The Cross-Section of Stocks and Options - Informed Trading & Volatility Mispricing

Philip Hamna[♠] & Caroline Holtsjo[♣]

May 18, 2015

Bachelor Thesis
Department of Finance
Stockholm School of Economics

Abstract

We examine the cross-section of stocks and options to confirm recent findings that past abnormal stock return predicts an increase in implied volatility and a decrease in realized volatility. In our extended empirical analysis we present new evidence indicating that future implied volatility is positively related to decreases in stock trading activity and negatively related to stock illiquidity. We also find indications that abnormal stock returns are positively and significantly related to informed trading. These patterns are consistent with rational models of informed trading and the results are robust under several regression specifications. Based on these findings, we present a zero-cost delta-hedge strategy that is short (long) volatility on stocks with high (low) past abnormal return and a decrease (increase) in past trading activity. This strategy generates monthly alphas of 98.2 bp. There are no indications that these alphas steam from neither naked positions in options nor in stocks, indicating a presence of volatility mispricing.

Keywords: Informed trading, Implied volatility, Realized volatility, Abnormal return, Cross-section

Thesis Advisor: Jungsuk Han[1]

♠ 22872@student.hhs.se

♣22902@student.hhs.se

[1] We are grateful for the rewarding help and support in writing this thesis.

Table of Contents

1 Introduction

In a complete and perfect capital market, where there are no transaction costs, no information asymmetry and continuous trading, options would be pointless financial instruments. As none of these assumptions holds true in the world as we see it today, option prices will, more often than not, deviate from the theoretical value of the underlying stock in combination with a risk-free asset [7]. This is why the informational role and price discovery process of options have been subject to academic research in finance for decades. A recent study by An et al. (2014), hereafter AABC, introduce a noisy rational expectation model of informed trading in both stock- and option markets. By examining the joint cross-sectional relationship between stocks and options, AABC finds evidence that informed trading simultaneously moves prices in both option- and stock markets. This implies that both options and stocks carry information with respect to future price movements. These results provides new evidence to the field of research concerning informed trading as it proves that informed traders use both options and stocks in their trading to take advantage of their private information.

Our empirical study builds on this noisy rational expectation model of informed trading. Based on AABCs methodology we find that CAPM alpha predicts a future increase in implied volatility by 3.17% and predicts a future decrease in realized volatility by 10.88%. In this cross-section of stocks and options, we find that the impact on future implied volatility increases by 15.77% as we convert CAPM alphas to Fama-French-Carhart adjusted alphas [8]. Furthermore, we find these Fama-French-Carhart adjusted alphas to be positively and significant related to the probability of informed trading, indeed indicating that there is a link to informed trading in abnormal stock returns. The intuition behind these predictive patterns is that prices adjust partly as informed traders trade simultaneously in both markets. Due to this, some future uncertainty is resolved, but at the same time, it creates a noise trading demand shock.

Additionally, we incorporate extensions both in terms of stock market liquidity and trading activity to distinguish on when informed traders use the stock market to exploit their superior information. These extensions are on one hand based on research indicating that informed investors have the ability to detect and circumvent price inefficiencies such as stock illiquidity [20][27]. On the other hand, the extensions are based on research presenting evidence that the probability of informed trading is closely connected to the trading activity in the stock market [14]. Supported by these findings we assert that the fraction of informed trading in the stock and thus, the demand shock in the option, is; increasing in falling stock trading activity and; decreasing in stock illiquidity.

In our empirical tests, we estimate firm level cross-sectional regressions on a monthly basis with lagged variables to predict the impact on implied volatility the following month. These lagged variables are Quoted Spread, Amihud, and percental change in stock trading activity. Using an extensive sample of large US stocks during a time period stretching from April 2005 to December 2014, we find empirical support for our extended predictions. Particularly, as trading activity is included into the original framework, we find a more pronounced impact on future implied volatility. On the other hand, as the stock becomes more illiquid, we observe a less pronounced impact on implied volatility. These results confirms the underlying theories and they are robust under several regression specifications, and also under the inclusion of year fixed effects. Furthermore, we find these results to be both of statistical and economic significance as we derive two self-financed trading strategies, which seeks to extract returns in the spread of implied- and realized volatility.

We construct a delta-hedge strategy that is long stocks and short the equivalent call options in decile portfolio 10, and short stocks and long the corresponding call options in decile portfolio 1. As we sort these decile portfolios based on past month Fama-French-Carhart alphas and change in stock trading activity, the monthly four-factor alpha is 98.2 bp. Also, during the entire sample period, the cumulative abnormal returns are 46.75 percentage points higher compared to when we sort stocks solely on past month's Fama-French-Carhart alphas.

To our knowledge, there is a limited field of research investigating the predictive relation of lagged stock return in this context. Furthermore, we are not aware of any studies incorporating trading activity in relation to stock alphas to estimate future impact on implied volatility. Nevertheless, the results presented in this paper have three main implications. First, our results supports the notion that informed traders use both option- and stock markets in their trading. Second, our findings are in line with the results indicating that informed investors are able to detect and circumvent stock market inefficiencies such as illiquidity. Third, we present new evidence suggesting a significant link between informed trading, trading activity and future implied volatility, something that we have not found in the academic research as of today.

The remaining parts of this paper are organized as follows. In section 2, we present previous research related to our study. Section 3 presents the underlying model, our extension and the main predictions which steam from this extension. Section 4 presents the data used in the empirical test and section 5, we define the key measures. Section 6 presents the methodology and in section 7, we describe and deliberate on our empirical findings and the self-financed strategies. Finally, in section 8 we conclude the main implications of our results and suggest starting points for future research.

2 Previous Literature

Up until today, there have been substantial research focusing on the price discovery process in the cross-section of stocks and options. This research is often connected to informed trading as option values frequently deviate from their theoretical values. The findings are rather conflicting and there are researchers supporting either the idea of the option markets leading the stock markets or on the other hand, stock markets leading option markets. Also, there are extensive research investigating the issue of which market the informed traders prefer to use to exploit their private information. In this section, we present previous literature related to this empirical study.

Easley et al. (1998) examines whether option volume contain information about future stock prices using intraday option data. By deriving an asymmetric information model where informed traders are allowed to trade in either stocks or options, they find significant evidence for their hypothesis that option volume contain valuable information with respect to future stock returns. Pan & Poteshman (2006) also find evidence that option volume contain information about future stock prices. They construct put-call volume ratios (P/C) to find that stocks with low P/C ratio outperforms stocks with high P/C ratio. Another volume approach is taken by Johnson & So (2012). They create ratios of option volume in relation to stock volume (O/S) and find a negative correlation between O/S and future firm value. They find evidence that this is due to high short-sale costs in the stock market. Consequently, as informed investors have negative information, they will prefer to trade in option markets to circumvent the high transaction costs in the stock market.

Investigating the price discovery process, Chakravarty et al. (2004) finds that both stock and option markets contribute to the price discovery process. They estimate that 17 percent of the stock price discovery can be attributable to option markets, and that this mostly is linked to leverage, spread and volume in the option markets. They conclude that this is in line with theoretical models suggesting that informed traders use both option- and stock markets in their trading. On the contrary, Muravyev et al. (2013) uses the put-call-parity to examine how price quotes adjusts in both option- and stock markets. They find that the option-implied price quotes adjusts in accordance with the actual stock price quote. Consequently, they find no significant price discovery in the option markets and thus, indicating that price discovery only occurs in the stock markets.

Concerning the cross-section of index performance and index options, Amin et al. (2004), investigates option prices in connection to the performance of the underlying asset. They suggest that past index returns exert an important influence on index option prices. They present evidence that option prices are influenced by the momentum effect of the underlying asset. Indicating that call (put) prices would increase as a result of increases (decreases) in the underlying asset. In line with research suggesting that informed traders use both markets, An et al. (2014)[2], derives a noisy rational expectation model where informed traders are allowed to trade in both option- and stock markets simultaneously. They show that the cross-section of past abnormal stock return predicts option implied volatilities; stocks with high abnormal return in the past month tend to give rise to increases in implied volatility over the next month. Contrariwise, Goyal & Saretto (2009) show that low past stock returns predicts an increase in implied volatility over the next month. However, they find evidence that this is due to investor over- and under reaction to the current month's stock raw returns and consequently, suggest a behavioural explanation for this predictive pattern. Also examining return predictability, Cremers & Weinbaum (2010) investigate stock returns from violations of the put-call parity. They find that stocks with relatively expensive call options, estimated from the put-call-parity, outperforms those stocks with cheaper call options.

Another important and related field of research is the subject of which market informed traders prefers to capitalize on their private information. As with the lead-lag relationship, there are different findings. On one hand, there is Black (1975) arguing that informed traders may be attracted to the option market due to the higher leverage available in option trading. He claims that the leverage generates a higher yield on the private information held by informed traders. On the other hand there is Kyle (1985), arguing that informed traders seek to hide their trading activity. Therefore, to disguise their trading, informed traders prefer the more liquid stock market. Similarly, Admati & Pfleiderer (1988) find that informed investors prefer to trade when their trading has little effect on prices, and therefore avoids the most illiquid market.

3 Model of Informed Trading

In this section we start by introducing the intuition behind the model derived by AABC and the main implication for our study. After which, we describe the underlying intuition of our extension to the original model, where we strive to incorporate trading activity, private information and market efficiency.

3.1 The Original Model

In the noisy rational expectation model, the economy consists of risk averse informed traders and noise traders and one risk averse market maker. The informed traders and the market maker are identical apart from the fact that the informed traders receive a signal about the true value of the firm's cash flow. This signal, denoted $\emptyset$, is received before the true value of the cash flow is released. The extent of the informed trading

[2] An et al. (2014) is referred to as AABC in the rest of this study

activity is partly dependent on the level of private information, and partly on the noise trading activity. This implies that the underlying model does not impose restriction on where the informed trading occurs. Consequently, the model allows for informed trading in both stock- and option markets simultaneously. However, the model assumes that noise traders cannot trade across these markets.

The cash flows are released in the end of period $t = 2$ and trading occurs in time $t = 1$. Additionally, the firm is assumed to be born in time $t = 0$. If no signal is received, the informed trader buys one half of a stock and the market maker buys the other half, and no trading shocks is assumed to take place in neither of the two market. Given this setting and the implications of the model, AABC have derived the following relationships:

$$\begin{aligned} Cov(F - S_1, C_1\,) &= Cov(F, S_1) - Var(S_1) \\ &= E[\big(F - E(F) - S_1 + E(S_1)\big) * \big(S_1 - E(S_1)\big)] \end{aligned} \tag{1}$$

$$Cov(C_2 - C_1, S_1) = E[\big(C_2 - E(C_2) - C_1 + E(C_1)\big) * \big(S_1 - E(S_1)\big)] \tag{2}$$

where F is the firm cash flows, C_1 , C_2 and S_1 is call option and stock prices in time $t = 1$ and $t = 2$. These two relationships show a joint cross sectional relationship - between options and stocks and between stocks and options. A third relationship is connected to expression (1) and links the implications of the resolved uncertainty that follows if $Cov(F - S_1, C_1) > 0$. Intuitively, if the call option predicts future stock prices, the uncertainty the subsequent period will decrease and thus, the variance of the stock returns decrease. The expression that follows is:

$$Var(F - S_1) < \ Var(F - S_0) \tag{3}$$

If we treat the stock at time $t = 0$ as a constant we obtain:

$$2Cov(F, S_1) - Var(S_1) = Cov(F, S_1) + (Cov(F - S_1, S_1) > \ 0) \tag{4}$$

where we find the expression that reveal the relationship from expression (1). The implications of the model is that prices adjusts partly towards the value indicated by the informed traders signal. Intuitively, as informed trading occurs in both markets, both option- and stock prices incorporates new information simultaneously. This causes a slow price adjustment towards the new value indicated by the informed traders signal. As the noise traders observe this adjustment, their demand will increase and this causes the profit maximizing informed traders to increase their holdings while the price still is below the value indicated by their signal. Putting it differently, due to the slow diffusion of information into stock- and option prices, the model predict that a demand shock will take place in the end of $t = 1$.[3]

3.2 Additions to the Original Model

In our extension to the model, we seek to highlight the predictive relationship between past stock performance, trading activity and the demand shock. Consequently, we aim to improve the proxy signal of informed traders in the stock market. There is evidence that the probability of informed trading in the stock market is closely connected to the trading activity. Namely, that the probability of informed trading is decreasing in an increased stock volume [14]. The hypothesis is that the demand shock will be more pronounced if the fraction of informed trades increases thus, if the trading activity during the preceding period was relatively low. We use the following expression:

$$\chi = \omega + (\alpha * Volume\ factor) \tag{5}$$

where χ is the total fraction of informed trades, ω describes the fraction of informed trades in the stock if the volume factor is zero. Putting it differently, if the volume is unchanged between month $t-1$ and $t-2$ (we define the volume factor in section 5.4). α is a negative coefficient that describes the relationship between the fraction of informed trades and the volume change between the two preceding months. Notice that the negative coefficient in expression (5) implies that the relationship holds for illiquidity factors as well. This follows from research indicating that informed traders are able to detect and circumvent stock inefficiencies such as stock illiquidity [20][27].

