The Journal of Alternative Investments

The Decline of Informed Trading in the Equity and Options Markets

Charles Cao, David Gempesaw and Timothy Simin

JAI 2018, 21 (2) 16-29
doi: https://doi.org/10.3905/jai.2018.1.064
http://jai.iijournals.com/content/21/2/16

The Decline of Informed Trading in the Equity and Options Markets

CHARLES CAO, DAVID GEMPESAW, AND TIMOTHY SIMIN

CHARLES CAO
is the Smeal Chair Professor of Finance in the Smeal College of Business at Pennsylvania State University in University Park, PA.
qxc2@psu.edu

DAVID GEMPESAW
is a doctoral student in the Smeal College of Business at Pennsylvania State University in University Park, PA.
dcg175@psu.edu

TIMOTHY SIMIN
is an associate professor of finance and Smeal Research Fellow in the Smeal College of Business at Pennsylvania State University in University Park, PA.
tts3@psu.edu

A substantial portion of the academic research on security markets has focused on identifying the presence of informed investors and evaluating the information content of their trading activity. Theoretically, informed investors possess private information about a stock's fundamental value and incorporate this information into asset prices, eliminating mispricing and improving market efficiency. This process can occur directly through trades in the equity market or the options market. In practice, the business of active portfolio management relies on trading with an informational advantage in order to generate excess returns.

The focus of this article is to test whether measures of informed trading in both the equity market and the options market provide useful information for forming alpha-generating portfolios. Specifically, we obtain measures of the information content of trading in both equities and equity options markets and then document the properties of these measures. We subsequently relate the trends in these measures to the excess returns of active portfolio managers. We specifically examine equity hedge fund excess returns because these are among the most active and informed traders on both the long and short side of trades and measures of hedge fund returns are widely available.

We find that spread portfolios of firms sorted on the previous period amount of informed trading in either the equity market or the options market yield risk-adjusted returns of roughly 9% per year. Using both measures of informed trading together to sort firms into portfolios increases the spread portfolio risk-adjusted returns, indicating that informed traders in these markets are basing trades on different information. When used to predict market returns, we find that aggregate informed trading in the equity market is not particularly useful, but aggregate informed trading in the options market is a stronger predictor than the dividend yield, arguably the most popular variable used for predicting market returns. We also demonstrate a consistent decline over time in the value of information content found in the measures of informed trading.

In this article, the information content of informed trading in the equity market is measured as net arbitrage trading, following Chen, Da, and Huang (2016). Net arbitrage trading is defined as the difference between abnormal hedge fund equity holdings and abnormal short interest on a stock. A positive value of net arbitrage trading is representative of a net increase in ownership by arbitrageurs relative to the average over the past four quarters. In the options market, informed trading is measured as the call–put implied volatility spread, following Bali and Hovakimian (2009) and Cremers and Weinbaum (2010). The difference between call and put implied volatility captures relative price pressure between calls and puts for a given stock. If informed traders trade in the options market in addition to the equity market, then violations of put–call parity reflected through the implied volatility spread may carry private information about future stock prices.

Consistent with the past studies on these measures of informed trading, we find that net arbitrage trading and the call–put implied volatility spread independently predict the future cross section of stock returns. Sorting stocks into equal-weighted quintile portfolios according to net arbitrage trading (call–put implied volatility spread) produces a difference in average returns between the extreme quintiles of 0.78% (0.91%) a month. In contrast to previous studies, we show that return predictability of informed trading in each market remains significant even after controlling for informed trading in the other market. Using an independent portfolio sorting approach, the predictive ability of each variable exists within each quintile portfolio formed on the other variable. These results indicate that the explanatory power of informed equity trading for the cross section of future returns is different from that of informed options trading. The average cross-sectional correlation between these two measures of informed trading is only 1.1%, further reinforcing the conclusion that they capture different types of information.

