_{\mathrm{i}j\tau} \mathrm{d}\tau \right] \tag{1.2}$$

By assuming equal pairwise implied correlations between all the pair stock returns, $\rho_{\mathrm{i}j\tau} = \rho_\tau$, and given that it is not possible to estimate the second term of the previous equation, we can use the following approximation,

$$\begin{aligned} &\sum_{\mathrm{i}=1}^{N} \sum_{\mathrm{i} \neq j} w_\mathrm{i} w_j E_t^{\mathcal{Q}} \left[\int_t^T \sigma_{\mathrm{i}\tau} \sigma_{j\tau} \rho_{\mathrm{i}j\tau} \mathrm{d}\tau \right] \\ &\approx \sum_{\mathrm{i}=1}^{N} \sum_{\mathrm{i} \neq j} w_\mathrm{i} w_j \sqrt{E_t^{\mathcal{Q}} \left[\int_t^T \sigma^2_{\mathrm{i}\tau} \mathrm{d}\tau \right]} \sqrt{E_t^{\mathcal{Q}} \left[\int_t^T \sigma^2_{j\tau} \mathrm{d}\tau \right]} E_t^{\mathcal{Q}} \left[\int_t^T \rho_\tau \mathrm{d}\tau \right] \end{aligned} \tag{1.3}$$

Then, it is straight forward to derive the expression for aggregate implied correlation $\mathrm{IC}_t = E_t^{\mathbb{Q}}\left[\int_t^T \rho_\tau \mathrm{d}\tau\right]$ by rearranging the equations above,

$$\mathrm{IC}_t \quad = \quad \frac{E_t^{\mathbb{Q}}\left[\int_t^T \sigma_{I\tau}^2 \mathrm{d}\tau\right] - \sum_{\mathrm{i}=1}^{N} w_{\mathrm{i}}^2 E_t^{\mathbb{Q}}\left[\int_t^T \sigma_{\mathrm{i}\tau}^2 \mathrm{d}\tau\right]}{\sum_{\mathrm{i}=1}^{N}\sum_{\mathrm{i}\neq j} w_{\mathrm{i}} w_j \sqrt{E_t^{\mathbb{Q}}\left[\int_t^T \sigma_{\mathrm{i}\tau}^2 \mathrm{d}\tau\right]}\sqrt{E_t^{\mathbb{Q}}\left[\int_t^T \sigma_{j\tau}^2 \mathrm{d}\tau\right]}} \tag{1.4}$$

IC_t represents the market's expectation of future market-wide correlation, implied by the option prices of the index and the prices of options on its components. It summarizes the pairwise correlations among all the individual components. An increase in IC_t is associated with a deterioration of the market's expectations of the portfolio diversification benefits.

1.2 Estimation of IC

To calculate the implied variance of the index and the implied variances of the index components, we use the risk-neutral variance of simple returns, which can be estimated from the strike of a simple variance swap. Specifically, we employ the strike on a simple variance swap on the market and the strikes on simple variance swaps on the index constituents. Martin [2011] introduces this financial contract with different properties to those of a standard variance swap. For instance, simple variance swaps can be hedged in the presence of jumps, and they measure the risk-neutral variance of simple returns. According to Martin [2011], it also provides a natural way to calculate implied correlations, since the decomposition of the index variance given in equation 1.1 refers to simple returns, not log returns.

The strike of a simple variance swap is defined as

$$V(0,T) \equiv \frac{2\mathrm{e}xp^{rT}}{F_T^2}\left(\int_0^{F_T} \mathrm{put}_T(K)\mathrm{d}K + \int_{F_T}^{\infty} \mathrm{call}_T(K)\mathrm{d}K\right), \tag{1.5}$$

where F_T denotes the underlying asset's forward price to time T at time 0, $\mathrm{put}_T(K)$ and $\mathrm{call}_T(K)$ are the put and call option prices with maturity date T and strike price K, respectively, and r is the continuously compounded interest rate. The integral is defined over an infinite set of strike prices. By assuming that the available strike prices of the put options belong to the interval $[K_{min}^P, K_{max}^P]$ where $0 < K_{min}^P < K_{max}^P < +\infty$, we solve the integral numerically using the trapezoidal method. Thus, the first term on the right-hand side of equation 1.5 is approximated as follows:

$$\frac{2}{F_T^2}\left(\int_{K_{min}^P}^{K_{max}^P} \mathrm{put}_T(K)\mathrm{d}K\right) \approx \frac{K_{max}^P - K_{min}^P}{m}\sum_{k=1}^{m}\left(\frac{\mathrm{put}_T(K_{\mathrm{i}})}{F_T^2} + \frac{\mathrm{put}_T(K_{\mathrm{i}-1})}{F_T^2}\right) \tag{1.6}$$

In a similar manner, we numerically approximate the second term of the right-hand side of equation 1.5 to finally obtain the estimates of the implied variance for the S&P100 index and its individual stocks. Similarly to Martin [2011], we approximate the forward price to the spot price. The implied variance is estimated for different maturities and, by interpolating, we construct daily estimates with 30 days time- to-maturity. Monthly time-series are given by the estimates at the end of each month. Once we have approximated the implied variances for the index and its constituents, we can finally obtain the aggregate implied correlation from equation 1.4.

1.3 Option Data and Variables

We use data from both the equity option and stock markets from the Ivy DB database provided by OptionMetrics. We use daily data for S&P100 index options and for individual options on all the stocks included in the S&P100 index from January 1996 to December 2014. For the option data we select the best bid and ask closing quotes, open interest, trading volume, implied volatility and the option's delta. For the underlyings we select the best bid and ask closing quotes, shares outstanding, closing price, and trading volume. We use the shares outstanding and the closing price to calculate the firm's market capitalization and thus obtain the value-weighted results of our portfolios in section 3.

We apply filters to the option data. We eliminate all the observations that violate arbitrage bounds, such as $K \geq P \geq \max(0, K\mathrm{e}^{-rT} - S)$ where P is the put option price, S is the underlying stock price, K is the strike price, T is time to maturity, and r is the risk-free rate. We remove observations with non positive bid price and for which the option open interest is equal to zero, in order to eliminate options with no liquidity.

For the time series analysis and later post-formation regressions for the cross-section, we utilize frequently used variables with predictive power and risk measures. The monthly CBOE Volatility Index (VIX) from the CBOE database. We employ Amihud [2002] illiquidity measure,

$$\mathrm{LIQ}_t = \frac{1}{T}\sum_{t}^{T} \frac{r_t}{\$V_t}, \tag{1.7}$$

where r_t and V_t are the absolute stock market return and dollar volume on day t respectively, and the time horizon $T - t$ is 30 days, as a control variable in the in-sample regression in section 2.1 and in the sorting process in section 3.1. We compute CRP_t as the difference between IC_t, obtained from 1.4, and RC, the average realized correlation, calculated as the sum of the value-weighted pairwise correlations,

$$\mathrm{RC}_t = \sum_{\mathrm{i}=1}^{N} \sum_{\mathrm{i} \neq j} w_{\mathrm{i}t} w_{jt} \hat{\rho}_{\mathrm{i}jt} \quad , \tag{1.8}$$

where $\hat{\rho}_{\mathrm{i}jt}$ is the sample correlation for each pair of constituents i and j of the S&P100 index, each month, obtained as

$$\hat{\rho}_{\mathrm{i}jt} = \frac{\hat{\sigma}_{\mathrm{i}jt}}{\hat{\sigma}_{\mathrm{i}t}\hat{\sigma}_{jt}} \quad , \tag{1.9}$$

where $\hat{\sigma}_{\mathrm{i}t}$ is the realized volatility of stock i and $\hat{\sigma}_{\mathrm{i}jt}$ is the covariance between stocks i and j. We calculate the realized volatility for each constituent i and each month as the square root of the realized variance of constituent i, where the latter is computed as

$$\mathrm{RV}_{\mathrm{i}t} = \sum_{t=1}^{T} \left((1 + r_{\mathrm{i}t}) - \frac{1}{T} \sum_{t=1}^{T} (1 + r_{\mathrm{i}t}) \right)^2 \quad , \tag{1.10}$$