3.2.1 Empirical Predictions

Given the implications of the original model; that informed traders are allowed to trade in both markets simultaneously; that uninformed investors are prohibited to trade across markets, and that we set $\alpha < 0$, we obtain:

i. Abnormal stock return will predict an increase in implied volatility and a decrease in realized volatility (Prediction from Original Model)
ii. Abnormal stock return in combination with a decreased volume will predict a higher fraction of informed trading in the stock and thus, a more pronounced demand shock (Main Empirical Prediction)
iii. Illiquidity in the stock will predict a decreased fraction of informed trading in the stock and thus, a less pronounced demand shock (Robustness Prediction)

[3] For further mathematical and technical aspects of the Original Model, see AABC Appendix A

3.3 Intuition and Derivation of Main Empirical Prediction

Conferring to the setting of the original model, the traded volume of any stock can be seen as a function of informed trading and noise trading. We define volume in any stock and any period as:

$$V_i = (p\ (\emptyset)\ * V_i^{Informed}) + (z * V_i^{Noise}) + \varepsilon_i^j \tag{6}$$

where V_i is the total traded volume, $V_i^{Informed}$and V_i^{Noise} are the informed- and noise traders volume fractions of the total volume and p symbolizes the probability of the informed traders receiving a signal. Further, $\emptyset$ represent the informed traders signal, z is the noise traders' stock demand and ε_i^j is an error term. According to the underlying model, we set $p\ (\emptyset)\ = 1$, implying that there is no uncertainty in the information received by the informed trader. In other words, the true value of the cash flows in the next period is completely known by the informed traders. Once again, according to the underlying model, we assume the noise traders' stock demand, z, to follow an independent normally distributed function. Accordingly, the right hand side can be simplified:

$$p\ (\emptyset)\ * V_i^{Informed} = V_i^{Informed} \tag{7}$$

$$z * V_i^{Noise} = V_i^{Noise} * z \sim N(0, \sigma_z^2) \tag{8}$$

The underlying hypothesis is that the fraction of informed trades increase, implying a more pronounced demand shock, if; the trading activity has decreased during the previous period, and if; it is accompanied by a signal which from the underlying model can be assumed to be abnormal stock return (alpha). Accordingly, we set:

$$V_{t-1} < V_{t-2} \tag{9}$$

This implies that the traded volume during month $t-1$ is less than the traded volume during month $t-2$ and that this volume change is accompanied by an informed signal that is of highest possible quality, $p\ (\emptyset)\ = 1$. The fact that the informed traders know the value of the cash flows in the next period drives the informed trading activity. But as the underlying model predicts, we cannot be confident about where the informed traders choose to trade. However, as the abnormal return grows, it can be seen as a proxy for informed trading in the stock market since the informed traders drives prices towards the value given by the informed signal. By this, we obtain the following expressions:

$$V_{t-1} = \left(V_{t-1}^{Informed}\right) + \left(V_{t-1}^{Noise} * z \sim N(0, \sigma_z^2)\right) + \varepsilon_{t-1}^{i} \tag{10}$$

$$V_{t-1} - \left(V_{t-1}^{Informed}\right) = \left(V_{t-1}^{Noise} * z \sim N(0, \sigma_z^2)\right) + \varepsilon_{t-1}^{i} \tag{11}$$

$$V_{t-1} \geq \left(V_{t-1}^{Informed}\right) > \left(V_{t-1}^{Noise} * z \sim N(0, \sigma_z^2)\right) + \varepsilon_{t-1}^{i} \tag{12}$$

It should be emphasized that this derivation requires two empirical test. Both in terms of the hypothesis that alpha is connected to informed trading and in terms of trading activity in connection to informed trading. This steams from the fact that any stock movements can be caused by irrational trading behavior of the noise traders. Due to this, the abnormal return cannot implicitly be assumed to be connected to informed trading.

4 Data

Our daily data on option implied volatilities are from Bloomberg L.P. The implied volatility data from Bloomberg is equating the Black-Scholes option pricing model to European style options [7]. Furthermore, Bloomberg excludes options that mature within 10 days of computation of implied volatility[4]. One key advantage of using an interpolated data set of implied volatilities is that we do not have to make any subjective assessments on which options to include and exclude. As AABC, we look at ATM options that are 30 days from expiration and our sample data covers a time period stretching from April 2005 through December 2014. Due to data availability issues we look at implied volatility data on European style call options on the stocks included in the Standard & Poor's 500 index as of March 23, 2015.

The stock data is collected from Center for Research in Security Prices (CRSP), which covers all stocks listed on NYSE, AMEX and NASDAQ. The daily CRSP data covers daily close prices, closing bid- and ask prices, number of shares outstanding, number of shares traded and individual tickers. The data also contains daily stock returns, adjusted for stock splits and dividend payments where any dividend payments are assumed to be reinvested on the ex-distribution date. We also collect Fama-French-Carhart factors from Kenneth French's website at Dartmouth and CBOE VIX data from WRDS. We also collect quarterly Probability of Informed Trading (PIN) data from Stephen Brown's website at Robert H. Smith. The PIN data is computed using the model by Venter and De Jongh (2006), which is an extension from the original findings by Easley et al. (1996). Due to data availability issues, the PIN data only

[4] Further description on the technical aspects of the interpolations methods used in the data set on implied volatilities from Bloomberg L.P can be found at: https://msb040.msb.edu/faculty/bodurthj/unrestricted/teaching/programs/Bloomberg˙Implied˙Volatility˙Method˙2008˙new.pdf (2015-04-20, 09.47)

covers a sample period stretching from April 2005 to December 2010. The data contains the crude PIN computation, probabilities of information based events and informed and uninformed trading intensity.

Our final data set, which combines the Bloomberg data, CRSP data and the Fama-French-Carhart factors, includes a total amount of 901.142 daily observations, split over 117 months. The average number of unique tickers in any given month amounts to 376 and the total number of unique tickers is 497, implying that three firms have been eliminated due to missing implied volatility data.

5 Variable Description

In this section we describe the definition of the key variables used in the empirical test. We start by defining the variables used to confirm the relationship derived in AABC after which, we define the variables used in our extended empirical analysis.

5.1 Realized Volatility Innovation

To compute the realized volatility of each stock we follow the most commonly used method - the standard deviation of daily returns over the past trading month. The equation is:

$$Realized\ Volatility_{n,j}^{i} = \sqrt{\left(\frac{252}{n} * \sum_{i=j}^{n} r_j^2\right)} \tag{13}$$

where r_j^2 is the squared daily return and n is the number of days in the estimation window. To match the realized volatility with the implied volatility, we set the estimation window to 30 days. Thus, for each given day, we have an annualized realized volatility that can be compared to the annualized implied volatility. The final variable is defined as:

$$\Delta RVOL_t^i = Realized\ Volatility_{j+n}^{i} - Realized\ Volatility_j^i \tag{14}$$

where we compute the first difference between the realized volatility on the last day, $j + n$, and the first day j, of month t.

5.2 Implied Volatility Innovation

We define the implied volatility variable in accordance with AABC, which simply is the first difference between the implied volatility on the last day, $j + n$, and the first day, j, of month t. Thus, the definition of the implied volatility variable is:

$$\Delta IV_t^i = \ Implied\ Volatility_{j+n}^i - Implied\ Volatility_j^i \tag{15}$$

As AABC points out, it is important to remember that this definition ignores the fact that implied volatilities are predicable both in the time-series and in the cross-section in terms of autocorrelation over time and stock characteristics. Nevertheless, as we strive to confirm and use the original relationship, we do not alter this definition of the implied volatility innovation.

5.3 Alpha

Alpha is the intercept in the regression of the excess stock return, r_j^i, on various systematic factor loadings. Following the method of AABC, we use the CAPM alpha (16). In addition, we use the Fama-French-Carhart model (17) (hereafter four-factor model) to compute alpha [8]. Each of the models are used for each stock at each day, j, in month $t-1$[5]. Alphas are calculated using the following expressions:

$$\alpha_j^i = r_j^i - \ [\beta_1(r_j^{mkt} - r_j^f) + \varepsilon_j^i] \tag{16}$$

$$\alpha_j^i = r_j^i - \ [\beta_1(r_j^{mkt} - r_j^f) + \ \beta_2 SMB_j + \beta_3 HML_j + \beta_4 UMD_j \ + \varepsilon_j^i] \tag{17}$$

where $(r_j^{mkt} - r_j^f)$, is the excess return of the market portfolio, SMB_j is the small firm-minus- big firm factor, HML_j is the high book-to-market minus the low book-to-market factor, UMD_j is the momentum factor and ε_j^i is an error term. The logic behind our inclusion of four-factor alphas steam from research of Fama & French (1993) and Jegadeesh & Titman (1993). As they find evidence that SMB_j, HML_j and UMD_j are systematic risk factors in the stock market, we include these factors in our empirical analysis. As a result, investors should not be able to earn abnormal returns that steam from exposure to these systematic risks. Consequently, the four-factor alpha should give a more accurate estimate due to the elimination of non-diversifiable risks. Further, we define the alpha variable as:

$$Alpha_{t-1}^i = \sum_{i=j}^{j+n} \alpha_j^i + .. + \alpha_{j+n-1}^i + \alpha_{j+n}^i \tag{18}$$

where α_j^i is the first day alpha of month $t-1$ and α_{j+n}^i is the alpha of the last day in month $t-1$. One advantage of using the cumulative sum of the daily estimates rather

[5] As AABC points out, the effect of nonsynchronous beta estimation has a small impact in this type of calculation. They perform both a regular beta computation and nonsynchronous adjustments, and the difference is small. According to this, we use the standard approach.

than monthly estimates is that we include more data points and thus, possibly, capture the trading behavior in a more realistic way.

5.4 Volume Factor

To capture the trading activity in the stock market, we compare the traded share volume between month $t-1$ and month $t-2$. We define the Volume factor as:

$$Volume\ factor_{t-1}^{i} = \left(\frac{\overline{Volume}_{t-1}^{i} - \overline{Volume}_{t-2}^{i}}{\overline{Volume}_{t-2}^{i}}\right) \tag{19}$$

where $\overline{Volume}_{t-1}^{i}$ is the average traded volume for stock i during month $t-1$ and $\overline{Volume}_{t-2}^{i}$ is the average traded volume for the same stock during month $t-2$. Thus, this variable captures the change in traded volume between the two previous months. Chordia et al. (2001) categorizes volume as a measure for trading activity and since we define our measure as the difference between months, we are able to capture any potential change in trading activity and aggregate trading behavior in the stock. As we seek to measure the demand shock that arises in month t, this measure captures one important aspect of the activity in the stock in the month before the estimation window.

5.5 Probability of Informed Trading

We use the Probability of Informed Trading (PIN) measure to test the proposed relationship between alpha and informed trading. Fundamentally, PIN is a measure capturing any information asymmetry in the financial markets. It is based on a sequential trade model and the measure itself is dependent on number of sell- and buy initiated orders and the probability of an information based event in a specified event window. According to Easley et al. (1996) and further extension by Venter and De Jongh (2006), we use the following definition of PIN:

$$PIN = \frac{\vartheta * \rho}{(\vartheta * \rho + (2 * \tau))} \tag{20}$$

where ϑ is the trading intensity (number of trades per day) derived from informed traders and τ is the same measure derived from the uninformed traders. ρ is the probability of any information based event on any day in the estimation window. This probability measure should not be confused with the probability used in section 3.3. The PIN measure is computed quarterly for each stock and it enables us to test the relationship between informed trading and alpha from a more theoretical and technical perspective.

5.6 Liquidity Measures

Continuing in our extended empirical analysis, we include two variables that aims to capture stock liquidity. The reason behind this inclusion is that academic researchers have found that informed traders have the ability to exploit fundamental price deviations, such as inefficiencies in terms of stock illiquidity [20][27]. We use the Quoted Spread and Amihud as proxies for stock liquidity.