Using stock-level Fama and MacBeth (1973) return regressions, we show that the magnitude and statistical significance of the average slope coefficient estimates on net arbitrage trading and the call–put implied volatility spread are essentially the same in regressions excluding the other variable as they are in regressions where both variables are included. The information content of each variable also appears to be unrelated to a measure of mispricing proposed by Stambaugh, Yu, and Yuan (2015) that captures overpricing or underpricing across a number of well-known return anomalies. We find a stronger positive time-series correlation between aggregate measures of informed equity trading and informed options trading. However, only aggregate informed options trading has significant ability to predict future market returns.

Although our results suggest that incorporating measures of informed trading is useful for equity investors looking for an alternative investment strategy, a subsample analysis indicates that this usefulness has declined over time. Risk-adjusted returns of spread portfolios of stocks sorted by informed trading in the previous period fall to roughly 3% to 6% a year by the end of our sample period. The ability of the informed trading measures to explain the cross section of stock-level alpha has also declined. Notably, rolling alphas of the informed trading spread portfolios exhibit significant negative trends that are highly correlated with decreases in the alpha of equity hedge funds. The decrease in the ability of informed trading to generate alpha combined with the decline in hedge fund alphas are consistent with the argument that capacity constraints in the hedge fund industry are increasingly binding.

In summary, we demonstrate that the return predictive power of informed equity trading is different from that of informed options trading and that informed trading from both the equity and options markets is useful in portfolio formation, but this usefulness is declining over time. Although prior papers in this area focused on informed trading in a single market, we control for informed trading in other connected markets. We also identify differences between informed equity trading and informed options trading, specifically in the long-term predictive ability of the two measures, the average characteristics of the stocks targeted by informed traders in each market, and the ability of these measures to forecast future returns.

This article is organized as follows. The next section describes the data from the equity and options markets. The following section presents evidence of cross-sectional predictability using measures of informed equity and options trading jointly. The subsequent section investigates the cross-sectional differences between the two informed trading measures in their long-term predictive power and their relationship to stock characteristics. We then examine time-series predictability of market returns using both measures of informed trading and include a subsample analysis of our results. Concluding remarks are provided in the final section.

DATA

Since 1978, all institutional investors with at least \$100 million in assets under discretionary management are required by the SEC to report long positions in common stocks greater than 10,000 shares or \$200,000 in market value on a quarterly basis. Thomson Reuters classifies institutions into five categories (banks, insurance companies, investment management companies, investment advisors, others) and does not separately distinguish hedge funds from investment advisors or other types of institutional investors. To construct the measure of informed trading in the equity market, we use data from Thomson Reuters 13F institutional ownership database in conjunction with the sample of hedge fund companies constructed by Cao, Liang, Lo, and Petrasek (2018) and Cao and Petrasek (2014). To identify hedge fund ownership, the authors collected hedge fund company names from six hedge fund databases (TASS, HFR, CISDM, Barclay Hedge, Morningstar, and Bloomberg) and matched these names to the 13F data. Unmatched investment advisors and other institutions are manually checked to determine whether they are a hedge fund company. Each identified company is then manually checked to ensure that the company's primary business is hedge fund operation. The final sample consists of 1,517 hedge fund management companies managing more than 5,000 individual hedge funds from 1981 to 2012. Quarter-end short interest data are obtained from the Compustat Short Interest file. Because NASDAQ stocks are not included in the Compustat Short Interest file until after 2003, short interest data from Bloomberg are also used whenever the Compustat data are missing.

Using the institutional ownership data and short interest data, the following variables are constructed at a quarterly frequency, as in Chen, Da, and Huang (2016):

1. *Hedge fund ownership* (HF) is the ratio of shares owned by hedge funds to the number of shares outstanding. If a stock is not held by any hedge fund, HF is set to zero.
2. *Short interest* (SR) is the ratio of shares held short to the number of shares outstanding. If a stock is not included in the short interest data, SR is set to zero.
3. *Abnormal hedge fund holdings* (AHF) is the current value of HF minus the moving average of HF over the past four quarters.
4. *Abnormal short interest* (ASR) is the current value of SR minus the moving average of SR over the past four quarters.
5. *Net arbitrage trading* (AHFSR) is the difference between abnormal hedge fund holdings and abnormal short interest. We use net arbitrage trading as the measure of informed trading in the equity market.