$r_{\mathrm{i}t}$ is the return on trading day d. In the same way as CRP_t, the VRP_t is computed as the difference between the risk-neutral expectation and the physical expectation of the market variance,

$$\mathrm{VRP}_t = E_t^{\mathcal{Q}} \left[\int_t^T \sigma_\tau^2 \mathrm{d}\tau \right] - E_t^{\mathcal{P}} \left[\int_t^T \sigma_\tau^2 \mathrm{d}\tau \right] \quad , \tag{1.11}$$

where again the time horizon $T - t$ is 30 days. The implied variance (IV) is estimated by numerically solving the risk-neutral expectation of the simple return variance defined in equation 1.5 using the trapezoidal rule described in equation 1.6. The physical expectation of the market variance is approximated as the realized variation of the index from t to T described in equation 1.10. We include $\mathrm{vVRP}_{\mathrm{i}}$, calculating it as the value-weighted average of the variance premia on all the index constituents.

Further, we obtain the daily and monthly Fama and French [1993] factor returns: Market (MKT-Rf), Size (SMB), Value (HML), and the Carhart [1997] momentum factor (Mom) from Kenneth French's data library and use the T-bill rate as the risk-free rate.

1.4 Option Returns

We follow Broadie et al. [2009] to get a continuous series of hold-to-maturity options returns. We construct portfolios of options and their underlying, based on the information available on the first trading day following the third friday of the month. We select options that mature the next month and the contracts which are closest to ATM (moneyness between 0.95 and 1.05). We take the mean of the best bid and ask as the initial price, then we hold the portfolio one month, and at the expiration date (usually the third friday of the month) we use the payoff of exercise (or not) as the final value.

Thus, the hold-to-maturity return of put options are

$$r_{t,T}^{P} = \frac{(K - S_{t+T})^{+}}{P_{t,T}(K, S_t)} - 1 \tag{1.12}$$

where $x^{+} \equiv \max(x, 0)$, and $P_{t,T}(K, S_t)$ is the price at t, of a put option on the underlying S_t, with strike price K, and maturity T.

After the expiration, for the next month, we select a new pair of call and put contracts, the closest ones to ATM, that have one month to expiration. Thus, for each stock, we form a time-series of monthly options returns.

Also we require some volatility control variables in options for both in-sample analysis and strategies post-formation regressions. We use the Coval and Shumway [2001] zero-beta straddle factor (ZBStraddle) and the Goyal and Saretto [2009] delta-hedged call factor (DHCall) as aggregate volatility factors constructed by ourselves following the procedure described in the papers. The zero-beta straddle factor is the excess return of a zero-beta S&P100 index ATM straddle, calculated as

$$r_{ZBS} = \frac{-C\beta_C + S}{P\beta_C - C\beta_C + S} r_C + \frac{P\beta_C}{P\beta_C - C\beta_C + S} r_P \quad . \tag{1.13}$$

r_{ZBS}, r_C, r_P are the returns of the straddle, call and put, respectively, and β_C is the market beta of the Call:

$$\beta_C = \frac{S}{C} \Delta_C \quad , \tag{1.14}$$

where Δ_C is the delta of the option. The delta-hedged call factor is the delta-hedged S&P100 index call, a strategy consisting in a long position in the option contract and a short position in delta shares of the underlying, return calculated as

$$r_{DH}^{C} = \frac{\max(S_T - K, 0) - |\Delta_C| S_T}{C - |\Delta_C| S_0} - 1 \tag{1.15}$$

1.5 Descriptive Analysis

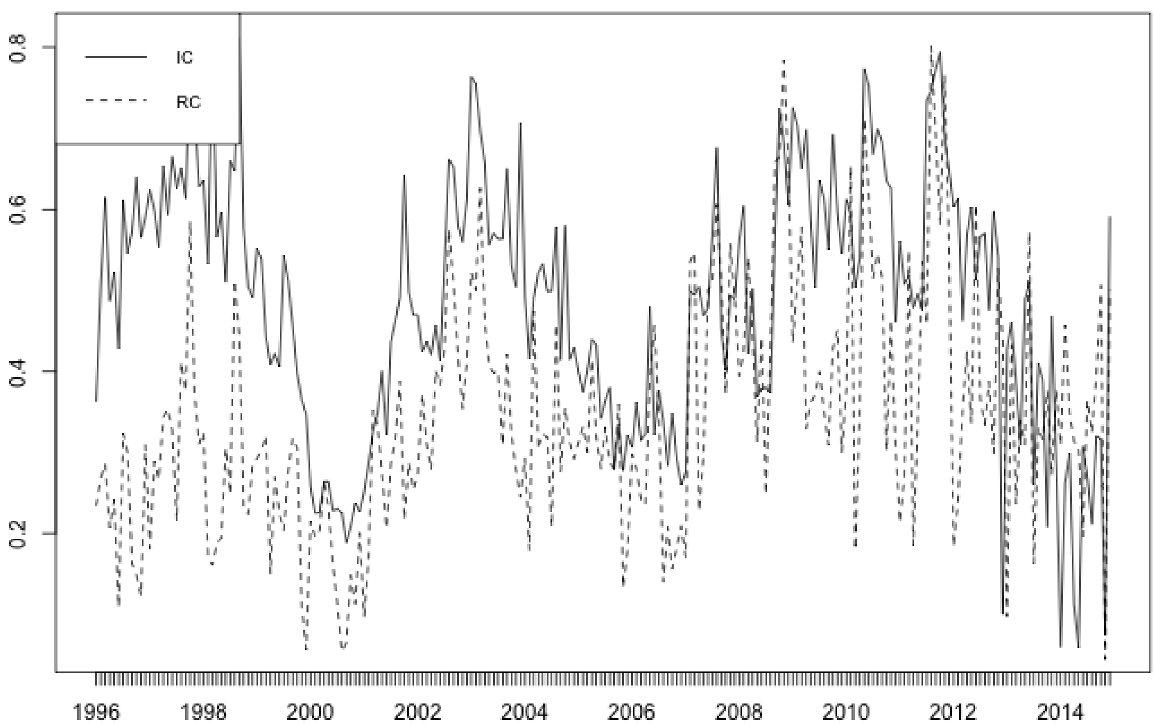
We compute the summary statistics of the time-series of our correlation estimates. Table 1.1 reports a monthly mean of 0.49, with values ranging from 0.06 to 0.81 and a standard deviation of 0.16 for IC_t. The mean of the realized correlation is 0.34, which indicates a positive average correlation risk premium of approximately 14% for the sample period. In Figure 1.1, we plot the implied and realized correlation estimates from January 1996 to December 2014. Visual inspection indicates that the implied correlation index is highly serially correlated with a first order autocorrelation coefficient of 0.75. Hence, we employ daily innovations in implied correlation, as the first difference in IC_t ($\Delta\mathrm{IC}_t$), in the cross-section analysis, displayed in the last row of Table 1.1. $\Delta\mathrm{IC}_t$ has a mean of zero, a standard deviation of 0.08, and a first order autocorrelation coefficient of -0.39.