5.6.1 Quoted Spread

The first liquidity measure is the Quoted Spread. Quoted Spread is a common measure for liquidity and as it captures transaction costs and thus, can be seen as a proxy for market tightness. Trading costs in terms of order processing, information asymmetry, inventory and market structure are assumed to be reflected in the Quoted Spread. It is defined as:

$$Quoted\ Spread_{t-1}^{i} = \frac{1}{n}\sum_{j=1}^{n}\frac{(Ask_j^i - Bid_j^i)}{Ask_j^i} \tag{21}$$

where Ask_j^i is the closing ask price for stock i during day j and Bid_j^i is the closing bid price for the same stock, during the same trading day. The sum of all daily Quoted Spreads is divided by n, which is the number of trading days in month $t-1$. Thus, we define the $Quoted\ Spread_{t-1}^{i}$ as the average Quoted Spread for each firm during month $t-1$. This measure is attractive since it is easy to compute and it takes both implicit and explicit trading costs into account [28].

5.6.2 Amihud

The second liquidity measure in our empirical tests is Amihud, which also is a widely used liquidity proxy, developed by Amihud (2002). Compared to the Quoted Spread, Amihud more effectively measures the price impact in combination to stock liquidity, which makes it a more appropriate proxy for market breadth [10]. The price impact in this setting refers to the decrease of stock price after a sell-initiated trade order or the price increase after a buy-initiated trade order [5]. The Amihud measure is defined as:

$$Amihud_{t-1}^{i} = \frac{1}{n}\sum_{j=1}^{n}\frac{1.000.000 * \left|r_j^i\right|}{Closing\ Price_j^i * Volume_j^i} \tag{22}$$

where r_j^i is the daily return for stock i and $Closing\ Price_j^i * Volume_j^i$ is the dollar traded volume for the same stock during the same day. By summing the daily Amihud measures during month $t-1$ and dividing with n, which is the number of trading days in month $t-1$. Consequently, we obtain an average Amihud measure for each stock during month $t-1$. Amihud (2002) show that less liquid stocks on average carry a higher illiquidity premium and thus, a positive relationship between returns and illiquidity. Putting it differently, a high return per dollar traded volume indicates an illiquid stock.

5.7 Model Complexity and Measurement Issues

The underlying model used in this study imposes a dimension of complexity and elusiveness. This due to the fact that the model allows for informed trading in both option- and stock market simultaneously, and that the determinant on where the informed trading takes place is the noise trading activity. As the model does not impose any restrictions on the informed traders trading activity, one can argue that this model approaches the reality in a more representative way compared to models derived by other academic researchers, such as Easley et al. (1998), Pan & Poteshman (2006), Johnson & So (2012).

When it comes to our liquidity measures and our proxy for trading activity, it is important to highlight that no single measure will capture all the aspect of stock liquidity and activity. It is also important to highlight that Quoted Spread has been used as a signal for information asymmetry and thus, can be seen as a proxy for informed trading on a stand-alone basis [12]. However, as described in section 5.6.1, the Quoted Spread is a tempting measure to use since it captures both implicit and explicit trading costs and it is easy to compute. One possible drawback with this measure is that it might, systematically, underestimate the liquidity of the stock since many trades during one trading day occur within the Quoted Spread [21]. Also, since we only use closing prices, we miss out on any intraday high and low values, which also could give a biased perspective on stock liquidity. Amihud might be more effective as it captures price impact in connection to market breadth. However, Goyenko et al. (2009) finds that Amihud might give a more accurate measure over longer time periods and also, that the correlation between Amihud and price impact only amounts to roughly 50%. Finally, as our activity measure is compelling in an intuitive way, any changes between months might arise due to seasonal patterns rather than fundamental changes in trading activity [11].

6 Methodology

In this section we describe the framework used in the empirical tests. We start by identifying the regression specifications used to confirm the underlying relationship from AABC. Second, we specify the main regression used in our primary test, where we incorporate our volume factor into the original framework. Third, we explain our three robustness test specifications and at last, we introduce the framework for our trading strategy to confirm any potential economic significance of our findings. All regression specification from (23) to (37) are firm level cross-sectional regressions.

6.1 Confirmation of Predictive Relationship

A natural start is to introduce the regression specifications used to confirm the underlying relationships derived by AABC. It serves as a good rhetorical benchmark for further analysis and also, as an introduction to the framework used in this paper. We define the regression specification in the following way:

$$\Delta IV_t^i = \lambda_{0_{t-1}}^i + \lambda_1^i Alpha_{t-1}^i + \varepsilon_{t-1}^i \tag{23}$$

where ΔIV_t^i is the implied volatility innovation defined in equation (15) and $\lambda_{0_{t-1}}^i$is the intercept in the regression. $Alpha_{t-1}^i$ is the CAPM-estimate of alpha on each stock during the previous month, $t-1$, defined in equation (16) and (18), and ε_{t-1}^i is an error term. According to the underlying model of informed trading, we expect the regression coefficient to be positive and significant. This since we anticipate a demand shock to occur in month t, causing the call option prices to increase and thus, causing the implied volatility measure (ΔIV_t^i) to increase. The other underlying relationship used in this paper is how past stock performance influence future realized volatility, we specify the regression as:

$$\Delta RVOL_t^i = \lambda_{0_{t-1}}^i + \lambda_1^i Alpha_{t-1}^i + \varepsilon_{t-1}^i \tag{24}$$

where $\Delta RVOL_t^i$ is the change in realized volatility during month t, defined in (14) using the same methodology as in the implied volatility innovation. All other variables are defined as in (23). Here we expect a negative and significant regression coefficient. The intuition is that if abnormal returns grows during month $t-1$, uncertainty with respect to future variance in return and cash flow surprises is resolved and thus, the coefficient should be negative. This is also based on the implications of underlying model.

6.2 Fama-French-Carhart Extension

Based on AABC indications that alpha can be seen as a proxy for informed trading, we test the four-factor adjusted alpha as a replacement for of the CAPM estimate. The regression specification is defined as in (23):

$$\Delta IV_t^i = \lambda_{0_{t-1}}^i + \lambda_1^i Alpha_{t-1}^i + \varepsilon_{t-1}^i \tag{25}$$

apart from $Alpha_{t-1}^{i}$ which now is defined as in (17) and (18). We still expect the regression coefficient to be positive and significant. If the hypothesis that alpha can be seen as a proxy for informed trading is valid, we should expect the coefficient to increase and still be significant. Intuitively, as we adjust the alpha estimate by eliminating additional systematic risk factors, we create a more sophisticated measure of abnormal returns. By doing so, we only preserve aspects of alpha that hypothetically could originate from informed trading. If this is true, four-factor adjusted alpha should be a more refined measure of informed trading and by this causing a more pronounced demand shock, hence the expectations of an increased coefficient.

6.3 Main Test

In our main test, we investigate the relationship between past stock performance, trading activity and their impact on implied volatility. We use the volume measure defined in expression (19) as a proxy for stock trading activity. We base this on the hypothesis and empirical predictions in section 3.2.1. The main regression specification is defined as:

$$\Delta IV_t^i = \lambda_{0_{t-1}}^i + \lambda_1^i Alpha_{t-1}^i + \lambda_2^i Volume_{t-1}^i + \varepsilon_{t-1}^i \tag{26}$$

where all variables are defined as in (25) apart from $Volume_{t-1}^i$, defined in (19). Since we seek to measure the impact of the demand shock in time t, we expect the regression coefficient to be negative and significant. Intuitively, if a negative volume change is combined with a negative coefficient, the impact on ΔIV_t^i will be positive. The underlying logic is that if the past volume has been relatively low and thus, the trading activity has decreased, the impact of the demand shock in time t will be more pronounced.

As seen in our derivation of the model extension, in cases were alpha is combined with a decreased volume, we predict a higher fraction of informed trades. This prediction is in line with results by Easley et al. (1996), which finds evidence that the probability of informed trading increases in a decreased stock trading volume. This might seem contradictive since it is widely argued that informed traders seek to hide their trade in the stock market if its liquidity or volume is high [24]. However, as our variable is defined, a negative change does not implicitly suggest that the volume in the stock is low and thus, that the stock is illiquid. It is rather a measure of how trading activity changes between months. In the setting of the model, this implies that as noise traders observe the proxy signal of informed trading, they will increase their activity. Since the trading activity was relatively low during the previous period, this increase in demand will be more pronounced. As noise trading activity increases, the informed traders will become more aggressive (as described in section 3.2), thus a more positive impact on future implied volatility. To simplify the interpretation of this hypothesis in the regression, we include a dummy which is defined as:

$$Dummy_{t-1}^{i} = \begin{cases} 1 & if\ Volume\ Decile = 1 \\ 0 & otherwise \end{cases} \quad (27)$$

Consequently, the regression specification is:

$$\Delta IV_t^i = \lambda_{0_{t-1}}^i + \lambda_1^i Alpha_{t-1}^i + \lambda_2^i Volume_{t-1}^i + \lambda_3^i Dummy_{t-1}^i + \varepsilon_{t-1}^i \quad (28)$$

where the dummy, defined in (27), captures observations where the volume for stock i during month $t-1$ has the largest negative change. All other variables are defined as in (26). We expect the coefficient of the dummy to be positive and significant. The logic is once again that if the volume has decreased, the demand shock should be more pronounced and thus predict a positive change on ΔIV_t^i. To further test the main regression model, we include a time fixed effect according to:

$$\Delta IV_t^i = \lambda_{0_{t-1}}^i + \lambda_1^i Alpha_{t-1}^i + \lambda_2^i Volume_{t-1}^i + \lambda_3^i Dummy_{t-1}^i + Y_T + \varepsilon_{t-1}^i \quad (29)$$

Here we test the model in a setting where we eliminate fixed effects over time. Y_T is a dummy which takes unique values each year and all other variables are defined as in (28). By capturing these fixed effect over time, we test the statistical validity of the model and thus, make more correct assessments with respect to systematic shocks. If our model holds and the predictive relationship between alpha, trading activity and their impact on ΔIV_t^i persist over time, we should not expect any major changes in the coefficients from specification (28).

6.4 Robustness Tests

Since the model allows informed traders to trade in both markets, our robustness tests aims to distinguish on where the informed traders place their trades to exploit their private information. Firstly, we use PIN as a proxy to test the relationship between alpha and informed trading. Secondly, we use the Quoted Spread and Amihud as proxies for stock liquidity and stock market efficiency.

6.4.1 Probability of Informed Trading

As previously mentioned, AABC suggests that alpha can be seen as a proxy signal for informed trading. Therefore, to test the relationship between alpha and informed trading, we use the Probability of Informed trading (PIN) measure. The measure is defined in expression (20) and the regression specifications are as follows:

$$PIN_t^i = \lambda_{0_t}^i + \lambda_1 Any\ Event_t^i + \varepsilon_t^i \tag{30}$$

where $Any\ Event_t^i$ is the probability of any information based event on any given date in the estimation period. The estimation period stretches from 10 to 70 days for any stock in the sample. Based on the underlying intuition of the model, suggesting that informed traders take positions in accordance with private signals, we expect the coefficient to be positive and significant. Furthermore, we include an extension defined as:

$$PIN_t^i = \lambda_{0_t}^i + \lambda_1 Any\ Event_t^i + \lambda_2 Bad\ Event_t^i + \varepsilon_t^i \tag{31}$$

where all variables are defined as in (30) exempt from $Bad\ Event_t^i$ which is the probability of $Any\ Event_t^i$ being negative on the event date. This variable has a more instinctive intuition. As the informed traders receives a bad signal, they will take short positions to exploit their information. Recall, PIN does not make a difference on the direction the informed traders place their orders, it is simply a measure of the fraction of informed trades. Therefore, we still expect a positive and significant coefficient on $Bad\ Event_t^i$. This implies that we expect the informed trading to increase in the wake of any signal as well in the wake of a negative signal. To take the noise trading activity into account, we use:

$$PIN_t^i = \lambda_{0_t}^i + \lambda_1 Any\ Event_t^i + \lambda_2 Bad\ Event_t^i + \lambda_3 Noise_t^i + \varepsilon_t^i \tag{32}$$

where all variables are defined as in (31) apart from the inclusion of $Noise_t^i$, which is the fraction of trading intensity derived from the noise traders per trading day. We expect the regression coefficient on $Noise_t^i$ to be negative and significant. Intuitively, as uninformed investors decrease their trading, the Probability of Informed Trading increases. This is in line with indications steaming from our model extension in section 3.3 and theoretical results presented by Easley et al. (1996). Finally, in our main PIN-regression, we include alpha:

$$PIN_t^i = \lambda_{0_t}^i + \lambda_1 Any\ Event_t^i + \lambda_2 Bad\ Event_t^i + \lambda_3 Noise_t^i + \lambda_4 Alpha_t^i + \varepsilon_t^i \tag{33}$$

where all variables are defined as in (32) apart from $Alpha_t^i$, which is the sum of all daily four-factor alphas for each stock during each quarter in the sample period. According to the hypothesis that alpha can be seen as a proxy for informed trading, we expect the coefficient on $Alpha_t^i$ to be positive and significant. Putting it differently, as alpha grows, the Probability of Informed Trading should grow if our hypothesis is valid. In the setting of the underlying model - if the informed traders receives a signal with respect to the true value of the firm's cash flow, they will cause the stock price to adjust towards the value indicated by their signal and thus, create abnormal returns.

