A study of Implied Volatility Spreads and their relation to Stock Returns

To be defined

Marie Dyveke Olafsen
201808138

Supervisor: Jonas Nygaard Eriksen
Associate Professor, Aarhus University

Aarhus University, Department of Economics and Business Economics
M.Sc. Finance Master's Thesis

Hand-in date: June 1st 2023
The thesis may be published

Abstract

Contents

1 Introduction

Starting with classic asset pricing theory -> talking about risk factors -> which are prominent at the moment? -> some have shown that options are a factor -> Harvey and Siddique (2000) show skewness is relevant -> at odds with a mean variance optimizing investor -> more in line with prospect theory and other behavioural finance aspects, such as loss aversion

-> the option market is inherently forward looking - see appendix regarding theorem -> options in general or mispricings of options? -> use implied spread as a proxy for the mispricing of options -> the implied vol spread should be negative, as puts are more expensive, can protect against jump risk -> which factors affect the implied vol spread? -> why is this implied vol spread a good proxy for systematic risk in stocks?

-> is it just a sign of shortsale constraints? Or informed trading? -> or Pan and Poteshman proxy this by looking at trades initiated by the buyer, Ofek et al. also looks at the relation and find significant predictability from the short sale constraint

-> put call parity -> fundamentals -> borrowing lending rates differ -> transactions costs -> margin requirements and taxes -> option on stocks are american and not european (which means the put call parity is not an equality but an inequality)

-> Inherent forward looking property of the option market -> we can estimate the risk-neutral distribution of the stock price given a lot of different options on the stock according to formula (2) in Breeden and Litzenberger (1978) (and elaborated in appendix A in Isakas and Melenberg) which is under the Q-measure but can be approximated to the P-measure and used for forecasting stock prices. This is done by evaluting the differentiated call price with respect to the strike twice at a lot of different strikes at one maturity. This is however not feasible to estimate across a lot of different stocks, as their options might not be liquid enough to provide an accurate market view (due to either reduced interest, high transaction costs or too high margin requirements for the sellers).

2 Setting

The setting of this study is elaborated in this section. I will start out by explaining the implied volatility spread, how it is derived and why it is a relevant signal. Then I will go on to discuss the formation of portfolios and why this approach was decided on. Afterwards, the included test for monotonicity is described, and then followed up by the models I will employ to evaluate the signal and its significance compared to previous known risk factors. Lastly, I will draw up some usefull thoughts from the utility framework and include some relevant measures for evaluating the signal.

2.1 Implied Volatility

The Black-Scholes-Merton formula for a European call price is:

$$c = SN\left(d_1\right) - Ke^{-r(T-t)}N\left(d_2\right) \tag{2.1}$$

where $d_1 = \frac{\ln(S_0/K)+\left(r-\delta+\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}$ and $d_2 = \frac{\ln(S_0/K)+\left(r-\delta-\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}$.

The corresponding price for an European put is:

$$p = Ke^{-r(T-t)}N\left(-d_2\right) - SN\left(-d_1\right) \tag{2.2}$$

Following the notation in Cremers and Weinbaum (2010) the implied volatility spread is calculated as:

$$IV_{i,t}^{call} - IV_{i,t}^{put} = VS_{i,t} \tag{2.3}$$

and the signal is then estimated across time and stocks:

$$VS_{i,t} = \sum_{j=1}^{N_{i,t}} w_{j,t}^{i}\left(IV_{j,t}^{i,call} - IV_{j,t}^{i,put}\right) \tag{2.4}$$

where t, i and j specifies the time, stock and combination of strike and maturity. $w_{j,t}^{i}$ is the weight for the specific maturity, strike and time. This weight can be both set to a simple average or based on the open interest in the pair of options.

2.2 Portfolios

The analysis of the relation between stock returns and the signal defined above is conducted using portfolios. Thus I will split the different firms into portfolios with the use of the signal.

The choice of forming portfolios compared to analyzing specific stocks over a period of time, is made based on the flexibility of the setup. The approach requires no a priori assumptions about the effect, and also allows for the discovery of more than just a linear effect. This approach also has some diasadvantages, as the flexibility gives rise to a bigger risk of datasnooping.