Table 1.1: Summary statistics

	Mean	Median	St. Dev.	Max	Min
IC_t	0.486	0.498	0.156	0.811	0.059
RC_t	0.343	0.319	0.145	0.801	0.045
$\Delta\mathrm{IC}_t$	0	-0.001	0.079	0.327	-0.249

Figure 1.1 shows that the implied correlation is higher than the realized correlation for most of the sample period, which indicates a positive correlation risk premium. Consistent with other studies (see for instance Cosemans et al. [2011]; Driessen et al. [2013]), the figure also reveals that correlation increases at times of stress or during periods of market uncertainty. We observe that some of the peaks for both measures of correlation take place at the same time as events such as the Long Term Capital Management default and the Russian crisis in 1998, the Iraq war in 2003, the 2008 Global Financial Crisis, and the European debt crisis. During the times of these events, we also observe that the difference between these two measures decreases, making the correlation risk premium negative in some periods.

Figure 1.1: Monthly IC and RC estimates



Chapter 2

Time Series Analysis

We study the predictive power of the implied correlation in the options market. The empirical methodology relies on a standard regression model of the market risk premium on the lagged implied correlation, and standard control predictors for different return horizons h:

$$r^o_{t+h} = \alpha_h + \beta_h \mathrm{IC}_t + \gamma_h \mathrm{X}_t + \varepsilon_t,$$

We examine forecasts of the cumulative option returns[1]. Hence, $r^o_{t+h} = \frac{1}{h}\sum_{\mathrm{i}=1}^{h} r^o_{t+\mathrm{i}}$, is the (scaled by the horizon h) cumulative option market return from $t+1$ to $t+h$, where $r_{t+\mathrm{i}}$ is the hold-to-maturity naked put return of the S&P100 contract at time $t+\mathrm{i}$ calculated as 1.12. Since the returns are monthly hold-to-maturity, there are no overlapping days when calculating the accumulated return. The implied correlation is our main independent variable. The coefficient of interest, β_h, is expected to be negative. This is consistent with risk averse investors perceiving states of high marketwide correlation as an increase in aggregate risk, which induces a search for hedge strategies, such as options, rising contract prices and dropping future returns. Finally, X_t is a set of control predictors including: implied variance and volatility factors, variance risk premium, correlation risk premium, and volatility factors in options such as delta hedged call and zero beta straddle.

To help with the interpretation of the estimated coefficients, all the explanatory variables are standardized to have mean zero and unit variance. Since we take cumulative returns, the regression involves overlapping observations, which induces serial correlation in the residuals. We adjust the calculation of standard errors using the Hodrick and Prescott [1997] procedure. In the same way, the variation of the dependent variable explained by IC_t must be interpreted carefully, as Boudoukh et al. [2008] and Cochrane [2005] argue, since the adjusted R^2 when using overlapping observations tends to increase with the return horizon.

[1]We also examine predictions of the monthly option returns h months in the future.

2.1 In-Sample Results

Table 2.1 presents the simple regression results for the cumulative returns from one month (1) up to six months (6) in the future, regressed on the lagged IC_t. We find that the implied correlation is significant at 1%, decreasing at larger forecast horizons, therefore the effect of the IC_t is immediately incorporated into the Put prices. The economic importance is high; a one standard deviation increase in IC_t translates into a 22% decrease when predicting cumulative returns three months ahead.

Table 2.1: Simple Regression

	Cumulative Naked Put Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
IC	-0.222^{***}	-0.226^{***}	-0.220^{**}	-0.206^{**}	-0.207^{**}	-0.198^{*}
	(0.078)	(0.166)	(0.273)	(0.375)	(0.494)	(0.616)
Constant	-0.243^{**}	-0.244^{**}	-0.242^{**}	-0.240^{**}	-0.240^{**}	-0.242^{**}
	(0.101)	(0.203)	(0.306)	(0.407)	(0.511)	(0.617)
Adjusted R^2	0.017	0.039	0.056	0.062	0.080	0.084
Note:						*p<0.1; **p<0.05; ***p<0.01

The results show statistical and economic significance between IC_t and future option returns. Table 2.2 shows the results for quarterly cumulative returns when IC_t is included along with different control variables. We present different combinations, including IC_t and typical risk factors as regressors, such as volatility, variance risk premium, implied volatility, correlation risk premium. Furthermore, if the implied correlation reflects liquidity effects, then part of the predictive power of IC_t could be due to this liquidity component. Hence, we also include the illiquidity measure for the stock market LIQ_t.

Table 2.2: Quarterly Regressions

	Cumulative Naked Put Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
IC	-0.220^{**}	-0.238^{**}	-0.234^{**}	-0.202^{**}	-0.258^{**}	-0.202^{**}
	(0.273)	(0.354)	(0.343)	(0.282)	(0.323)	(0.278)
VIX		0.033				
		(0.285)				
IV			0.028			
			(0.259)			
LIQ				-0.061		
				(0.294)		
VRP					-0.078	
					(0.464)	
vVRPi					0.015	
					(0.494)	
CRP					0.070	
					(0.340)	
dhCall						-0.261^{**}
						(0.340)
zbStraddle						-0.261^{*}
						(0.411)
Constant	-0.242^{**}	-0.243^{**}	-0.243^{**}	-0.243^{**}	-0.243^{**}	-0.243^{**}
	(0.306)	(0.306)	(0.306)	(0.305)	(0.306)	(0.306)
Adjusted R^2	0.056	0.052	0.052	0.061	0.048	0.064
Note:						*p<0.1; **p<0.05; ***p<0.01

2.2 Out-of-Sample Results

The main findings indicate that the implied correlation has strong in-sample predictive power on the return of Put options. We want to extend this result to an out-of-sample analysis. In this section we explore whether implied correlation also exhibits out-of-sample forecasting power, in the sense of beating the historical average return.