6.4.2 Quoted Spread

The second robustness test incorporates Quoted Spread as the first stock liquidity proxy. The regression specifications are as follows:

$$\Delta IV_t^i = \lambda_{0_{t-1}}^i + \lambda_1^i Alpha_{t-1}^i + \lambda_2^i Quoted\ Spread_{t-1}^i + \varepsilon_{t-1}^i \tag{34}$$

$$\Delta IV_t^i = \lambda_{0_{t-1}}^i + \lambda_1^i Alpha_{t-1}^i + \lambda_2^i Quoted\ Spread_{t-1}^i + Y_T + \varepsilon_{t-1}^i \tag{35}$$

where the regression specifications are as in (25) apart from $Quoted\ Spread_{t-1}^i$, which is defined in equation (21). In specification (35) we include a year fixed effect, Y_T, to test the validity of the model with respect to time fixed effects, all other variables are defined as in (34).

We expect a negative and significant coefficient on the $Quoted\ Spread_{t-1}^i$ variable and after the inclusion of the year fixed effect, we expect similar results. The underlying logic steams from research presenting evidence on how informed traders are able to detect price inefficiencies and thus avoid stock trading if the stock is illiquid [20][27]. Therefore, informed traders will prefer use the option market to trade on their private information as stock illiquidity grows. As they do so, the option price will incorporate new information in a higher pace than the underlying stock price. This will drive down the potential demand shock at time t in the option market as more of the new information already has been priced in the option market due to the informed trading.

6.4.3 Amihud

The third robustness test incorporates the Amihud variable as a second proxy for stock liquidity. The regression specifications are as follows:

$$\Delta IV_t^i = \lambda_{0_{t-1}}^i + \lambda_1^i Alpha_{t-1}^i + \lambda_2^i Amihud_{t-1}^i + \varepsilon_{t-1}^i \tag{36}$$

$$\Delta IV_t^i = \lambda_{0_{t-1}}^i + \lambda_1^i Alpha_{t-1}^i + \lambda_2^i Amihud_{t-1}^i + Y_T + \varepsilon_{t-1}^i \tag{37}$$

Specification (36) is defined as in (34) besides the liquidity measure, which now is Amihud. This variable is defined in equation (22). In specification (37) we once again include a year fixed effect, Y_T, to test to model further and the other variables are defined as in (36). According to the intuition behind the Quoted Spread, we expect to see a negative and significant coefficient on the Amihud variable as well. This follows from the fact that both a high Quoted Spread and a high Amihud measure is a signal of stock illiquidity. Applying the same intuition as in the case of the Quoted Spread, if the stock is illiquid, informed traders will prefer to trade in the equivalent call option, reducing the potential demand shock in time t.

6.5 Trading Strategy

To test the statistical relationships derived in this paper, and investigate if there is any economic significance in our findings, we propose a self-financed trading strategy. We use the underlying predictive relationship between alpha, and implied- versus realized volatility. Furthermore, we also test whether the inclusion of stock trading activity during month $t-1$ have any implications for the economic significance in our empirical findings.

We implement a buy-hold delta-hedge strategy rather than a dynamically rebalanced delta-hedge strategy to reduce any trading costs [18]. It should be mentioned that our strategy deviates from a pure volatility strategy since we implement a buy-hold strategy. As we do so, each position in the strategy will generate a stock exposure as time passes by. We sell volatility (write call options) on stocks in decile portfolio 10. These stocks have had a high alpha during month $t-1$. We hedge the price exposure in the stock market and hold our positions for one month. In decile portfolio 1, we take the reversed position. We buy volatility (buy call options) on stocks with low or negative alphas during month $t-1$ and once again hedge our price exposure in the stock market and hold our positions for one month. This position formation will, on average, generate a positive payoff based on the respective spread in volatility measures [9].

The first strategy (hereafter "Original") is constructed by sorting stocks into decile portfolios based on individual four-factor alphas during month $t-1$. We rebalance these equally weighted portfolios each month and take new positions on the first day of each month. The second strategy (hereafter "Volume") is based on an interaction term. According to our main regression test, we sort stocks based on past month four-factor alpha and past month's change in trading activity. In portfolio 1, there are stocks with high increase in trading activity and low alphas and in portfolio 10, there are stocks with high decrease in trading activity and high alphas. As in the Original strategy, we take positions on the first day of each month and rebalance each portfolio on a monthly basis. This translates into the following expression:

$$r_t^{strategy} = \frac{1}{n_1}\sum_{i=1}^{n_1} r_{i,t}^{1} + \frac{1}{n_{10}}\sum_{j=1}^{n_{10}} r_{j,t}^{10} \tag{38}$$

where n_1 is the number of stocks in portfolio 1 and n_{10} the number of stocks in portfolio 10 and $r_{i,t}^{1}$ and $r_{j,t}^{10}$ is the monthly returns on the positions of each stock with its corresponding call option. Notice that the strategy returns, $r_t^{strategy}$, is computed by taking portfolio 1 plus portfolio 10. This is due to the fact that we have adjusted each portfolios positions in accordance to the direction indicated by the volatility spread.

Furthermore, we regress these monthly returns, minus the risk free rate in both the CAPM and in the four-factor model to estimate monthly alphas. The regression specifications are as follows:

$$\alpha_t^{strategy} = (r_t^{strategy} - r_t^f) - \left[\beta_1\left(r_t^{mkt} - r_t^f\right) + \varepsilon_t^i\right] \tag{39}$$

$$\begin{aligned}\alpha_t^{strategy} &= (r_t^{strategy} - r_t^f) \\ &\quad - \left[\beta_1\left(r_t^{mkt} - r_t^f\right) + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 UMD_t + \varepsilon_t^i\right]\end{aligned} \tag{40}$$

where $\left(r_t^{strategy} - r_t^f\right)$ is the monthly excess return on our two strategies, $\left(r_t^{mkt} - r_t^f\right)$ is the excess return on the market portfolio, SMB_t is the small-minus-big factor, HML_t is the high book-to-market minus the low book-to-market factor and UMD_t is the momentum factor. We run both these regressions on the Original and on the Volume strategy to see how the monthly alphas behave when we include the trading activity in the portfolio formation.

To further test the strategy, we decompose the positions into naked option- and stock positions respectively. We create one dependent variable that captures the short stock positions from portfolio 1 and long stock positions from portfolio 10, and a second variable that captures the long option positions from portfolio 1 and short option positions from portfolio 10. The regression specifications are as in (39) and (40), apart from the change in dependent variable, which now is excess returns on both stock- and option positions respectively. Since we intend to extract returns from the spread in the volatility measures, this decomposition will display in how the naked positions in both stocks and options perform. Consequently, this can be seen as an approximation to estimate where the returns steam from. Due to the same reasoning, we follow the methodology of Goyal & Saretto (2006) and include the VIX-risk factor. The intuition behind this is once again that we seek to profit from volatility exposure. This makes the inclusion of VIX a natural extension to the regressions:

$$\alpha_t^{strategy} = (r_t^{strategy} - r_t^f) - \left[\beta_1\left(r_t^{mkt} - r_t^f\right) + \beta_2 VIX_t + \varepsilon_t^i\right] \tag{41}$$

$$\alpha_t^{strategy} = (r_t^{strategy} - r_t^f) - \left[\beta_1\left(r_t^{mkt} - r_t^f\right) + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 UMD_t + \beta_4 VIX_t + \varepsilon_t^i\right] \tag{42}$$

where all variables are defined as in (39) and (40) apart from VIX_t, which is the monthly excess return on the CBOE VIX Index.

7 Empirical Results

In this section we describe the empirical results from the tests of volatility prediction, trading activity and market efficiency. We start by describing the descriptive aspects of our data in connection to our main variables. We also describe the volatility spread and monthly returns based on different portfolio categorizations. Secondly, we describe the test constructed to confirm the underlying predictive relationship, the extended empirical analysis and the robustness tests. At last, we describe the results of the proposed trading strategy and explain any economic significance. All regressions presented under section 7 are adjusted for heteroscedasticity.

7.1 Descriptive Statistics

Table 1 presents descriptive statistics per year for market capitalization, share volume per day, implied volatility and realized volatility. The average market capitalization increases throughout the sample period, from 24 billion USD in 2005 to 36.9 billion USD in 2014. However, during the most recent financial crisis in 2008 and 2009 there was a sizeable decrease in market capitalization. This decrease reached 19.4 billion USD in 2009, after which, there have been five consecutive years of increasing values. Volume, defined as the number of shares traded per day, is on average 5.5 million shares per day. Again, there is a sizeable change in pattern in the wake of the financial crisis in 2008. During this period the average share volume peaked at almost 8 million.

Throughout the sample period, the average implied volatility is to be found on a higher level compared to the corresponding realized volatility. Once again the exempt from this pattern arises within the years of 2008 and 2009. Notice, that during this time, both implied- and realized volatility increased considerable and peaked at 97.88% and 124.18% in 2008, respectively. This is expected since there is a negative correlation between volatility and returns. Interesting is that the standard deviation of realized volatility always exceeds the standard deviation of implied volatility.

Table 2 reports mean statistics over time for some of our key variables. Both the Quoted Spread and Amihud decreases over the sample period. This decrease in illiquidity indicates that the aggregate stock market liquidity has improved and thus, an enhanced stock market efficiency. However, the liquidity measures also peaked during the financial crisis at 0.1336 and 0.0229, respectively. The volume factor is the yearly average of the change in traded share volume between two sequential months. This factor is slightly positive, indicating the overall volume increase during the sample period. This increase in share volume is also displayed in Table 1.

Table 3 reports the monthly volatility spreads for 10 equally weighted portfolios sorted based on past month four-factor alphas. This spread is estimated by taking the implied volatility innovation minus the realized volatility innovation. In portfolio 1, this spread is on average -0.848% during the sample period. The corresponding spread for portfolio 10 is on average 3.13%. This supports AABCs predictions that alpha predicts a spread in volatility measures. Our goal is to extract returns from this spread in volatility measures.

Table 4 presents future returns and standard deviation for 10 equally weighted portfolios. The returns is computed as writing a call option and hedging the position in the stock market. Each position is taken on the first day of month t, and held for one month. In panel A we sort stocks based on four-factor alphas during month $t-1$ and rebalance the portfolios each month. This generates an average monthly return in portfolio 10 of 1.44% while portfolio 1 has an average monthly return of 1.23%. Thus, the strategy of being long portfolio 10 and short portfolio 1 generates approximately 21.5 bp on average per month. This is consistent with our findings in Table 3 and theoretical models suggesting that option strategies in combination with stock market hedging can extract returns from the spread in volatility measures.

Panel B, C and D displays the average monthly returns for portfolio 1 and portfolio 10 based on the interaction term deciles: $(Alpha * Volume\ factor)$, $(Alpha * Quoted\ spread)$ and $(Alpha * Amihud)$ respectively. In panel B, portfolio 1 earns an average monthly return of 0.91% while portfolio 10 earns an average monthly return of 1.22%. Consequently, the combination of being long portfolio 10 and short portfolio 1 on average generates a monthly return of 30.5 bp. This interaction term is of particular interest since it steam from our main regression test in section 6.3. As described, we predict that the combination of high alpha and high decrease in stock trading activity will forecast a higher probability of informed trading. Due to this, we expect a more pronounced demand shock and thus, a more precise forecast of future returns. This is confirmed as portfolio 1 generates a return of 0.91%, in comparison to 1.23% if we sort solely based on four-factor alpha. Accordingly, on average, the combination of portfolio 10 minus portfolio 1, earns 9 bp more per month if we compare to the sorting based exclusively on four-factor alphas.

Conferring to our results in the robustness specifications using the liquidity proxies, panel C and panel D show the inverse pattern. In panel C, portfolio 1 generates a monthly of 1.97% on average while portfolio 10 generates 0.89% on average per month. Panel D displays the same relationship. Portfolio 1 generates a monthly return of 1.96% on average and portfolio 10 generates an average monthly return of 1.31%. This implies that the combination of being long portfolio 10 and short portfolio 1 earns -108 bp and -65 bp respectively for panel C and panel D. This indicates that portfolio 1 earns higher returns than portfolio 10 if stocks are sorted on the interaction deciles referring to the liquidity measures.