Data on options are available beginning in 1996 from OptionMetrics. The OptionMetrics volatility surface file contains the interpolated volatility surface for a set of standardized options. The volatility surface incorporates information from listed options with various strikes and maturities and is calculated separately for calls and puts. The primary measure of informed trading activity in the options market is the call–put implied volatility spread (CPVOL), defined as the difference in call implied volatility (CVOL) and put implied volatility (PVOL) following Bali and Hovakimian (2009) and Cremers and Weinbaum (2010). We construct this measure at a monthly frequency from the volatility surface data using the month-end implied volatilities of calls and puts (deltas of 0.5 and −0.5, respectively) with 30 days to maturity. In addition to CPVOL, a number of other measures of informed trading activity in the options market has been proposed, including the put–call volume ratio (PCR), implied volatility innovations (ΔCVOL − ΔPVOL), implied volatility skew (VSKEW), and realized–implied volatility spread (RVOL−IVOL).[1] For robustness, we use these other variables as controls in cross-sectional regressions.

In addition to the main variables of interest, we obtain or construct other stock-level variables from CRSP and Compustat. Firm size is the market capitalization of equity at the end of the month in billions of dollars. Book-to-market ratio (B/M) is measured using the book value of equity in the latest fiscal year ending in the prior calendar year and the market value of equity at the end of December of the prior calendar year. Market beta is estimated using monthly regressions of excess daily returns on the excess market return, as well as the lag and lead values of the market return to account for nonsynchronous trading, and beta is the sum of the three estimated slope coefficients. Stock illiquidity is

[1] Following Pan and Poteshman (2006), the put–call volume ratio (PCR) is measured as total put volume divided by total option volume during the month. Following An, Ang, Bali, and Cakici (2014), we compute the difference between innovations in call implied volatility and innovations in put implied volatility (ΔCVOL − ΔPVOL). Following Xing, Zhang, and Zhao (2010), we measure the implied volatility skew (VSKEW) as the difference between implied volatility of out of the money puts (delta of −0.2) and at the money calls. Following Bali and Hovakimian (2009), the realized–implied volatility spread (RVOL–IVOL) is measured as the realized volatility of daily returns over the month minus the average of the call implied volatility and put implied volatility. Each variable is constructed at a monthly frequency.

measured following Amihud (2002) as the monthly average of the ratio of absolute daily return to dollar trading volume. Realized volatility (RVOL) is measured as the annualized standard deviation of daily returns over a given month. The momentum return is the cumulative 12-month return from month $t-12$ to month $t-1$. In our multivariate analysis, we also use a measure of mispricing constructed by Stambaugh, Yu, and Yuan (2015). This variable takes on values from 0 to 100 and represents the degree that a stock is over- or underpriced based on 11 documented return anomalies.[2] Risk-adjusted monthly returns (alphas) are measured using a regression of daily excess returns within a given month on the Fama and French (2015) five-factor model (FF5).[3] The monthly alpha is defined as the sum of the estimated intercept and residual across all days within the month.

The sample period is from January 1996 to December 2012 based on the intersection of the data on informed trading in each market. The final sample consists of NYSE, AMEX, and NASDAQ common stocks. At the end of each quarter, stocks with a share price of less than \$5 and market capitalization below the 20th percentile of NYSE firms are excluded from the sample. All explanatory variables are winsorized on both ends of the distribution at the 1% level to mitigate the influence of extreme observations in the data.