We split the data, into an estimation window and a testing window. With the data from the estimation window we adjust our model, to further estimate the error terms of the model (ε_1) and a benchmark model (ε_2) using the data from the testing window. We use the historical average as our benchmark model.

$$
\begin{aligned}
\varepsilon_{1,t+h} &= rx^m_{t+h} - \widehat{rx^m_{t+h}}, \\
\varepsilon_{2,t+h} &= rx^m_{t+h} - \overline{rx^m_t},
\end{aligned}
$$

Then we calculate the average of the errors of each of the models: $\text{MSE}_1 = \sum_{t=1}^{T_{\text{test}}} \varepsilon^2_{1,t+h}$ and $\text{MSE}_2 = \sum_{t=1}^{T_{\text{test}}} \varepsilon^2_{2,t+h}$. We evaluate the predictive performance by employing the out-of-sample R^2, following Campbell and Thompson [2008]; and MSE-F, that is McCracken [2007] F-statistic.

$$
\begin{aligned}
R^2_{\text{out}} &= 1 - \frac{MSE_1}{MSE_2} \\
\text{MSE-F} &= (T-h+1) \text{ x } \left(\frac{\text{MSE}_2 - \text{MSE}_1}{\text{MSE}_1}\right)
\end{aligned}
$$

A positive out-of-sample R^2 indicates that the predictive regression displays a lower average mean squared error than the historical average return. In addition, due to small samples, we obtain critical values from a bootstrap procedure, to provide statistical significance. The bootstrap follows Kilian [1999] and imposes the NULL of no predictability for calculating the critical values. The data generating process is assumed to be

$$
\begin{aligned}
y_{t+1} &= \alpha + u_{1t+1} \\
x_{t+1} &= \mu + \rho \cdot x_t + u_{2t+1}
\end{aligned}
$$

and the bootstrap for calculating power assumes the data generating process is

$$
\begin{aligned}
y_{t+1} &= \alpha + \beta \cdot x_t + u_{1t+1} \\
x_{t+1} &= \mu + \rho \cdot x_t + u_{2t+1}
\end{aligned}
$$

and both coefficients, β and ρ, are estimated by OLS using the full sample of observations, with the residuals stored for sampling. We then generate 1000 bootstrapped time series by drawing with replacement from the residuals. And from this time series we obtain the critical values to compare with.

Table 2.3 presents the results for our out-of-sample analysis. The first two rows show our findings when the forecast horizon is quarterly, whereas the last two rows show the results for semi-annual return horizons. Consistent with the in-sample results, we find that the implied correlation delivers positive out-of-sample R^2s for both forecast horizons. The mean of the difference in the squared errors is always positive, which indicates that predictions using historical average returns present a higher prediction error than those employing implied correlation. The difference is statistically significant at a 1% level when the prediction horizon is six months ahead. The last column show the results for a specification that includes the interest rate, and achieve a much higher explanatory power. This finding could lead to future new research involving the term structure. Taking into account the short sample period used to perform out-of-sample experiments, this finding provides weak evidence of option returns predictability at a quarterly and semi-annual forecast horizon.

Table 2.3: Out-of-Sample Results

	Naked Put	
	IC	IC + IRrel
	h=3	
MSE-F	0.75***	42.35***
R^2	0.4%	8%
	h=6	
MSE-F	7.71***	76.29***
R^2	4%	9%

In summary, these findings document evidence consistent with the in-sample results. The implied correlation exhibits better out-of-sample performance than the historical average market excess returns, supported by a positive out-of-sample R^2 and a MSE-F statistic higher than the critical values obtained by bootstrapping. For robustness of our findings, we also check the MSE-T statistic, the Diebold and Mariano [1995] t-statistic modified by Harvey et al. [1997], calculated as

$$\text{MSE-T} = \sqrt{T + 1 - 2 \cdot h + h \cdot (h-1)/T} \cdot \frac{\text{MSE}_2 - \text{MSE}_1}{\widehat{SE}(\text{MSE}_2 - \text{MSE}_1)}$$

and we get the same results.

Chapter 3

Cross Section Analysis

3.1 Portfolio Strategies

To examine how aggregate implied correlation affects the cross section of put option returns, we sort equity returns based on their sensitivities to the aggregate implied correlation innovation factor using the following multifactor model:

$$r_t^{\mathrm{i}} = \alpha_t^{\mathrm{i}} + \beta_{\Delta\mathrm{IC}}^{\mathrm{i}} \Delta\mathrm{IC}_t + X_t + \varepsilon_t^{\mathrm{i}}, \tag{3.1}$$

where ΔIC represents the innovations in the aggregate implied correlation factor and $\beta_{\Delta\mathrm{IC}}^{\mathrm{i}}$ are the loadings on aggregate implied correlation innovations. X_t represents other risk factors. We include the three factors of Fama and French [1993] and Amihud [2002] liquidity factor calculated for the stock market.

r_t^{i} corresponds to the equity return of the S&P100 constituent i at time t. At the end of each month, we sort stocks into quintiles, based on their past loadings $\beta_{\Delta IC}^{\mathrm{i}}$. Past loadings are obtained by running the regression over the previous 3 months [1] consistent with our findings in 2.1. Hence, portfolio 1 contains the firms with the lowest $\beta_{\Delta IC}^{\mathrm{i}}$ loadings, whereas portfolio 5 is composed by the companies with the highest $\beta_{\Delta IC}^{\mathrm{i}}$ loadings. Then, we construct equal weighted portfolios using put option returns written on those firms.

Table 3.1 reports summary statistics for the time-series option returns of quintile portfolios. The first two rows report the mean and standard deviation of monthly total simply returns. The last column reports the statistics for the spread in monthly returns between portfolio 5 and portfolio 1. We also exhibit the maximum and minimum return obtained by each portfolio during the period studied. Consistent with our intuition, the mean returns of the portfolios show a monotonic increasing from the first portfolio, which contains the lowest β of the regression 3.1, with a -15% monthly average, to last portfolio, which contains the firms with the highest β, with a monthly average of -5%. The extreme portfolios present a mean difference of 10% significant at the 1%. The last row shows the β of each portfolio, increasing monotonically (by construction) from -0.07 to 0.12.

[1] We check the robustness of our findings by employing 6 and 9 months of estimation window.

Table 3.1: Portfolios Summary Statistics

Statistic	1	2	3	4	5	$5-1$
Mean	-0.15	-0.13	-0.11	-0.10	-0.05	0.10
St. Dev.	0.88	0.90	0.90	0.82	0.92	0.52
Max	4.23	4.05	4.01	3.34	3.72	2.52
Min	-1.00	-1.00	-1.00	-0.98	-1.00	-1.65
β	-0.07	0	0.02	0.05	0.12	

Hence, a trading strategy based on selling put option contracts in portfolio 1 and buying put option contracts in portfolio 5 every month, reports a 13% monthly return. Also this return is less volatile than the portfolios, with a 0.52 standard deviation, below the range of the portfolios between 0.82 and 0.92. Likewise, since the strategy involves two positions, the minimum monthly return obtained is less than that of portfolios, with -165%, however the maximum is less, with a 252% compared to an average 387% of 5 portfolios. In the appendix we exhibit the value-weighted portfolios and their summary statistics. The difference between the extreme portfolios is 13% in average, significant at 1%, and with a 0.76 standard deviation.

We further explore a second strategy, which we called *an implied correlation timing strategy*, where investors change the long-short position depending on their forecast of implied correlation the next period. If an investor forecast an increase in implied correlation then she goes long on portfolio 5 and short on portfolio 1. On the contrary, if she forecasts a decrease in implied correlation, then she follows exactly the opposite strategy and goes long on portfolio 1 and short on portfolio 5.