Table 5 presents delta-exposure per month and portfolio. The portfolios are equally weighed and rebalanced monthly based on preceding month four-factor alpha. The mean monthly exposure for portfolio 1 is 0.0404. According to a dynamically rebalanced delta-hedge strategy, we would have to adjust our stock holdings to eliminate this slight increase in stock exposure. Putting it differently, we would have to sell, on average, 0.0404 stocks more per position to keep portfolio 1 delta-neutral. In portfolio 10, the corresponding value is 0.0392. Important to highlight is that portfolio 10 writes call, implying that in practice, this number is negative. However, we seek to exclude any subjective assessments in which portfolio the short call positions begins and the long call positions ends. This is why we present all mean delta-exposures in absolute values. Using the same logic, we would have to buy, on average, 0.0392 stocks more per position in portfolio 10 to keep the delta-neutrality. Therefore, the strategy of being short volatility in portfolio 10 and long volatility in portfolio 1, creates a positive stock exposure on average, with a delta value of 0.00127 per month. This result indicates that the strategy keeps a delta fairly close to zero despite our choice of implementing a buy-hold strategy rather than a dynamically rebalanced strategy.

7.2 Test of Predictive Relationship

The descriptive statistics in Table 3 and Table 4 indicates that there is a double sided predictive relationship from alpha. On one hand we see an increase in implied volatility and on the other hand, a decrease in realized volatility. We verify this relationship by running the regression specifications in (23) and (24) and the results are presented as average coefficients in Table 6. Column (1) display how implied volatility (ΔIV_t) during month t behave in relation to CAPM alpha during month $t-1$. The regression coefficient on alpha is 3.167 (t-stat 4.18). The interpretation is that if alpha grows by 1% during month $t-1$, implied volatility is predicted to increase by 3.167% the following month.

Column (2) display the same methodology apart from the dependent variable which now is the change in realized volatility ($\Delta RVOL_t$) during month t. The coefficient on alpha converts to negative with a value of -10.879 (t-stat -5.07). Once again, the interpretation is that if alpha grows by 1% during month $t-1$, realized volatility is predicted to decrease by 10.879% the subsequent month. In both (1) and (2) is the regression constant insignificantly different from zero. This indicates that if CAPM alpha is zero during month $t-1$, the impact on ΔIV_t and $\Delta RVOL_t$ cannot be assumed to be either positive or negative. These results support the findings presented in Table 3, the results of AABC and confirms the intuition of the underlying model.

Continuing with our extended empirical tests, the focus is to forecast the demand shock emerging in month t. To do this we seek to improve the model to capture factors which have a more pronounced impact on ΔIV_t. As described in section 6.2 we present a hypothesis that four-factor alphas should be a more sophisticated proxy signal for informed trading as it removes systematic risk exposure. By running the regression specification (25), this test is presented as average coefficients in Table 7. In column (1) we present the original findings derived from the CAPM alpha, these results are the same as in column (1) in Table 6. In column (2) we replace the CAPM alpha with the new four-factor adjusted alpha. The regression coefficient on alpha increases by 15.77% to 3.677 (t-stat 4.99), indicating that the impact on implied volatility increases as we transform the CAPM alphas to four-factor adjusted alphas. Due to the usage of separately computed alphas is it possible to rule out any statistical inference such as multicollinearity to be the driver behind this result. This results confirms our hypothesis, indicating that four-factor alphas possibly can be seen as a more sophisticated proxy for informed trading. As earlier, if alpha grows by 1% during month $t-1$, implied volatility is predicted to increase by 3.677% the succeeding month.

7.3 Main Test

Having confirmed the underlying predictive relationship, we reinforce our focus on predicting the impact on ΔIV_t. We use regressions specifications (26) to (29) to test the relationship between four-factor alpha, trading activity and the projected demand shock. Table 8 presents the average coefficients of main regression tests, where we also include regression specification (25) in column (1) as a benchmark. In column (2) we include the volume factor, which aims to capture the change in stock trading activity during month $t-1$. The coefficient on alpha is still positive and significant. Notice that the coefficient is negative with a value of -1.881 (t-stat -2.02). Indicating that the demand shock and thus, the impact on ΔIV_t, will be positive if the trading activity during month $t-1$ has decreased. This is in line with the prediction (ii) and our model extension in section 3.3. This also confirms the results presented by Easley et al. (1996) and the intuition that the demand shock will be more pronounced if the stock trading activity the month preceding the estimation of ΔIV_t has decreased.

In column (3) we include a dummy which takes a value of 1 in volume factor decile 1 and 0 otherwise. Putting it differently, in cases where the trading activity has decreased the most during month $t-1$, the dummy takes on a value of 1. This dummy makes the interpretation of trading activity's impact on implied volatility during month t somewhat easier. The other variables are fundamentally the same and so are their significance. Notice that the dummy coefficient have a positive value of 0.665 (t-stat 3.17), confirming the hypothesis that a decreased trading activity during month $t-1$ has a positive and significant impact on ΔIV_t.

In column (4) we include a year fixed effect to test the relationship with respect to systematic shocks. This corresponds to regression specification (29). All values are qualitatively the same with small or virtually non-existing changes and still significant on all conventional level. This strengthens the predictions presented in column (3) and thus, supports the hypothesis that trading activity and alpha truly have a significant influence on future implied volatility.

7.4 Robustness Test

In this section we discuss and describe the three robustness test performed. Firstly, we present the test of our PIN measure in connection to alpha. Secondly, we present robustness test where we link stock illiquidity to the impact on implied volatility using the Quoted Spread and Amihud.

7.4.1 Probability of Informed Trading

Table 9 presents the first robustness test, where we strive to test the relationship between PIN and four-factor alphas. This test is different from the two other robustness tests since it does not test the relation to the impact on ΔIV_t. It rather investigates the hypothesis that alpha can be seen as a proxy for informed trading. Important to highlight is the fact that due to data availability issues, this test only covers half of the total sample period. Column (3) and (4) presents regression specifications (32) and (33) respectively. In (3) the coefficient on Any event is positive with a value of 0.214 (t-stat 26.14) and the coefficient on Bad event is 0.0106 (t-stat 4.189). These two coefficients suggest that the predictions of the underlying model are valid. Implying that, as informed investors receives a private signal, independent of direction of the signal value, the probability of informed trading increases.

The coefficient on the Noise variable is -0.486 (t-stat -22.71). As uninformed demand decreases, the probability of informed trading increases, supporting our extension of the underlying model in section 3.3 and confirming the results of Easley et al. (1996). In column (4) we use the same specification, but include the main variable, the four-factor alpha. The other coefficients are fundamentally the same and their significance does not change. The coefficient on alpha is positive with a value of 0.0016 (t-stat 1.883) indicating a significance close to 5%. This truly supports our prediction in section 3.2.1, where we hypothesize that alpha and informed trading is positively connected. To our knowledge there are no scholars who has presented similar indications. This robustness test suggests that there is a relation between alpha, stock trading activity and informed trading.

7.4.2 Quoted Spread

In Table 10 we present the result from our second robustness test. We include the Quoted Spread as proxy for stock liquidity and the regression specifications are as in (34) and (35). As earlier, in column (1), the result from column (2) in Table 7 is presented as a benchmark. We include the Quoted Spread in column (2). The coefficient on Quoted Spread has a negative coefficient with a value of -3.649 (t-stat -2.26). This indicates that the impact on implied volatility is decreasing in stock illiquidity. This result confirms the prediction (iii) under section 3.2.1 and research indicating that informed traders are able to circumvent and profit from fundamental mispricing, such as illiquidity in the stock. Consequently, as stock illiquidity increases, the informed traders will prefer to exploit their private information in the equivalent call option. As they do so, the potential demand shock decreases since the call option price will have incorporated more of the new information steaming from the informed traders activity. This is in line with findings presented by Kyle (1985) and Admati & Pfleiderer (1988).

Finally, in column (3), we include a year fixed effect which captures systematic shocks. As in Table 8, the changes in the coefficients and their significance is practically non-existing. This supports the findings that there is a negative relationship between the impact on implied volatility in month t, and stock illiquidity in month $t-1$.

7.4.3 Amihud

Table 11 presents the third robustness test. Using regression specifications (36) and (37) we test how the predictive relationship behaves when including the second liquidity proxy, Amihud. Column (1) displays the benchmark finding from column (1) in Table 7. In column (2) the alpha coefficient is still positive and significant. Notice that the coefficient on Amihud is negative with a value of -7.213 (t-stat -2.02). This indicates, as in the robustness test with the Quoted Spread variable, that the impact on ΔIV_t decreases in stock illiquidity. This is due to the fact that a high Amihud and a high Quoted Spread are proxies for stock illiquidity.

In column (3) we include a year fixed effect to once again test our specifications with respect to systematic shocks. The results are in line with earlier results from Table 8 and 10. Putting it differently, the change of the values on the coefficients are small and their significance are fundamentally the same. This supports the findings that the impact in ΔIV_t is truly decreasing in stock illiquidity.

7.5 Trading Strategy

Using the statistically significant relationships established in Table 6 and in Table 8, we now investigate whether there is any economic significance in our findings. In Table 12, we present the first trading strategy (Original strategy), where we sort stocks into decile portfolios based on four-factor alphas during month $t-1$. In Table 13, we present the findings on our second strategy (Volume strategy), using the interaction term where we include the volume factor into the original sorting.

Table 12 displays monthly alphas for the Original strategy, both with and without adjusting for Fama-French-Carhart factors. In columns (1) and (4) we present the monthly factor exposures and alphas for portfolio 1 and in columns (2) and (5), the corresponding results for portfolio 10. Column (3) displays the Original strategy exposure to the market portfolio and the monthly CAPM alpha. Notice that the strategy has a negative exposure to the market with a coefficient of -1.415 (t-stat -18.75) and a monthly CAPM alpha of 66.4 bp (t-stat 3.314). Column (6) displays the equivalent result when we adjust for Fama-French-Carhart factors. It is important to highlight the fact that in both column (3) and (6), the strategy alphas is computed by taking portfolio 10 plus portfolio 1. This is due to the fact that the portfolio returns have been adjusted to the directional positions that each of the two portfolios take. The market exposure is still negative with a value of -1.044 (t-stat -14.05). The strategy has positive exposure to the high-minus-low factor, 0.0443 (t-stat 0.415), and the momentum factor, 0.655 (t-stat 4.958). The small-minus-big factor is negative, -0.585 (t-stat -4.745).

All factor exposures, exempt high-minus-low, is statistical significant on all conventional levels. The monthly four-factor adjusted alpha decreases to 51.8 bp (t-stat 2.497) from 66.4 bp (t-stat 3.314) if using the CAPM. This implies that the Original strategy have positive four-factor alphas in 95% of all months and positive CAPM alphas in 99% of all months. Thus, the Original strategy earns abnormal returns that is both of statistical and economic significance. In Graph 1, we display the abnormal returns generated by the Original strategy. Over the sample period, the Original strategy earns four-factor adjusted abnormal returns of 65.93%.

Table 13 display the corresponding results referring to the Volume strategy introduced in section 6.5. In this strategy, we sort stocks into deciles based on the interaction term of four-factor alphas and trading activity during month $t-1$. In column (1) and (4) we display the monthly alphas and factor exposures for portfolio 1, both using CAPM and the extended four-factor model. Column (2) and (5) displays the corresponding results for portfolio 10. In column (3), the Volume strategy is presented using CAPM. As in the Original strategy, the exposure to the market portfolio is negative and have a coefficient of -1.831 (t-stat -15.87). The monthly CAPM alpha is 1.17% (t-stat 3.880), which is 50.6 bp higher than what the Original strategy generated under the same model specification.

In column (6), we show the corresponding results using Fama-French-Carhart adjustments. Once again, the coefficient of the market portfolio is significant and negative, with a value of -1.332 (t-stat -12.34). The value is 0.228 higher in absolute terms than the exposure that the Original strategy had given the same model. The increase in market beta indicates that the Volume strategy is more volatile if compared to the Original strategy. The small-minus-big factor, -0.602 (t-stat -3.138) and momentum factor, 1.050 (t-stat 4.756), are still significant at all conventional levels. Notice that the momentum factor exposure increases considerable when we include the volume factor in the sorting process. The high-minus-low factor is still statistically insignificant with a value of 0.101 (t-stat 0.664). The monthly four-factor alpha is somewhat decreasing to 98.2 (t-stat 3.182) from CAPM alpha of 1.17% (t-stat 3.880), but is still significant. Once again, if this is compared to the four-factor alpha generated by the Original strategy, it is 46.4 bp higher.