The choice of using portfolios for the analysis is further made based on the extensive literature within asset pricing which makes the same choice. Furthermore, the choice is based on an approach looking the the effect of the signal on the entire American stock market, and not on specific stocks, therefore, an aggregation and combination of these stocks provides a clearer picture of the effect from the signal and reduces the noise from other individual risk factors.

The portfolios are formed using breakpoints in distribution of the signal. Thus I start by choosing the amount of portfolios I would like. The traditional choice would be to follow and make three portfolios with breakpoints at the 30 percentile and 70 percentile, to make portfolios consisting of the lowest 30% of the signal, then the middle 40% and the last portfolio of the highest 30% from the sample. In Cremers and Weinbaum (2010) they use five equal sized portfolios, and I will follow their example, but compare the results with portfolios sorted in the traditional way as mentioned above, and also with 10 portfolios.

The use of equal-sized portfolios stems from an ambition to evaluate the effect of the entire distribution of the signal and not just evaluate the tails as done in the traditional approach. Furthermore, my main analysis will be based on 5 equal sized portfolios to ensure a sufficient amount of stocks are in each. This means that I have a smaller sample of portfolio returns to evaluate, which I thought a reasonable tradeoff given the inclusion of the results from the analysis with different amount of portfolios in the appendix.

Furthermore, the signal identified above might not be the sole signal that we are interested in. A combination of the absolute value of the implied volatility spread and the recent change in the implied volatility spread might also provide some insights.

Therefore, I will conduct an analysis focusing only on the univariate portfolio sorting based on the value, and following that a bivariate portfolio sorting with both a dependent and independent sorting with the implied volatility spread as the first sorting factor and the recent change in the implied volatility spread as the second sorting factor.

The *univariate* portfolio sorting has the following steps:

Step 1: Computing the breakpoints Starting from the cross-sectional distribution of the signal, $S_{i,t}$, from the number of portfolios I need, denoted n_ρ, I find for all the $n_\rho - 1$ breakpoints the k'th breakpoint as follows:

$$\mathcal{B}_{k,t} = \text{Percentile}_{\rho_k}\left(S_{i,t}\right) \tag{2.5}$$

The sample used for the breakpoints determination is the entire available dataset. This is done to ensure equal sized portfolios across the analysis. It might however affect the liquidity aspect of the conclusions, as inclusion of smaller and less liquid stocks might skew the picture. This is due to the fact that smaller and less liquid stocks often have higher transactions costs and tighter borrowing constraints, and therefore not priced precisely as the markets expect.

Step 2: Forming the Portfolios The portfolios are then formed using the breakpoint, $\mathcal{B}_{k,t}$, found in the previous step. A stock, i, is then included in portfolio, k, if it's signal is within the breakpoints identified:

$$P_{k,t} = \{i | \mathcal{B}_{k-1,t} \leq S_{i,t} < \mathcal{B}_{k,t}\} \tag{2.6}$$

When forming the last portfolio (the portfolio consisting of stocks with the highest value of the signal), I will add the last few stocks with a signal value equal to the maximum signal value of the sample. In contrast some researchers include less than or equal sign on both side of the signal in the formation of their portfolios, which means that some stocks might be included in two portfolios. To counteract this I choose the above approach insted, which results in my last portfolio having slightly more stocks included, but no stocks will be in two portfolios at once.

Step 3: Calculating the Returns The returns will be calculated based on a value-weighted investment. Thus they are reported in simple value-weighted form with the market value, $MV_{i,t-1}$ observed as the shares outstanding timed the opening price at portfolio formation, and calculated as follows:

$$r_{k,t} = \frac{\sum_{i=1}^{N_{k,t}} MV_{i,t-1} \cdot r_{i,t}}{\sum_{i=1}^{N_{k,t}} MV_{i,t-1}} \tag{2.7}$$

I choose a value-weighted approach compared to equal-weighted, as I want to make sure that small and illiquid stocks, which are difficult to trade, have a small impact on the results.