If an investor forecasts an increase in implied correlation the next month, consistent with implied correlation being a risk factor, then stocks that correlates positively with implied correlation are more demanded, since they help to protect investors against periods of high correlation. Then the demand on put options written on those stocks should decrease. In other words, the demand on options of portfolio 5 decreases and price of put options decreases. On the other hand, when correlation is forecast to be high stocks that are negatively correlated to the implied correlation are riskier, since they behave bad in periods of high correlation. Consequently, the demand of put options written on those stocks increases, i.e., options of portfolio 1 are more demanded and their price increases. Therefore, it is optimal to go long on portfolio 5 and short on portfolio 1. When investors forecast a low correlation the next month, they switch strategy by going long on portfolio 1 and short on portfolio 5.

In particular at the end on each month $t+1$, we have a forecast $\widehat{IC}_{t+1}$ of implied correlation. By using the previous three months as estimation window, from $t-2$ to t, we estimate an AR(p) process for implied correlation. Then $\widehat{IC}_{t+1}$ is obtained as follows:

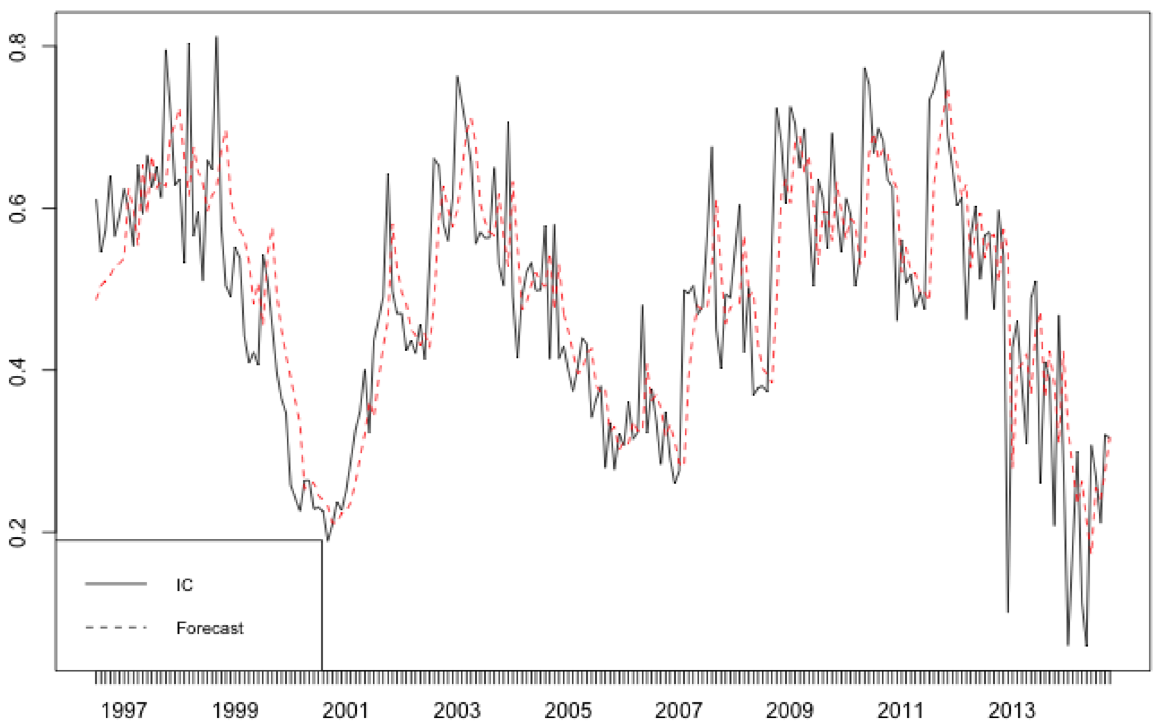
$$\widehat{IC}_{t+1} = \hat{\alpha}_0 + \hat{\alpha}_1 IC_t ... + \hat{\alpha}_p IC_{t-p+1},$$

where p, the number of lag terms is obtained employing the corrected Aikaike information criterion. Then, the *implied correlation timing strategy* is:

$$Strategy = \begin{cases} \text{Long portfolio 5, Short portfolio 1} & \text{if } IC_t < \widehat{IC}_{t+1} \\ \text{Long portfolio 1, Short portfolio 5} & \text{otherwise.} \end{cases} \tag{3.2}$$

For this trading strategy we use a 6-month estimation window, which reports a better forecast of the implied correlation, but the results are still significant for 3 and 9 months. The forecast and the corresponding realized implied correlation are shown in the Figure 3.1.

Figure 3.1: Forecast vs IC



This strategy reports a 10% monthly return on average, statistically significant at 1%, with a 0.52 standard deviation. In the same way as our previous strategy, we incurred in two options transactions, so the minimum return obtained is -142%, which is equivalent to losing money in both positions. On the other hand, the maximum return obtained in a month is 222%, less than the average of the maximum return of the 5 portfolios.

Both strategies, the 5-1 spread portfolio and the IC timing strategy, generate significant returns both economically and statistically.

3.2 Post-Formation Controls for Risk

We want to check if the returns of the previously described strategies answer to some known risk in the market or it is something new, arguably correlation risk. We regress the following linear factor model:

$$R_{p,t} = \alpha + \beta F_t + \varepsilon,$$

where $R_{p,t}$ are the returns of the strategies obtained in Section 3.1 and F_t denotes a matrix of risk factors. We consider different factors, that have been studied in the past, with a significant effect on the options returns, such as (1) Fama and French [1993] three factor model, the Carhart [1997] momentum factor model (2), Goyal and Saretto [2009] show that the difference between implied volatility and realized volatility significantly affect the cross-section of option returns, so we include the excess zero-beta straddle factor and the delta-hedged call factor return as option volatility factors (3); and more recently, Bai et al. [2019] find that firm-level variance risk premium significantly predicts future option returns, so we also use a Variance Risk Premium factor (4). Since all the factors are spread traded portfolios, the intercept α from the regression can be interpreted in the usual sense of mispricing relative to the factor model and hence, the option returns described in this paper are not related to standard aggregate sources of risk.

Table 3.2 reports the estimated coefficients. For the 5-1 strategy, in specifications (1), (2), and (3) we observe that alpha remains statistically significant, at 1%, and very similar, even higher, to the average raw returns. But for specification (4) we see a drop in alpha magnitude and significance, possibly attributable to the VRP variable. For the IC Timing Strategy, the loadings on Fama and French Market and Value factors are significant. The main difference between signs in loadings for both strategies relies on the HML and VRP loadings, this implies that our second strategy earn abnormal returns even with positive exposure to Value risk and positive Variance Risk Premium. The alpha of the IC Timing Strategy remains constant, and become more significant, both economically and statistically, in specification (4), reaching an alpha of 14%. We conclude that our returns can not be associated with any of this common source of risk. In the appendix we show the same results for the value-weighted portfolios.