In Graph 2, we display the abnormal returns generated by the Volume strategy. Over the entire sample period, the strategy earns four-factor adjusted abnormal returns of 112.68%. In Graph 3, we display a comparison of the adjusted abnormal returns generated by the two strategies. Over the 117 months, the Volume strategy earns 46.75 percentage points more than the Original strategy, indicating the economic significance of including trading activity in the prediction on implied volatility.

Table 14 presents the decomposed positions for both portfolio 1 and portfolio 10 in the Volume strategy. In column (1) and (4) we use the CAPM and the four-factor model to regress excess stock returns. The monthly CAPM alpha is -0.635 bp (t-stat 0.378) and the monthly four-factor adjusted alpha is -13.5 bp (t-stat -0.798). In column (2) and (5) the corresponding alphas for the decomposed option returns are, -26.2 bp (t-stat -0.131) and 18.5 bp (t-stat 0.0923). This indicates that the alpha generated by the Volume strategy does not steam from neither naked positions in stock nor in options, but rather the composition of the positions. This suggests that the alphas generated by our strategy steam from volatility mispricing. In (3) and (6) we include the VIX risk factor. The Volume strategy exposure to this factor is -0.0206 (t-stat 0.811) in the CAPM and 0.00915 (t-stat 0.384) in the four-factor model. After the inclusion of VIX, the alphas is virtually the same and still statistically significant.

8 Conclusion

This paper sets out to investigate the cross-sectional relationship between past abnormal stock return and future implied volatility. Building on research investigating this cross-sectional connection and research exploring stock illiquidity, trading activity and information asymmetry, we seek to improve the original relationship. By doing so, we find empirical evidence suggesting a negative relationship between stock trading activity and future implied volatility and a negative relationship between stock illiquidity and future implied volatility.

Furthermore, we find new evidence indicating that alpha is positively and significantly correlated to the probability of informed trading and that the exclusion of systematic risk factors from alpha generates a more sophisticated signal of informed trading. Additionally, we find results suggesting that informed traders use the call option to exploit their superior information if the stock is illiquid. As we use a lagged Fama-French-Carhart alpha variable and a lagged stock trading activity variable, we enhance the original relationship and find indications of a more pronounced impact on future implied volatility. These findings are robust under several firm level cross-sectional regression specifications and under the inclusion of year fixed effects.

Based on these empirical pattern we present a self-financed trading strategy which aims to extract returns from the spread in volatility measures. We develop one strategy which is based on lagged Fama-French-Carhart alphas and another strategy that incorporates a lagged stock trading activity variable. As this is done, the cumulative abnormal returns increases 46.75 percentage points over the sample period, stretching from April 2005 to December 2014. This suggests that stock trading activity has an important role both in the predicting future implied volatility and in identifying the probability of informed trading. In addition to this, we find that the abnormal returns of the strategies does not steam from neither naked positions in options nor in stocks, but rather the composition of the strategy. Suggesting a volatility mispricing.

This study contributes to the existing literature in two ways. First, as our results indicates a negative impact on implied volatility in the wake of stock illiquidity, the findings contributes to research investigating the trading behavior of informed traders. Second, as the inclusion of trading activity in the original framework improves the predictability of implied volatility, our findings also contributes to the field of research concerning volatility predictions and mispricing.

Finally, we want to underline the limitations of our study and also, highlight fields of future research connected to our paper. Our first issue is data related. As our data only covers large stocks in the US, this imposes a risk that the relations presented in this study are isolated to this particular investment universe. Another data related issue is that our data on Probability of Informed Trading (PIN) is quarterly and only covers half of our sample period. These data issues could impose risks of sample biases. The second issue is related to our extended empirical research. We use variables which potentially are more accurate over longer time-periods. Also, the liquidity measures used in this study can be seen as proxies for informed trading by on a stand-alone basis. Another issue connected to this is that we only investigate stock illiquidity in our extended empirical study. It is well documented that the illiquidity in options and thus, the transaction costs for trading in option markets, is somewhat higher than in stock markets. These aspects could also, to some extent, generate biased results. As for future research, we suggest two main starting points. Since we have established a positive and significant relationship between alpha and PIN on a quarterly basis, one future starting point is to refine this relationship. For instance, testing the relation in different timeframes and a larger sample could bring insights into the significance of alpha in the setting of informed trading. Another important aspect is to investigate the aggregate risks in relation to volatility returns and mispricing. Based on the empirical results that our strategy generates monthly abnormal returns of 98.2 bp, there could be an unknown systematic risk factor yet to be discovered.

As for our study, we conclude that our empirical results indeed suggests a significant relationship between informed trading, trading activity and volatility mispricing.

References

[1] Admati, Anat R., and Paul Peiderer, 1988, A theory of intraday patterns: Volume and price variability. *Review of Financial Studies* 1, pg. 3-40.

[2] Amihud, Yakov, 2002, Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets* 5, pg. 31-56.

[3] Amin, Kaushik, Joshua D. Coval, and H. Nejat Seyhun, 2004, Index option prices and stock market momentum, *Journal of Business* 77, pg. 835-873.

[4] An, Byeong-Je, Andrew Ang, Turan G. Bali, and Nusret Cakici, 2014, The Joint Cross Section of Stocks and Options. *The Journal of Finance* 69, pg. 2279-2337.

[5] Bessembinder, Hendrik, and Herbert M Kaufman, 1997, A comparison of trade execution costs for NYSE and Nasdaq-listed stocks. *Journal of Financial and Quantitative Analysis* 32, pg. 287-310.

[6] Black, Fisher, 1975, Fact and fantasy in the use of options, *Financial Analysts Journal* 31, pg. 36- 72.

[7] Black, Fischer, and Myron Scholes, 1973, The pricing of options and corporate liabilities. *The Journal of Political Economy*, pg. 637-654.

[8] Carhart, M. Mark, 1997, On Persistence in Mutual Fund Performance, *The Journal of Finance* 52, pg. 57-82

[9] Carr, Peter and Dilip Madan,. 2001, Towards a Theory of Volatility in J. Cvitanic,. E. Jouini and Marek Musiela.: Trading, Option Pricing, Interest Rates and Risk Management,. (Cambridge, *Cambridge University Press.*,) pg. 458-477

[10] Chakravarty, Sugato, Huseyin Gulen, and Stewart Mayhew, 2004, Informed trading in stock and option markets. *The Journal of Finance* 59, pg. 1235-1258.

[11] Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam, 2001, Market Liquidity and Trading Activity. The Journal of Finance 56, pg. 501-530

[12] Copeland, E. Thomas,. and Dan Galai, 1983, Information Effects on the Bid-Ask Spread, *The Journal of Finance* 38, pg. 1457-1469

[13] Cremers, Martijn, and David Weinbaum, 2010, Deviations from put-call parity and stock return predictability, *Journal of Financial and Quantitative* Analysis 45, pg. 335-367.

[14] Easley, David, Nicholas M. Kiefer, Maureen O'Hara, and Joseph B. Paperman, 1996, Liquidity, Information and Infrequently Traded Stocks. *The Journal of Finance* 51, pg. 1405-1436.

[15] Easley, David, Maureen O'Hara, and Pulle Subrahmanya Srinivas, 1998, Option volume and stock prices: Evidence on where informed traders trade. *The Journal of Finance* 53, pg. 431-465.

[16] Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, pg. 3-56.

[17] Goyal, Amit, and Alessio Saretto, 2006, Option Returns and Volatility Mispricing. *Working paper.*

[18] Goyal, Amit, and Alessio Saretto, 2009, Cross-section of option returns and volatility, *Journal of Financial Economics* 94, pg. 310-326.

[19] Goyenko, Ruslan Y., Craig W Holden, and Charles A Trzcinka, 2009, Do liquidity measures measure liquidity? *Journal of Financial Economics* 92, pg. 153-181.

[20] Hayunga, Darren K., Richard D. Holowczak, Peter P Lung, and Takeshi Nishikawa, 2012, Derivatives traders reaction to mispricing in the underlying equity. *Journal of Banking & Finance* 36, pg. 2438-2454.

[21] Huang, Roger D., and Hans R Stoll, 1996, Dealer versus auction markets: A paired comparison of execution costs on Nasdaq and the NYSE. *Journal of Financial Economics* 41, pg. 313-357.

[22] Johnson, Travis L., and Eric C So, 2012, The option to stock volume ratio and future returns. *Journal of Financial Economics* 106, pg. 262-286

[23] Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal of Finance* 48, pg. 65-91.

[24] Kyle, Albert S, 1985, Continuous auctions and insider trading. Econometrica: *Journal of the Econometric Society*, pg. 1315-1335.

[25] Muravyev, Dmitriy, Neil D. Pearson, and John Paul Broussard, 2013, Is there price discovery in equity options? *Journal of Financial Economics* 107, pg. 259-283.

[26] Pan, Jun, and Allen M. Poteshman, 2006, The information in option volume for future stock prices. *Review of Financial Studies* 19, pg. 871-908.

[27] Sadka, Ronnie, and Anna Scherbina, 2007, Analyst disagreement, mispricing, and liquidity. *The Journal of Finance* 62, pg. 2367-2403.

[28] Sarr, Abdourahmane, and Tonny Lybek, 2002, Measuring liquidity in financial markets. *Working paper.*

[29] Venter, Johannes H., and Dawid C.J. de Jongh, 2006, Extending the EKOP model to estimate the probability of informed trading. *Studies in Economics and Econometrics* 30, pg. 25-39

Appendix A

Table 1: Descriptive Statistics by Year

Table 1 reports descriptive statistics by year for Market Capitalization (in thousands USD), Volume (number of shares traded per day), Implied volatility (interpolated dataset from Bloomberg L.P) and Realized volatility (estimated by the standard deviation of daily returns using a 30 day estimation window). Statistics is grouped by year and measures are Mean, Median, Min, Max, Standard deviation and Number of observations. The sample period is April 2005 to December 2014 and covers 117 months, which equals 901.142 daily observations and 497 unique tickers.

Year	Variable	Mean	Median	Min	Max	Std.	N
2005	Market Cap	24000000	11200000	2000016	394000000	39700000	64768
	Volume	3501332	1469800	12200	254000000	7581212	64768
	Implied Vol	24.06%	23.00%	13.37%	40.00%	9.18%	64768
	Realized Vol	23.09%	20.79%	11.46%	42.55%	10.52%	64768
2006	Market Cap	25500000	12600000	2000156	393000000	40100000	88548
	Volume	4097523	1727850	27400	395000000	8829329	88548
	Implied Vol	24.92%	23.35%	13.27%	42.72%	9.99%	64768
	Realized Vol	23.65%	20.96%	10.92%	44.95%	11.23%	88559
2007	Market Cap	27900000	13900000	2003159	432000000	42800000	91898
	Volume	5072392	2353400	10400	288000000	8941274	91898
	Implied Vol	27.43%	26.56%	14.02%	45.06%	10.67%	91909
	Realized Vol	26.26%	24.10%	11.99%	48.06%	12.91%	91909
2008	Market Cap	23700000	10900000	2000197	384000000	36900000	92704
	Volume	7568494	3633150	88600	1200000000	15000000	92704
	Implied Vol	48.32%	41.56%	22.58%	97.88%	25.04%	92707
	Realized Vol	52.77%	41.78%	19.63%	124.18%	35.57%	92707
2009	Market Cap	19400000	8916240	2000158	279000000	30200000	91392
	Volume	8655137	3944800	49000	1230000000	24900000	91392
	Implied Vol	45.93%	41.26%	21.63%	86.31%	22.03%	91393
	Realized Vol	46.31%	39.05%	16.88%	102.14%	29.24%	91393

2010	Market Cap	23300000	11200000	2002221	300000000	35300000	94244
	Volume	6824013	3306750	62100	655000000	15000000	94244
	Implied Vol	30.24%	29.00%	16.90%	48.05%	9.72%	94244
	Realized Vol	27.91%	25.97%	12.85%	49.58%	11.80%	94244
2011	Market Cap	25700000	12600000	2009465	392000000	38600000	94407
	Volume	6612519	3205600	90300	860000000	15700000	94407
	Implied Vol	31.52%	29.23%	16.53%	54.44%	12.09%	94408
	Realized Vol	31.43%	27.65%	12.70%	62.63%	15.91%	94408
2012	Market Cap	27700000	13000000	2001464	658000000	46500000	93713
	Volume	5567823	2683942	51769	669000000	13600000	93713
	Implied Vol	25.76%	24.47%	13.59%	42.83%	9.30%	93713
	Realized Vol	24.22%	22.10%	11.33%	44.56%	11.04%	93713
2013	Market Cap	32100000	15400000	2002766	516000000	48000000	94651
	Volume	4672946	2300900	200	336000000	9678823	94651
	Implied Vol	22.63%	21.42%	13.73%	35.97%	7.33%	94651
	Realized Vol	21.48%	19.65%	11.84%	37.36%	8.94%	94651
2014	Market Cap	36900000	17500000	2001936	698000000	55400000	94817
	Volume	4301722	2250873	100	618000000	8059025	94817
	Implied Vol	21.55%	20.04%	13.20%	35.09%	7.05%	94817
	Realized Vol	20.33%	18.33%	10.58%	36.59%	9.00%	94817

Table 2: Descriptive Statistics of Key Variables by Year

Table 2 reports mean value for some of the key variables over time. Quoted Spread and Amihud have been converted into percent. Quoted Spread is defined as the average Quoted Spread for each stock during each month of the sample period. Amihud is defined as the average Amihud measure per month for each stock in the sample period. The volume factor is the percental change in traded volumes between two successive months. The sample period is April 2005 to December 2014 and covers 117 months.