For a *bivariate* portfolio sorting, the steps look a little different. First and foremost, this will make it possible to control for two different sorting signals compared to just one in the univariate portfolio sorting. The inclusion of bivariate sorting to take a second signal into account leaves more choices to be made. Of course, there is the amount of portfolios and the percentiles of the second signal, but another important choice regarding the sorting of the stocks is also relevant, namely whether to do independent or dependent sorting. Which means that when we do the breakpoints calculation of the second signal, we should either find them for the entire sample or from the grouping from the first signal.

In my case, I will look at the recent change in implied volatility spread as the second signal. Given the combination of the value of the same dataseries being the first sorting signal, it makes sense to do a dependent sorting. If an independent sorting were to be made, and given that the signal is relatively stable over the period as seen in figure OBS, there would be very few stocks with a low value of the signal (thus sorted in the first portfolio in the first sorting) and also having a big positive change in the signal value (sorted into the last portfolio in the second sorting) and vice versa. An independent sorting would therefore result in an almost empty portfolio of stocks with high (low) value and big negative (positive) change in the signal. A dependent sorting is chosen, and the following steps from above is adapted.

Adapted Step 1: Breakpoints for Dependent Bivariate Sorting To find the breakpoints dependent on the first sorting signal, I use the following approach. Note that as mentioned above, the last portfolios are including the stocks with maximum value of the signal. The first sorting breakpoints for portfolio k are defined as above in Equation 2.5, and the second sorting signal's breakpoints for portfolio j is defined below:

$$\mathcal{B}^2_{k,j,t} = \text{Percentile}_{\rho_j}\left(S^2_{i,t} | \mathcal{B}^1_{k,t} \leq S^1_{i,t} < \mathcal{B}^1_{k,t}\right) \tag{2.8}$$

Which leaves me with $n_{p_1} - 1$ breakpoints for the first portfolio sorting and $n_{p_1} \cdot (n_{p_2} - 1)$ breakpoints for the second portfolio sorting.

Adapted Step 2: Forming Portfolios with Dependent Bivariate Sorting Instead of allocating stocks to different portfolios according to only the first signal, I use both set of breakpoints simultaneously.

$$P_{k,j,t} = \left\{i | \mathcal{B}^1_{k-1,j,t} \leq S_{i,t} < \mathcal{B}^1_{k,j,t}\right\} \cap \left\{i | \mathcal{B}^2_{k,j-1,t} \leq S_{i,t} < \mathcal{B}^2_{k,j,t}\right\} \tag{2.9}$$

This results in $n_{p_1} \cdot n_{p_2}$ portfolios with approximately an equal amount of stocks included in each, with only the last portfolios including a slightly larger amounts of stocks.

Adapted Step 3: Calculating Returns with Dependent Bivariate Sorting The third step is only slightly changed to account for the second dimension of portfolios:

$$r_{k,j,t} = \frac{\sum_{i=1}^{N_{k,j,t}} MV_{i,t-1} \cdot r_{i,t}}{\sum_{i=1}^{N_{k,j,t}} MV_{i,t-1}} \tag{2.10}$$

This concluding the adapted steps for the bivariate sorting. Any structure of analysis from here will look largely the same.

With the portfolio returns being computed, it is time to analyze them. A relevant fourth step will be to do some descriptive statistics of the returns and dive further into the pattern.

2.3 Descriptive Statistics

Sharpe Ratio leading on to Utility

2.4 Monotonicity Test

The pattern of the returns have traditionally been inspected through visual analysis of the mean return of each portfolio. To do this in a more structured way, Wolak (1987, 1989) introduce a test for monotinicity in returns with a null hypothesis of a clear pattern, and Fama (1984) proposes a similar test using Bonferroni bounds. Patton and Timmermann (2010) introduces a test with the same purpose, but a different null hypothesis of no clear pattern in the returns across portfolios. This ensures that we only reject the null hypothesis of no distinct monotonically increasing returns across portfolios if there is enough evidence to support it, while the abovementioned tests would require the data to show a random pattern instead of a monotonically increasing one. In other words, I prefer to use a test which has no pattern as a prior and relies on the data to prove it wrong.