Table 3.2: Post Formation Regressions

	5-1 Strategy				*IC Timing Strategy*			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Alpha	0.11^{***}	0.10^{***}	0.09^{***}	0.06^{*}	0.11^{***}	0.11^{***}	0.11^{***}	0.14^{***}
	(0.03)	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)
Mkt.RF	-0.40	-0.05		0.09	-1.67^{**}	-1.68^{**}		-1.58^{*}
	(0.68)	(0.73)		(0.77)	(0.67)	(0.72)		(0.75)
SMB	-1.05	-1.07		-0.99	-0.92	-0.93		-0.89
	(1.23)	(1.23)		(1.23)	(1.21)	(1.21)		(1.20)
HML	-1.84^{*}	-1.38		-1.05	2.32^{**}	2.32^{**}		2.01^{*}
	(1.01)	(1.08)		(1.08)	(1.00)	(1.06)		(1.06)
Mom		0.91		1.19		0.0		-0.22
		(0.74)		(0.79)		(0.73)		(0.78)
dhCall			-3.09	-0.19			-0.39	0.29
			(2.47)	(2.76)			(2.45)	(2.70)
zbStraddle			-0.13	-0.04			0.10	0.11
			(0.11)	(0.12)			(0.11)	(0.12)
VRP				-0.69^{*}				0.66^{*}
				(0.29)				(0.29)
R^2	0.02	0.03	0.01	0.06	0.06	0.06	0.03	0.10

Note: *p<0.1; **p<0.05; ***p<0.01

3.3 Transaction Costs

Next, we investigate whether the strategies described in section 3.1 still have profits after considering transaction costs, measured by the bid-ask spread. Both strategies consist in sell and buy two Put options simultaneously, therefore we face transaction costs two times each month. To this end, we consider 4 measures of effective spread; 0%, 25%, 50% and 100%. Specifically, an effective spread equal to 0% indicates options with transaction cost exactly at the midpoint of the bid and ask price quotes. On the other hand, an effective spread of 25% or 50% would correspond to a transaction cost of 0.25 or half of the bid-ask spread, respectively. Finally, 100% is equivalent to paying the entire bid-ask quote. These transaction costs are considered each time we rebalance our position.

Table 3.3 reports the results. The first row of each sub-table presents the results for the time-series returns; the second row exhibits the alpha of the (1) specification of the post-formation regression as in Table 3.2; and the third row of each sub-table represents the alpha of the (4) specification. For the first strategy, the 5-1 spread, the impact is higher. The average return decrease drastically, from a 10% monthly return to a 2% monthly return taking the entire bid-ask spread into account, and the α of the post-formation regression decrease from 6% to -1% in specification (4) in table 3.2, and from 11% to 3% in specification (1). For our second strategy, the return and alphas are significant, and the economic relevance remains, notoriously on the (4) specification, which yields a 7% monthly return even at 100% of the effective spread. Muravyev and Pearson [2015] show that the effective spread measure taken into account by the high frequency trade timing ability is 53% of the quoted spread average, and Mayhew [2002] show that the effective spreads for equity options are large in absolute terms but small relative to the quoted spreads, resulting typically in a ratio of effective to quoted spread less than 0.5. So at 50% effective spread the returns are still significantly positive at the 5%, with an expected monthly return of 6% and an 11% alpha of the post-formation regression. We conclude that although transaction costs have an impact on the monthly returns, they do not eliminate the statistical nor economic significance of our results, at very reasonable estimates of effective spreads, particularly on the IC Timing Strategy.

Table 3.3: Transaction Costs

		ESPR/QSPR			
5-1	Midpoint	25%	50%	75%	100%
All	0.10***	0.08**	0.06*	0.04	0.02
Alpha (1)	0.11***	0.09**	0.07*	0.05	0.03
Alpha (4)	0.06*	0.04	0.03	0.01	-0.01
IC Timing Strategy	Midpoint	25%	50%	75%	100%
All	0.10***	0.08**	0.06*	0.05	0.03
Alpha (1)	0.11***	0.09***	0.07**	0.05	0.04
Alpha (4)	0.14***	0.13***	0.11***	0.09**	0.07*
Note:			*p<0.1; **p<0.05; ***p<0.01		

3.4 Sub-Sample Analysis

We investigate the robustness of our results in 3.1 over different sub-samples. To this end, we split the data in five periods, based on the NBER expansions and contractions periods: from January 1996 to February 2001, from March 2001 to November 2001, from December 2001 to November 2007, from December 2007 to June 2009, and from July 2009 to December 2014. Table 3.4 reports the summary statistics for the portfolios. The last two columns exhibit the monthly average of the returns for both strategies. The recessions periods are marked with an **R**.

Table 3.4: Summary Statistics Over Different Subsamples

Subperiod	1	2	3	4	5	$5-1$	*IC TS*
Jan 1996 - Feb 2001	-0.12	0.02	0.01	0.05	0.13	0.24	0.08
Mar 2001 - Nov 2001 **R**	0.50	0.58	0.45	0.60	0.62	0.12	-0.42
Dec 2001 - Nov 2007	-0.16	-0.18	-0.19	-0.17	-0.15	0.01	0.05
Dec 2007 - Jun 2009 **R**	0.02	0.11	0.16	0.12	0.30	0.28	-0.09
Jul 2009 - Dec 2014	-0.27	-0.30	-0.22	-0.25	-0.20	0.06	0.07

It is not surprising that the returns of the options are positive, and very high, in recessions, when most of the stocks in the market are in downside. In every sub-sample our 5-1 spread is positive, becoming even higher in recessions, up to 28% monthly return. On the other hand, our IC Timing Strategy varies with the economic cycle: during expansions remains positive, but during contractions it becomes negative. Possibly, our second strategy relies on expansive markets, in the third and fifth periods generates higher returns than our 5-1 spread; this conclusion is also justified by the significance of the Market factor in Table 3.2. It should be noted that recession periods have few observations.

Conclusion

This paper shows that aggregate implied correlation is an important indicator of market-wide risk. We estimate risk-neutral expectations of second moments from option prices of the S&P100 index and its individual constituents.

We document that implied correlation has strong forecasting power for the option market return, both in-sample and out-of-sample, and it is responsible for a notorious difference in the cross-section of individual option contracts. A strategy that take advantage of this spread produces a 10% average monthly return. A second strategy that uses a prediction of the IC to anticipate investor behavior, produces a 10% average monthly return.

We find that our strategies' returns are not related to obvious sources of risk, generating significative alphas after controlling for several factors, although it is not enough evidence to conclude that these are true alphas. The alphas of the post-formation regression become even higher, with a 14% for our IC Timing Strategy.

Including transaction costs in the form of effective spread at the time of construction of portfolios, our findings remain the same, at reasonable estimates of effective spreads for the 5-1 spread strategy, and even with inclusion of the entire effective bid-ask spread as transaction cost, our IC Timing Strategy produces 7% monthly return significant at 10%. And if we split the sample, to control for the effect of possible particular periods, we reach the same conclusions. Our 5-1 spread returns holds, regardless the economical cycle, but our IC Timing Strategy fails in contractions periods.

Bibliography

Yakov Amihud. Illiquidity and stock returns: cross-section and time-series effects. *Journal of financial markets*, 5(1):31–56, 2002.

Andrew Ang, Robert J Hodrick, Yuhang Xing, and Xiaoyan Zhang. The cross-section of volatility and expected returns. *The Journal of Finance*, 61(1):259–299, 2006.

Haorui Bai, Xiaoyan Zhang, and Hao Zhou. Option return predictability and variance risk premium. *Available at SSRN 3414841*, 2019.

Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of political economy*, 81(3):637–654, 1973.

Oleg Bondarenko. Why are put options so expensive? *The Quarterly Journal of Finance*, 4 (03):1450015, 2014.