Year	Quoted Spread	Amihud	Volume Factor
2005	0.06851	0.02415	0.03924
2006	0.05958	0.01763	0.05805
2007	0.09533	0.01402	0.08929
2008	0.13367	0.02297	0.10571
2009	0.08713	0.02143	0.01071
2010	0.04289	0.01305	0.03216
2011	0.03561	0.01407	0.05492
2012	0.0345	0.01168	0.02929
2013	0.02795	0.00998	0.02738
2014	0.0391	0.00832	0.05543
N	43042	43042	43042

Table 3: Descriptive Statistics of Volatility Spread by Portfolio

Table 3 reports mean value of the volatility spread by portfolio. Each decile portfolio is equal weighted and rebalanced monthly based on past month Fama-French-Carhart alpha. The volatility spread is the difference between the implied volatility innovation and realized volatility innovation. These innovations are estimated by taking the respective volatility measure of the last day in each month, minus the same volatility measure on the first day of the same month. Mean values are in percent. The sample period is April 2005 to December 2014 and covers 117 months.

	Mean	N
Portfolio 1	-0.848413	117
2	-0.804967	117
3	-1.465727	117
4	-1.205179	117
5	-1.058992	117
6	-0.928294	117
7	-0.400337	117
8	-0.069749	117
9	0.862999	117
Portfolio 10	3.137027	117
P10-P1	3.98544	

Table 4: Descriptive Statistics by Alpha and Interaction Portfolios

Table 4 reports the mean, min and max return and standard deviation for 10 equally weighted portfolios. We also display the returns on the combination of being long portfolio 10 and short portfolio 1. In panel A we sort stocks into 10 portfolios based on Fama-French-Cahart alpha during month $t-1$. Panel B, C and D show the same result for portfolio 1 and portfolio 10, when we sort stocks based on the interaction terms:($Alpha * Volume\ factor$), ($Alpha * Quoted\ spread$) and ($Alpha * Amihud$) respectively during month $t-1$. The sample period is April 2005 to December 2014 and covers 117 months.

Panel A: Descriptive Statistics by $Alpha$

	Mean	Min	Max	Std	N
Portfolio 1	.0123188	-.0959870	.0764953	.0290042	117
2	.0104104	-.0886831	.0978279	.0258158	117
3	.0084739	-.0866281	.0531696	.0213907	117
4	.0075536	-.0862920	.0498225	.0198480	117
5	.0094270	-.0451143	.0655364	.0190884	117
6	.0084182	-.0977164	.0592211	.0193448	117
7	.0092375	-.0588119	.0554914	.0190458	117
8	.0113865	-.0875307	.0693233	.0190112	117
9	.0127941	-.0364001	.0957132	.0190614	117
Portfolio 10	.0144713	-.0613167	.0835405	.0202945	117
P10-P1	.0021525				

Panel B: Descriptive Statistics by ($Alpha * Volume\ factor$)

	Raw Return	Min	Max	Std	N
Portfolio 1	.0091307	-.2395752	.1057747	.042204	117
Portfolio 10	.012173	-.0813101	.0898715	.0247385	117
P10-P1	.0030423				

Panel C: Descriptive Statistics by ($Alpha * Quoted\ spread$)

	Raw Return	Min	Max	Std	N
Portfolio 1	.0197692	-.1864986	.4823417	.0685943	116
Portfolio 10	.0089418	-.2236244	.1413307	.048162	116
P10-P1	-.0108274				

Panel D: Descriptive Statistics by ($Alpha * Amihud$)

	Raw Return	Min	Max	Std	N
Portfolio 1	.0196698	-.1803331	.1898491	.0467767	109
Portfolio 10	.0131333	-.1057555	.1234497	.035470	109
P10-P1	-.006537				

Table 5: Descriptive Statistics – Monthly Delta Exposure

Table 5 reports the mean, min, max, standard deviation and number of months on the delta change in month t for 10 equally weighted portfolios sorted on Fama-French-Carhart alpha during month $t-1$. The delta change is computed as the difference between the delta neutral positions at inception of month t and the final delta exposure for each position during the same month. We display statistics by each portfolio and the values are computed monthly. The sample period is April 2005 to December 2014 and covers 117 months.

	Mean	Min	Max	Std	N
Portfolio 1	.0404942	-.1641088	.1982985	.0738816	116
2	.0423201	-.1085567	.2046431	.0688033	116
3	.0401331	-.1862767	.1789523	.0658368	116
4	.0351204	-.1694063	.1897836	.0726074	116
5	.0463669	-.2552181	.2194012	.0784302	116
6	.0289616	-.1629642	.163725	.0679905	116
7	.0563796	-.1743624	.2781543	.0692123	116
8	.0439783	-.1547773	.2688552	.0794642	116
9	0482783	-.191205	.2983137	.0791402	116
Portfolio 10	.0392233	-.1579664	.2619957	.0767235	116

Table 6: Confirmation of Predictive Relationship

Table 6 displays the main predictive relationship used in this study and results are based on An et al. (2014) findings. We regress CAPM alpha for each stock during month $t-1$ on the first difference in implied volatility in (1) during month t and the first difference in realized volatility in (2) during month t. The monthly alpha is the sum of all daily alphas during month $t-1$ for each stock i. The first difference in implied- and realized volatility is the difference between the volatility of the last day in month t, and the volatility of the first day of month t. The table displays the average coefficients. The sample period stretches from April 2005 to December 2014 and covers 117 months.

	(1)	(2)
VARIABLES	ΔIV	ΔRVOL
Alpha	3.176***	-10.879***
	(4.18)	(-5.07)
Constant	-0.452	0.436
	(-0.99)	(0.52)
Observations	117	117

Robust t-statistics in parentheses
*** p< 0.01, ** p< 0.05, * p< 0.1

Table 7: Inclusion of Fama-French-Carhart Alpha

Table 7 displays the inclusion of Fama-French-Carhart (FFC) alphas. We regress both CAPM and FFC alphas for each stock during month $t-1$ on the first difference in implied volatility during month t. The monthly alpha is the sum of all daily alphas during month $t-1$ for each stock i. The first difference in implied- volatility is the difference between the volatility of the last day in month t, and the volatility of the first day of month t. The table displays the average coefficients. (1) displays the regression using CAPM estimates on alphas and (2) displays FFC estimates on alphas. The sample period stretches from April 2005 to December 2014 and covers 117 months.

	(1)	(2)
VARIABLES	ΔIV	ΔIV
Alpha	3.176***	3.677***
	(4.18)	(4.99)
Constant	-0.452	-0.474
	(-0.99)	(-1.04)
Observations	117	117

Robust t-statistics in parentheses
*** p< 0.01, ** p< 0.05, * p< 0.1

Table 8: Main Test

Table 8 presents the regression specifications used in the main test. We use Fama-French-Carhart (FFC) alphas defined as the sum of all daily alphas during month $t-1$ for each stock i, Volume is defined as the change in average traded volume between month $t-1$ and $t-2$, for each stock i. The dummy takes a value of 1 if volume decile equals 1 (largest decrease in average traded volume between month $t-1$ and $t-2$) and a value of 0 otherwise. The variable Year is a dummy, which captures year fixed effects. The table displays the average coefficients. (1) displays the regression using FFC alphas only, (2) displays the same regression when including the volume factor. In (3) we include the dummy and in (4) the year fixed effect. The sample period stretches from April 2005 to December 2014 and covers 117 months.

	(1)	(2)	(3)	(4)
VARIABLES	ΔIV	ΔIV	ΔIV	ΔIV
Alpha	3.677***	3.426***	3.492***	3.387***
	(4.99)	(4.70)	(4.80)	(4.68)
Volume Factor		-1.881**	-0.727***	-0.717***
		(-2.02)	(-4.29)	(-4.27)
Dummy			0.665***	0.487***
			(3.17)	(4.40)
Year				0.740**
				(2.56)
Constant	-0.474	-0.435	-0.673	-3.963**
	(-1.04)	(-0.93)	(-1.31)	(-2.52)
Observations	117	117	117	117

Robust t-statistics in parentheses
*** p< 0.01, ** p< 0.05, * p< 0.1

Table 9: Robustness Test 1 - Probability of Informed Trading (PIN)

Table 9 displays the first robustness test using quarterly PIN data from Stephen Browns' website at Robert H. Smith. PIN is based on Easley et al. (1996) and is a measure based on trading intensity of informed and uninformed investors in connection to any information event. Any Event is the probability of an information based event and Bad Event is the probability of Any Event to be bad news. Noise is the fraction of uninformed trading intensity during one trading day. The alpha variable is computed as earlier, by taking the sum of all daily Fama-French-Carhart alphas during the estimation window, which now is one quarter. The sample period is April 2005 to December 2010, and covers 69 months, 23 quarters and 2.102 unique observations

	(1)	(2)	(3)	(4)
VARIABLES	PIN	PIN	PIN	PIN
Any Event	0.171***	0.173***	0.214***	0.215***
	(13.08)	(11.78)	(26.14)	(26.26)
Bad Event		0.00452	0.0106***	0.0103***
		(0.896)	(4.189)	(3.998)
Noise			-0.486***	-0.486***
			(-22.71)	(-22.72)
Alpha				0.00160*
				(1.883)
Constant	-0.00491	-0.00766	0.331***	0.332***
	(-0.890)	(-0.949)	(26.92)	(26.99)
Observations	2,102	2,102	2,102	2,102

Robust t-statistics in parentheses
*** p< 0.01, ** p< 0.05, * p< 0.1

Table 10: Robustness Test 2 – Quoted Spread

Table 10 presents the second robustness test based on the regression specifications (34) and (35). We use Fama-French-Carhart (FFC) alphas defined as the sum of all daily alphas during month $t-1$ for each stock i. Daily Bid-Ask spread is defined as the difference between closing ask and bid prices divided by closing ask price for stock i during each day of month $t-1$. The Quoted Spread variable is defined as the average of all daily Bid-Ask spreads during month $t-1$ for each stock i. The variable Year is a dummy, which captures year fixed effects. The table displays the average coefficients. (1) displays the regression using FFC alphas only, (2) displays the same regression when including the Quoted Spread. In (3) we include the year fixed effect. The sample period stretches from April 2005 to December 2014 and covers 117 months

	(1)	(2)	(3)
VARIABLES	ΔIV	ΔIV	ΔIV
Alpha	3.677***	3.578***	3.484***
	(4.99)	(5.01)	(4.89)
Quoted Spread		-3.649**	-3.361**
		(-2.26)	(-2.05)
Year			0.491
			(1.45)
Constant	-0.474	-0.267	-2.671
	(-1.04)	(-0.64)	(-1.51)
Observations	117	117	117

Robust t-statistics in parentheses
*** p< 0.01, ** p< 0.05, * p< 0.1

Table 11: Robustness Test 3 – Amihud

Table 11 presents the third robustness test based the regression specifications (36) and (37). We use Fama-French-Carhart (FFC) alphas defined as the sum of all daily alphas during month $t-1$ for each stock i. Daily Amihud is defined the absolute return for stock i during each day of month $t-1$ divided by its dollar traded volume during the same day. The Amihud variable is defined as the average of all daily Amihud measures during month $t-1$ for each stock i. The variable Year is a dummy, which captures year fixed effects. The table displays the average coefficients. (1) displays the regression using FFC alphas only, (2) displays the same regression when including the Amihud variable. In (3) we include the year fixed effect. The sample period stretches from April 2005 to December 2014 and covers 117 months

	(1)	(2)	(3)
VARIABLES	ΔIV	ΔIV	ΔIV
Alpha	3.677***	3.717***	3.615***
	(4.99)	(5.04)	(4.92)
Amihud		-7.213**	-7.053**
		(-2.02)	(-1.96)
Year			0.731**
			(2.38)
Constant	-0.474	-0.396	-3.925**
	(-1.04)	(-0.89)	(-2.38)
Observations	117	117	117

Robust t-statistics in parentheses
*** p< 0.01, ** p< 0.05, * p< 0.1

Table 12: Original Strategy – Monthly Alphas

Table 12 reports monthly alphas for the Original trading strategy, where we sort stocks into equally weighted portfolios based on Fama-French-Carhart (FFC) alphas during month $t-1$. MKTRF is the excess return for the market portfolio, SMB is the small-minus-big factor, HML is the high BM-minus-low BM factor and UMD is the momentum factor. Portfolio 10 consists of long stock positions with written call options on each stock, and Portfolio 1 consist of short stock positions with bought call options on each stock. (1) and (4) displays the CAPM and FFC alphas on Portfolio 1, while (2) and (5) displays the corresponding results for Portfolio 10. (3) and (6) show the combined alphas of the Original strategy. The sample period is April 2005 to December 2014 which is 117 months.