The hypothesis of the Patton and Timmermann (2010) test is formally written as, using $\Delta_i = \mu_i - \mu_{i-1}$ as the difference between the average return of portfolio i and $i-1$:

$$\begin{aligned} H_0 &: \Delta \leq 0 \\ H_1 &: \Delta > 0 \end{aligned} \tag{2.11}$$

The test statistic is then computed as the minimum of the observed deltas, $J_T = \min(\Delta)$. As there is no preformed distribution of this test statistic on which we can evaluate significance, we will find the distribution through bootstrapping with replacement from the original sample. This is all done in the R package provided by Köstlmeier (2019).

The reason for the inclusion of this test is to identify if there is a monotonically increasing (or decreasing) pattern in the average returns of the portfolios. This will lead to a preliminary conclusion that it might be relevant to look at this signal, given that it when rejecting the null has a significant effect on the expected returns.

This test is, however, based on a view on the entire sample period and does not account for time varying monotonicity or if the pattern persists when corrected for other known risk factors. I will therefore in the further analysis look deeper into the time-varying predictability abilities of the signal and examine how it performs in the cross sectional study when confronted with common risk factors from Fama (1984).

2.5 Comparing with Other Factors

In this subsection I will elaborate on different approaches used for evaluating the signal against other (and acknowledged in the literature) factors. The first approach will be very basic and relate to the cross sectionality part of the analysis. The remaining approaches will relate to predictability of the

2.5.1 Fama-Macbeth Analysis

Models considered : time invariant + rolling window betas + simple model + extended models with FF5

Common issues -> Errors-in-variables + autocorrelation and heteroskedasticity -> using GMM to lessen issues + increasing the sample frequency to decrease uncertainty in measuring + using portfolios to cancel out noise + using Shanken's correction to the standard errors as to not accept some coefficient as being statistical different from zero.

2.6 Economic Evaluation

Why would it be useful to evaluate the portfolios, in not only a statistical way but also in an economic way? Looking at utility from investing in long-short portfolios seem obvious.

Utility from investing (expected) in the portfolio -> certainty equivalent -> mean-variance?

2.7 Optimal Allocation

Mean-variance? -> gaussian distribution (then mean-variance is the most relevant) -> finding the efficient frontier -> plotting the portfolios found along with FF 25 portfolios -> drawing the optimal allocation -> including the target return index

3 Data

Data is downloaded from OptionMetrics, CRSP and Compustat. The extensive data has been filtered to only include the stocks registered on American exchanges and are only included in the portfolio formation if a signal is observed in the relevant period, which is the preceding week in the return predictability section and in the week of the return in the cross sectional section. All data have been joined together using the SECIDs, PERMNOs and GKEY.

The conditions for excluding some of the data regarding the implied volatilities are introduced in Cremers and Weinbaum (2010). The conditions are elaborated below. Note, however, that as also discussed in Shang (2016), I will do the analysis for a dataset with the conditions imposed, but also include the results from the analysis without the conditions imposed in the appendix, to show the effect of the conditions and the robustness of the results.

The signal is the spread between the implied volatility of call and put options with the same maturity and strike. These values are observed at market close at a daily frequence. I will take an average across maturities and strikes everyday for each stock to have a simple signal. This average can both be a simple average across every available datapoint or a weighted average using the open interest reported at market close for each pair. Cremers and Weinbaum (2010) argue that the latter approach incorporates the liquidity aspect of the options. I assume that the more liquidity an option has, the more fair is the price, and the more reflective of the market's opinion is it. Thus the latter approach for weighting the implied volatility spread ensures that the signal incorporates the market's view.

The conditions for including the options in the signal is described below. The first condition is in regards to the implied volatility of each option.

Condition 3.1: Restrictions on Implied Volatility Level

$$0 \leq IV_{j,t}^{i,call} \leq 1.5 \text{ and } 0 \leq IV_{j,t}^{i,put} \leq 1.5$$

The implied volatility should be within these limits for both puts and calls for the spread to be included in the signal, which naturally limits the distribution of the implied volatility spread. The historical volatility of the stock markets represented by the VIX Index reached a max of 0.7 in 2017, therefore a limit of 1.5 on individual options is sensible.