Exhibit 1 reports cross-sectional and time-series summary statistics for the main variables. First, we compute the mean and standard deviation for each variable over the cross section at the end of each month, and then report the time-series average of each statistic in Panel A. On average, across the monthly cross sections, the mean value of AHFSR (CPVOL) is 0.14% (−0.73%). The average stock has a market capitalization of \$7.26 billion, B/M of 0.51, and market beta of 1.15. The average number of stocks in the sample in a given month is 1,533. The average cross-sectional correlation between the informed trading measures (untabulated)

Exhibit 1

Summary Statistics for the Main Variables

Panel A: Cross-Sectional Statistics

	Mean	SD
AHFSR	0.14	3.32
CPVOL	–0.73	5.35
PCR	0.37	0.24
RVOL–IVOL	–1.73	13.23
ΔCVOL–ΔPVOL	0.01	7.00
VSKEW	6.05	7.52
Size	7.26	15.52
B/M	0.51	0.36
Beta	1.15	1.18
Illiquidity	0.34	0.54
RVOL	40.95	19.55
Return (%)	1.08	10.56
Momentum (%)	20.63	48.17
Mispricing	48.42	12.80
Number of Stocks	1,533	107

Panel B: Time-Series Statistics for Aggregate Measures of Informed Trading

	AR(1)	Mean	SD
AHFSR	0.76	0.18	0.40
CPVOL	0.28	–0.64	0.76
Pearson Correlation (AHFSR & CPVOL)		0.37	
No. of Quarterly Observations		68	

Notes: For each variable in Panel A, the mean and standard deviation are first computed over the cross section at the end of each month and then averaged across the monthly time series. Measures of informed trading include net arbitrage trading (AHFSR), call–put implied volatility spread (CPVOL), put–call volume ratio (PCR), realized–implied volatility spread (RVOL–IVOL), spread in call and put implied volatility innovations (ΔCVOL − ΔPVOL), and implied volatility skew (VSKEW). Other stock characteristics include market capitalization in billions of dollars (size), book-to-market ratio (B/M), market beta, Amihud (2002) illiquidity, realized volatility (RVOL), 1-month return, cumulative 12-month return (momentum), and the Stambaugh, Yu, and Yuan (2015) mispricing measure. The last line in Panel A reports the mean and standard deviation of the number of stocks in the sample each month. Panel B reports statistics for the time series of aggregate measures of informed trading (AHFSR and CPVOL). These aggregate measures are computed as the value-weighted average values of AHFSR and CPVOL, respectively, across all stocks at the end of each quarter. The panel includes the mean, standard deviation, and first-order quarterly autocorrelation coefficient for each aggregate measure, the time-series Pearson correlation coefficient between the two measures, and the number of quarterly observations. The sample period is from January 1996 to December 2012.

[2] We use the data on mispricing of individual stocks from Robert Stambaugh's website: http://finance.wharton.upenn.edu/~stambaug/. The 11 return anomaly variables are net stock issues, composite equity issues, accruals, net operating assets, asset growth, investment-to-assets, distress, O-score (probability of bankruptcy), momentum, gross profitability, and return on assets.

[3] Factor model data are obtained from Kenneth French's website: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

is positive, but it is economically insignificant at only 0.011.

In Panel B of Exhibit 1, we report statistics on aggregate measures of informed trading. Later in the article, we use these measures to investigate the time-series predictive power of informed trading in the equity and options markets. Each aggregate measure is computed as the value-weighted average across all stocks at the end of each quarter. Aggregate AHFSR (CPVOL) has a quarterly autocorrelation coefficient of 0.757 (0.282). The time-series correlation between these two measures is 0.368 and is significant at the 1% level. Combined with the cross-sectional correlation results, this suggests that informed equity trading and informed options trading are more strongly related over time than they are in the cross section.

CROSS-SECTIONAL PREDICTABILITY

Portfolio Sorting on Informed Trading Measures

We begin the empirical analyses by using a portfolio sorting approach to demonstrate the predictive power of the two measures of informed trading. At the end of each month, all stocks with available data are ranked based on their values of AHFSR from the most recent quarter-end and sorted into quintiles. We perform a similar quintile sorting procedure separately based on month-end values of CPVOL. For each quintile portfolio, we compute a monthly time series of equal-weighted average returns, as well as the return difference between the highest and lowest quintiles (5–1). We also adjust for the risk exposure of each portfolio using the Fama and French (2015) five-factor model.[4] Newey and West (1987) *t*-statistics are presented for the 5–1 portfolio with a maximum lag order of six months.