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Portfolio 1	Portfolio 10	P10+P1	Portfolio 1	Portfolio 10	P10+P1
MKTRF	-1.486***	0.0713***	-1.415***	-1.146***	0.102***	-1.044***
	(-20.50)	(4.245)	(-18.75)	(-16.23)	(5.479)	(-14.05)
SMB				-0.518***	-0.0669***	-0.585***
				(-4.307)	(-2.864)	(-4.745)
HML				0.105	-0.0608**	0.0443
				(1.016)	(-2.347)	(0.415)
UMD				0.643***	0.0121	0.655***
				(4.906)	(0.965)	(4.958)
ALPHA	0.000783	0.00586***	0.00664***	-0.000573	0.00575***	0.00518**
	(0.407)	(11.72)	(3.314)	(-0.287)	(11.48)	(2.497)

Robust t-statistics in parentheses
*** p< 0.01, ** p< 0.05, * p< 0.1

Table 13: Volume Strategy – Monthly Alphas

Table 13 reports monthly alphas for the Volume trading strategy, where we sort stocks into equally weighted portfolios based on Fama-French-Carhart (FFC) alphas during month $t-1$ and change in trading activity between month $t-1$ and $t-2$. MKTRF is the excess return for the market portfolio, SMB is the small-minus-big factor, HML is the high BM-minus-low BM factor and UMD is the momentum factor. Portfolio 10 consists of long stock positions with written call options on each stock, and Portfolio 1 consist of short stock positions with bought call options on each stock. (1) and (4) displays the CAPM and FFC alphas for Portfolio 1, while (2) and (5) displays the corresponding results for Portfolio 10. (3) and (6) show the combined alphas of the Volume strategy. The sample period is April 2005 to December 2014 which is 117 months.

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Portfolio 1	Portfolio 10	P10+P1	Portfolio 1	Portfolio 10	P10+P1
MKTRF	-1.895***	0.0637***	-1.831***	-1.410***	0.0784***	-1.332***
	(-16.69)	(3.988)	(-15.87)	(-13.40)	(4.035)	(-12.34)
SMB				-0.576***	-0.0260	-0.602***
				(-3.073)	(-0.799)	(-3.138)
HML				0.170	-0.0688*	0.101
				(1.166)	(-1.685)	(0.664)
UMD				1.062***	-0.0118	1.050***
				(4.852)	(-0.642)	(4.756)
ALPHA	0.00798***	0.00375***	0.0117***	0.00613**	0.00368***	0.00982***
	(2.716)	(6.087)	(3.880)	(2.045)	(5.939)	(3.182)

Robust t-statistics in parentheses
*** p< 0.01, ** p< 0.05, * p< 0.1

Table 14: Decomposed Volume Strategy – Monthly Alphas

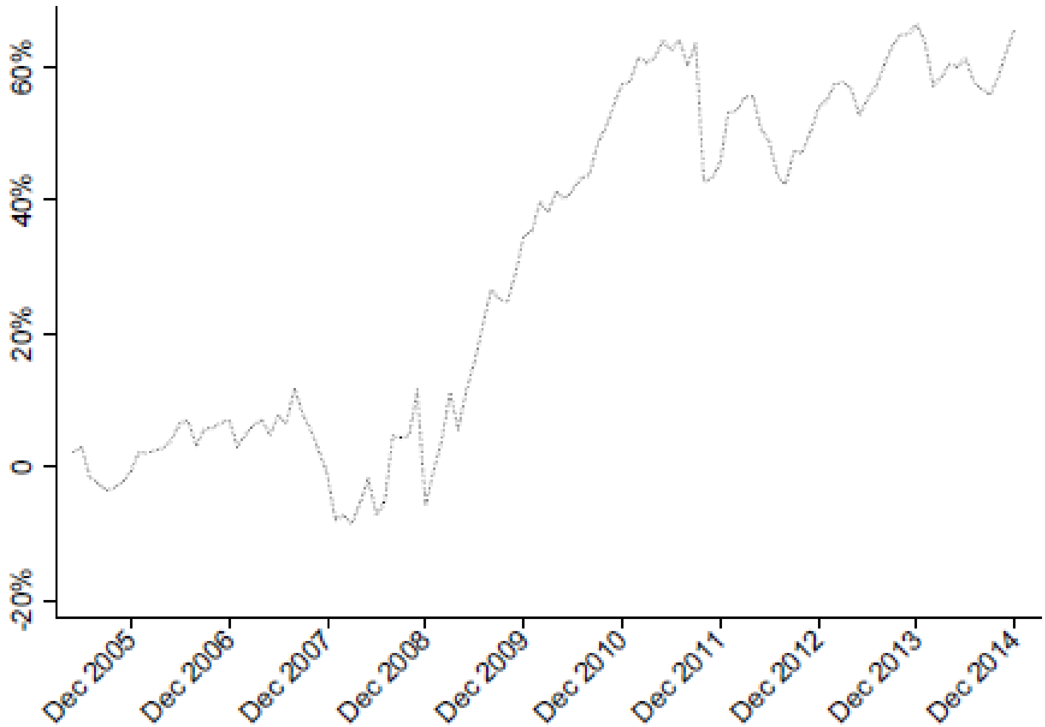
Table 14 reports monthly alphas for the Volume trading strategy, where we sort stocks into equally weighted portfolios based on Fama-French-Carhart (FFC) alphas during month $t-1$ and change in trading activity between month $t-1$ and $t-2$. MKTRF is the excess return for the market portfolio, SMB is the small-minus-big factor, HML is the high BM-minus-low BM factor and UMD is the momentum factor. (1) and (4) displays the decomposed CAPM and FFC exposures on excess stock returns. This is a variable that captures the combination of long stock positions from portfolio 10 and short stock positions from portfolio 1. (2) and (5) displays the corresponding factor exposures on the decomposed option positions. The variable captures the written call options from portfolio 10 and the bought call options from portfolio 1. (3) and (6) displays the CAPM and FFC alphas for the complete Volume strategy (as presented in Table 13) when we include VIX in the regression. The sample period is April 2005 to December 2014 which is 117 months

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Excess stock (P10+P1)	Excess option (P10+P1)	P10+P1	Excess stock (P10+P1)	Excess option (P10+P1)	P10+P1
MKTRF	-0.420***	3.113***	-1.898***	-0.241***	2.201***	-1.299***
	(-7.070)	(6.606)	(-11.90)	(-3.848)	(3.807)	(-8.499)
SMB				-0.209**	0.514	-0.603***
				(-2.420)	(0.501)	(-3.129)
HML				-0.0108	-0.515	0.0846
				(-0.117)	(-0.570)	(0.598)
UMD				0.354***	-2.425***	1.051***
				(4.435)	(-3.832)	(4.756)
VIX			-0.0206			0.00915
			(-0.811)			(0.384)
ALPHA	-0.000635	-0.00262	0.0127***	-0.00135	0.00185	0.00938**
	(-0.378)	(-0.131)	(3.590)	(-0.798)	(0.0923)	(2.576)

Robust t-statistics in parentheses
*** p< 0.01, ** p< 0.05, * p< 0.1

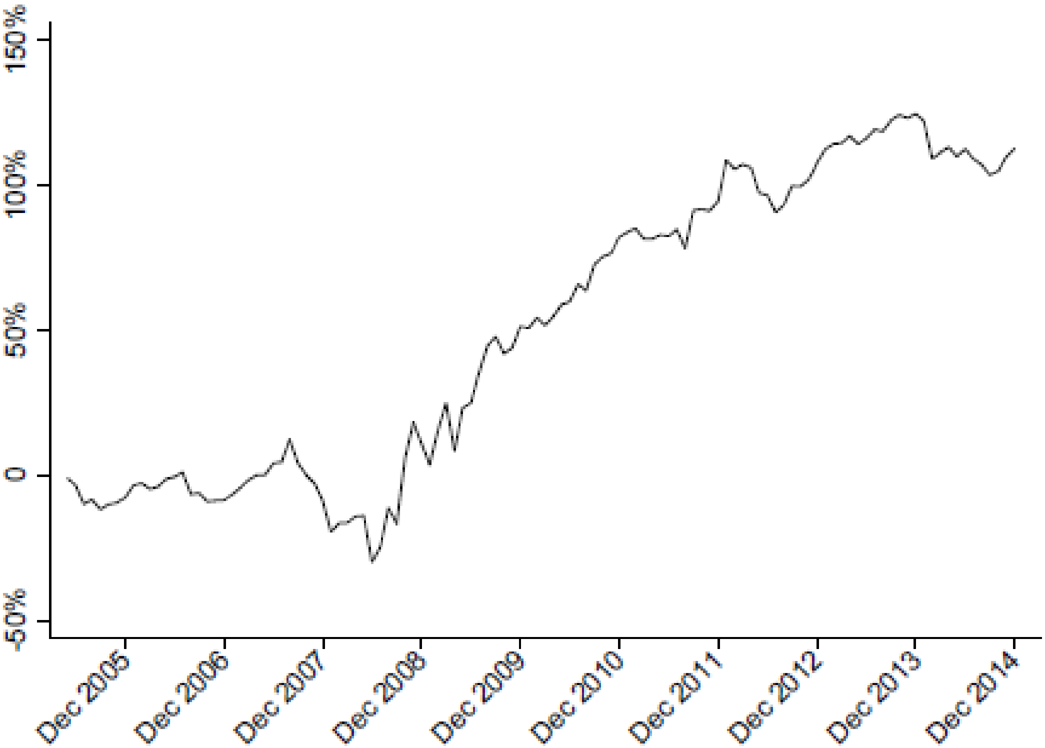
Graph 1: CARs of Original Strategy

Graph 1 presents the cumulative abnormal returns of the Original strategy, where we sort stocks into equally weighted portfolios based on past months Fama-French-Carhart alpha. Portfolio 1 consists of stocks with low or negative alpha and portfolio 10 of stocks with high alpha. The abnormal returns steam from the composition of long stock positions with written call options on each stock in portfolio 10, and of short stock positions with bought call options on each stock in portfolio 1. Abnormal returns are computed as the difference between the excess return on the Original strategy, minus the expected value from Fama-French-Carhart asset pricing model. The sample period stretches from April 2005 to December 2014 and covers 117 months.



Graph 2: CARs of Volume Strategy

Graph 2 presents the cumulative abnormal returns of the Volume strategy, where we sort stocks into equally weighted portfolios based the interaction term of past month Fama-French-Carhart alpha and changes in trading activity. Portfolio 1 consists of stock with low or negative alphas in combination with a high increase in trading activity. Portfolio 10 consists of stocks with high alpha and high decrease in trading activity. The abnormal returns steam from the composition of long stock positions with written call options on each stock in portfolio 10, and of short stock positions with bought call options on each stock in portfolio 1. Abnormal returns are computed as the difference between the excess return on the Volume strategy, minus the expected value from Fama-French-Carhart asset pricing model. The sample period stretches from April 2005 to December 2014 and covers 117 months.



Graph 3: Comparison of CARs between Original and Volume Strategy

Graph 3 presents the cumulative abnormal returns of the Original (dashed line) and Volume strategy (continuous line). In the Original strategy, we sort stocks into equally weighted portfolios based on past month Fama-French-Carhart alphas. Portfolio 1 consists of stocks with low or negative alpha and portfolio 10 of stocks with high alpha. In the Volume strategy, we sort stocks into equally weighted portfolios based the interaction term of past months alpha and changes in trading activity. Portfolio 1 consists of stock with low or negative alphas in combination with a high increase in trading activity. Portfolio 10 consists of stocks with high alpha and high decrease in trading activity. The abnormal returns steam from the composition of long stock positions with written call options on each stock in portfolio 10, and of short stock positions with bought call options on each stock in portfolio 1. Abnormal returns are computed as the difference between the excess return on the two strategies, minus the expected value from Fama-French-Carhart asset pricing model. The sample period stretches from April 2005 to December 2014 and covers 117 months.