The second condition relates to the time to maturity of the option pairs, measured in days.

Condition 3.2: Restrictions on Time To Maturity

$$7 \leq TTM_{days} \leq 365$$

Options with time to maturity within these limits are the most liquid and by excluding the imminent maturing options, the signal will only contain data on options with maturity prevailing the return period.

For the last condition, the forward price of the stock is estimated at the point of daily closing prices for the individual option. Optionmetrics describes the calculation as follows: "The forward security price is calculated based on the last closing security price, plus the interest, less projected dividends" - OptionMetrics (2015).

Condition 3.3: Restrictions on Moneyness

$$0.7 \leq \frac{F_{0,i}}{K} \leq 1.3$$

The moneyness (ratio between forward price of the underlying stock and the strike price of the option) should be between those limits, which ensures that the options included are somewhat close to being at-the-money and decreases the noise from illiquid deep-in-the-money or deep-out-of-the-money options.

The average of these is taken per day, and the observed signal for return predictability is the latest day within the last 7 days before the return period begins. Furthermore, for the cross sectional analysis, the implied volatility spread is averaged across the entirety of the return period, giving an average of at most 5 days observations. If there is no observed spread in this period, the stock is not considered for portfolio formation, in either scenario.

The data is splitted into portfolios based on the absolute value of the signal and on the change in the signal over the last week.

Why are we only interested in at-the-money options?

Meta data: amount of stocks over time, signal over time (the four plots of being weighted and filtered) + split the sample into different levels of open interest and tabulate the difference in mean / var etc. for 5 portfolios.

4 Cross Sectional Analysis

The portfolios ::: univariate on both variables -> which shows the most economic tendencies + bivariate sorting with both and arguments as to which sorting is best.

4.1 Descriptive Statistics

4.2 Analysis

4.3 Optimal Allocation

4.4 Conclusion

5 Discussion and Further Research

Choice of signal -> value -> evaluating the mispricing with a lower value equal to higher potential of losses absolute value -> evaluating mispricings and assuming both mispricing would lead to lower future return only the negative values -> only looking at the overpriced puts and if this is priced differently in the market squared mispricing ->

mean-variance optimizing investors -> does it make sense ? how have it been disproved what other masurements are there

References

D. T. Breeden and R. H. Litzenberger. Prices of state-contingent claims implicit in option prices. *Journal of business*, pages 621–651, 1978.

M. Cremers and D. Weinbaum. Deviations from put-call parity and stock return predictability. *Journal of Financial and Quantitative Analysis*, 45(2):335–367, 2010.

E. F. Fama. Term premiums in bond returns. *Journal of Financial economics*, 13(4):529–546, 1984.

C. R. Harvey and A. Siddique. Time-varying conditional skewness and the market risk premium. *Research in Banking and Finance*, 1(1):27–60, 2000.

J. Isakas and B. Melenberg. The option implied risk-neutral distribution.

S. Köstlmeier. *monotonicity: Test for Monotonicity in Expected Asset Returns, Sorted byPortfolios*, 2019. URL `https://CRAN.R-project.org/package=monotonicity`. R package version 1.3.1.

OptionMetrics. Ivydb us, file and data reference manual. Technical report, December 2015. Version 5.0.x.

A. J. Patton and A. Timmermann. Monotonicity in asset returns: New tests with applications to the term structure, the capm, and portfolio sorts. *Journal of Financial Economics*, 98(3):605–625, 2010.

D. Shang. *Option markets and stock return predictability.* PhD thesis, The University of Arizona, 2016.

F. A. Wolak. An exact test for multiple inequality and equality constraints in the linear regression model. *Journal of the American Statistical Association*, 82(399):782–793, 1987.

F. A. Wolak. Testing inequality constraints in linear econometric models. *Journal of econometrics*, 41(2): 205–235, 1989.

Appendices

List of Tables

List of Figures

A Figures

B Tables