Jacob Boudoukh, Matthew Richardson, and Robert F Whitelaw. The myth of long-horizon predictability. *The Review of Financial Studies*, 21(4):1577–1605, 2008.

Mark Broadie, Mikhail Chernov, and Michael Johannes. Understanding index option returns. *The Review of Financial Studies*, 22(11):4493–4529, 2009.

Adrian Buss and Grigory Vilkov. Measuring equity risk with option-implied correlations. *The Review of Financial Studies*, 25(10):3113–3140, 2012.

John Y Campbell and Samuel B Thompson. Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies*, 21(4):1509–1531, 2008.

Mark M Carhart. On persistence in mutual fund performance. *The Journal of finance*, 52 (1):57–82, 1997.

John H Cochrane. Asset pricing (revised edition), 2005.

Mathijs Cosemans et al. The pricing of long and short run variance and correlation risk in stock returns. Technical report, Working Paper, 2011.

Joshua D Coval and Tyler Shumway. Expected option returns. *The journal of Finance*, 56 (3):983–1009, 2001.

Francis X Diebold and RS Mariano. Comparing forecast accuracy. *Journal of Business and*, 1995.

Joost Driessen, Pascal J Maenhout, and Grigory Vilkov. The price of correlation risk: Evidence from equity options. *The Journal of Finance*, 64(3):1377–1406, 2009.

Joost Driessen, Pascal J Maenhout, and Grigory Vilkov. Option-implied correlations and the price of correlation risk. 2013.

Eugene F Fama and Kenneth R French. Common risk factors in the returns on stocks and bonds. *Journal of*, 1993.

Amit Goyal and Alessio Saretto. Cross-section of option returns and volatility. *Journal of Financial Economics*, 94(2):310–326, 2009.

David Harvey, Stephen Leybourne, and Paul Newbold. Testing the equality of prediction mean squared errors. *International Journal of forecasting*, 13(2):281–291, 1997.

R. Hodrick and C. E. Prescott. Postwar US business cycles: an empirical investigation. *Journal of Money Credit & Banking*, 29:1–16, 1997.

Lutz Kilian. Exchange rates and monetary fundamentals: what do we learn from long-horizon regressions? *Journal of applied Econometrics*, 14(5):491–510, 1999.

Ian Martin. Simple variance swaps. Technical report, National Bureau of Economic Research, 2011.

Stewart Mayhew. Competition, market structure, and bid-ask spreads in stock option markets. *The Journal of Finance*, 57(2):931–958, 2002.

Michael W McCracken. Asymptotics for out of sample tests of granger causality. *Journal of econometrics*, 140(2):719–752, 2007.

Dmitriy Muravyev and Neil D Pearson. Options trading costs are lower than you think. *The Review of Financial Studies*, 2015.

Appendix

Figure 3.2: Daily IC innovation ($\Delta \mathrm{IC}_t$)

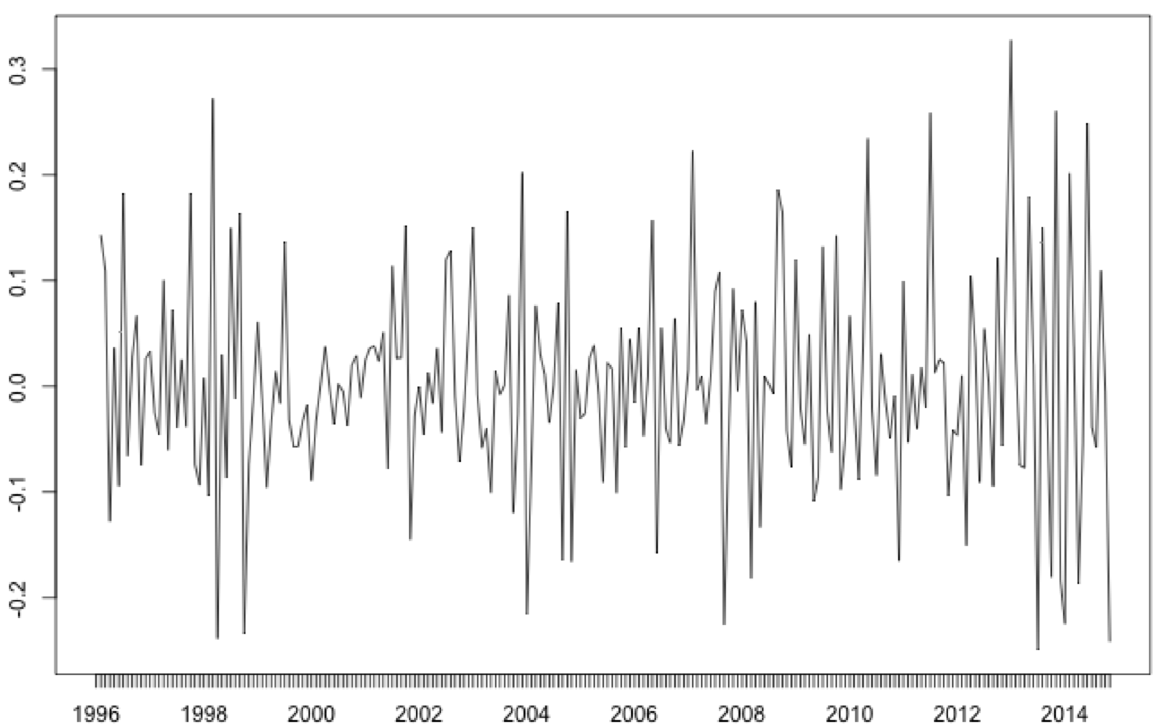


Table 3.5: Predictive Return Regressions

	Dependent variable:					
	Cumulative Naked Put Returns h=1					
	(1)	(2)	(3)	(4)	(5)	(6)
IC	-0.222^{***}	-0.240^{***}	-0.248^{***}	-0.213^{***}	-0.217^{**}	-0.184^{**}
	(0.078)	(0.091)	(0.090)	(0.080)	(0.107)	(0.082)
VIX		0.033				
		(0.094)				
IV			0.052			
			(0.086)			
LIQ				-0.031		
				(0.080)		
VRP					-0.105	
					(0.221)	
vVRPi					-0.055	
					(0.206)	
CRP					-0.018	
					(0.140)	
dhCall						-0.566^{*}
						(0.329)
zbStraddle						-0.489^{*}
						(0.278)
Constant	-0.243^{**}	-0.243^{**}	-0.243^{**}	-0.243^{**}	-0.243^{**}	-0.244^{**}
	(0.101)	(0.101)	(0.101)	(0.101)	(0.101)	(0.101)
Adjusted R^2	0.017	0.013	0.013	0.013	0.015	0.034
Note:				*p<0.1; **p<0.05; ***p<0.01		

Table 3.6: Predictive Return Regressions

	Dependent variable:					
	Cumulative Naked Put Returns h=2					
	(1)	(2)	(3)	(4)	(5)	(6)
IC	-0.226^{***}	-0.240^{**}	-0.241^{**}	-0.209^{**}	-0.233^{**}	-0.205^{**}
	(0.166)	(0.215)	(0.205)	(0.171)	(0.209)	(0.166)
VIX		0.025				
		(0.185)				
IV			0.030			
			(0.159)			
LIQ				-0.059		
				(0.156)		
VRP					-0.055	
					(0.310)	
vVRPi					-0.035	
					(0.316)	
CRP					0.007	
					(0.239)	
dhCall						-0.312^{**}
						(0.266)
zbStraddle						-0.338^{**}
						(0.306)
Constant	-0.244^{**}	-0.244^{**}	-0.245^{**}	-0.244^{**}	-0.244^{**}	-0.244^{**}
	(0.203)	(0.203)	(0.203)	(0.203)	(0.203)	(0.203)
Adjusted R^2	0.039	0.035	0.035	0.037	0.032	0.048
Note:				$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$		

Table 3.7: Predictive Return Regressions

	Dependent variable:					
	Cumulative Naked Put Returns h=4					
	(1)	(2)	(3)	(4)	(5)	(6)
IC	-0.206^{**}	-0.217^{*}	-0.212^{*}	-0.188^{*}	-0.232^{**}	-0.188^{**}
	(0.375)	(0.495)	(0.476)	(0.397)	(0.411)	(0.382)
VIX		0.020				
		(0.393)				
IV			0.013			
			(0.354)			
LIQ				-0.064		
				(0.313)		
VRP					-0.0004	
					(0.709)	
vVRPi					-0.052	
					(0.792)	
CRP					0.040	
					(0.395)	
dhCall						-0.223^{**}
						(0.358)
zbStraddle						-0.244^{**}
						(0.377)
Constant	-0.240^{**}	-0.240^{**}	-0.240^{**}	-0.239^{**}	-0.240^{**}	-0.241^{**}
	(0.407)	(0.407)	(0.407)	(0.408)	(0.407)	(0.407)
Adjusted R^2	0.062	0.059	0.058	0.064	0.055	0.071
Note:				$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$		

Table 3.8: Predictive Return Regressions

	Dependent variable:					
	Cumulative Naked Put Returns h=5					
	(1)	(2)	(3)	(4)	(5)	(6)
IC	-0.207^{**}	-0.227^{*}	-0.215^{*}	-0.189^{*}	-0.222^{**}	-0.201^{**}
	(0.494)	(0.653)	(0.627)	(0.520)	(0.535)	(0.496)
VIX		0.037				
		(0.524)				
IV			0.016			
			(0.465)			
LIQ				-0.063		
				(0.422)		
VRP					0.034	
					(0.746)	
vVRPi					-0.062	
					(0.791)	
CRP					0.018	
					(0.408)	
dhCall						-0.073
						(0.458)
zbStraddle						-0.075
						(0.466)
Constant	-0.240^{**}	-0.240^{**}	-0.240^{**}	-0.239^{**}	-0.241^{**}	-0.241^{**}
	(0.511)	(0.511)	(0.511)	(0.512)	(0.511)	(0.511)
Adjusted R^2	0.080	0.078	0.076	0.084	0.071	0.074
Note:				$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$		

Table 3.9: Predictive Return Regressions

	Dependent variable:					
	Cumulative Naked Put Returns h=6					
	(1)	(2)	(3)	(4)	(5)	(6)
IC	-0.198^{*}	-0.222^{*}	-0.207	-0.179^{*}	-0.216^{*}	-0.194^{*}
	(0.616)	(0.791)	(0.762)	(0.640)	(0.690)	(0.616)
VIX		0.043				
		(0.616)				
IV			0.018			
			(0.531)			
LIQ				-0.070		
				(0.517)		
VRP					0.037	
					(1.115)	
vVRPi					-0.046	
					(1.195)	
CRP					0.026	
					(0.445)	
dhCall						-0.058
						(0.499)
zbStraddle						-0.068
						(0.504)
Constant	-0.242^{**}	-0.242^{**}	-0.242^{**}	-0.241^{**}	-0.243^{**}	-0.242^{**}
	(0.617)	(0.617)	(0.617)	(0.618)	(0.616)	(0.617)
Adjusted R^2	0.084	0.083	0.081	0.091	0.075	0.078
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 3.10: Value-Weighted Portfolios Summary Statistics

Statistic	1	2	3	4	5	5 − 1
Mean	−0.09	−0.09	−0.04	−0.09	0.04	0.13
St. Dev.	0.94	1.00	1.03	0.89	1.11	0.76
Max	4.14	4.94	4.37	3.47	3.91	3.99
Min	−1.00	−1.00	−1.00	−1.00	−1.00	−2.35
β	−0.07	0	0.02	0.05	0.11	

Table 3.11: Value-Weighted Post Formation Regressions

	5-1 Strategy				IC Timing Strategy			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Alpha	0.15***	0.15***	0.14***	0.14**	0.10*	0.10*	0.09*	0.13**
	(0.05)	(0.05)	(0.05)	(0.06)	(0.05)	(0.06)	(0.05)	(0.06)
Mkt.RF	−1.31	−1.32		−1.03	−1.84*	−1.96*		−1.92
	(1.00)	(1.08)		(1.14)	(1.03)	(1.12)		(1.18)
SMB	−2.01	−2.01		−1.85	−1.03	−1.02		−0.95
	(1.80)	(1.81)		(1.83)	(1.88)	(1.88)		(1.90)
HML	−2.24	−2.25		−1.98	3.57**	3.41**		3.11*
	(1.48)	(1.59)		(1.62)	(1.54)	(1.65)		(1.67)
Mom		−0.01		0.04		−0.32		−1.00
		(1.10)		(1.18)		(1.13)		(1.22)
dhCall			−4.34	−1.51			−3.76	−4.29
			(3.65)	(4.12)			(3.79)	(4.24)
zbStraddle			−0.11	−0.04			−0.10	−0.14
			(0.17)	(0.18)			(0.17)	(0.18)
VRP				−0.40				0.67
				(0.43)				(0.45)
R^2	0.02	0.02	0.002	0.04	0.04	0.04	0.01	0.06
Note:					*p<0.1; **p<0.05; ***p<0.01			

Table 3.12: Value-Weighted Transaction Costs

		ESPR/QSPR			
5-1	Midpoint	25%	50%	75%	100%
All	0.13**	0.11**	0.10*	0.08	0.06
Alpha (1)	0.15***	0.13***	0.12**	0.10*	0.08
Alpha (4)	0.14**	0.12**	0.10*	0.08	0.06
IC Timing Strategy	Midpoint	25%	50%	75%	100%
All	0.09**	0.07*	0.06	0.04	0.02
Alpha (1)	0.09*	0.08*	0.06	0.05	0.03
Alpha (4)	0.13**	0.11*	0.10	0.08	0.06
Note:		*p<0.1; **p<0.05; ***p<0.01			

Table 3.13: Value-Weighted Summary Statistics Over Different Subsamples

Subperiod	1	2	3	4	5	$5-1$	*IC TS*
Jan 1996 - Feb 2001	-0.03	0.02	0.15	0.04	0.30	0.33	0.08
Mar 2001 - Nov 2001 **R**	0.37	0.66	0.72	0.41	1.09	0.72	-0.18
Dec 2001 - Nov 2007	-0.12	-0.11	-0.16	-0.15	-0.04	0.08	0.14
Dec 2007 - Jun 2009 **R**	0.12	0.05	0.25	0.16	0.31	0.19	-0.12
Jul 2009 - Dec 2014	-0.24	-0.27	-0.18	-0.23	-0.20	0.04	0.05