The univariate portfolio sorting results are presented in Exhibit 2. The first two columns report the results from sorting based on net arbitrage trading. On average, stocks with the lowest values of AHFSR have a monthly return of 0.636% during the next quarter, while stocks with the highest values of AHFSR have a monthly return of 1.413%. The average monthly return spread is 0.777%, which is statistically significant with a *t*-statistic of 5.96. After risk adjusting using the FF5 factor model, the alpha of each quintile decreases relative to the respective average raw return. However, the average spread in monthly risk-adjusted returns between the extreme quintiles remains significant. The 5–1 portfolio has a FF5 alpha of 0.705% with a *t*-statistic of 5.44. The last two columns report the results from sorting based on the call–put implied volatility spread. Moving from the lowest to the highest quintile, average monthly returns increase from 0.539% to 1.449%. The average difference in equal-weighted returns between extreme quintiles is 0.910% per month with a *t*-statistic of 5.49. This difference persists after risk adjustment. The 5–1 portfolio has a FF5 alpha of 0.894% with a *t*-statistic of 5.20. Computing value-weighted average returns within each portfolio leads to similar conclusions (untabulated). The average monthly returns and alphas

EXHIBIT 2

Cross-Sectional Predictability of Informed Trading in the Equity and Options Markets: Univariate Portfolio Sorting

	AHFSR		CPVOL	
Quintile	Return	FF5 Alpha	Return	FF5 Alpha
1	0.636	–0.227	0.539	−0.418
2	0.826	–0.067	0.832	−0.104
3	0.877	–0.066	0.989	0.112
4	1.127	0.115	1.060	0.150
5	1.413	0.478	1.449	0.476
5–1	0.777***	0.705***	0.910***	0.894***
t-Stat	(5.96)	(5.44)	(5.49)	(5.20)
Sharpe	0.439	0.414	0.539	0.542

Notes: At the end of each month, all stocks with available data are sorted into quintile portfolios based on a measure of informed trading. In the first two columns, stocks are sorted according to net arbitrage trading (AHFSR) from the most recent quarter-end. In the last two columns, stocks are sorted according to the spread between call and put implied volatilities (CPVOL). For each quintile portfolio, the exhibit reports the next month average raw return and risk-adjusted return (alpha) with respect to the Fama and French (2015) five-factor model. The row labeled "5–1" represents the difference in next-month average monthly raw return or risk-adjusted return between the highest and lowest quintile portfolios. Returns and alphas are reported in monthly percentage terms. Newey and West (1987) t*-statistics, significance levels, and Sharpe ratios are given for the 5–1 portfolio. The sample period is from January 1996 to December 2012.*

**** indicates significance at the 1% level.*

[4] While not tabulated, the following results are robust to the computation of value-weighted average portfolio returns as well as risk-adjustment using the CAPM, Fama and French (1993) three-factor model, or Carhart (1997) four-factor model.

EXHIBIT 3

Cross-Sectional Predictability of Informed Trading in the Equity and Options Markets: Independent Portfolio Sorting

	Fama and French (2015) Five-Factor Alpha						
	AHFSR(1)	**AHFSR(2)**	**AHFSR(3)**	**AHFSR(4)**	**AHFSR(5)**	**AHFSR(5–1)**	***t*-Stat**
CPVOL(1)	−0.821	−0.462	−0.493	−0.208	−0.076	0.745***	(3.48)
CPVOL(2)	−0.359	−0.429	−0.155	−0.121	0.526	0.885***	(5.29)
CPVOL(3)	0.130	−0.054	−0.142	0.195	0.527	0.397**	(2.08)
CPVOL(4)	−0.148	0.190	0.012	0.115	0.623	0.771***	(3.99)
CPVOL(5)	0.200	0.422	0.377	0.567	0.798	0.598**	(2.47)
CPVOL(5–1)	1.021***	0.883***	0.871***	0.776***	0.874***		
t-Stat	(4.11)	(3.59)	(4.35)	(3.91)	(3.77)		

Notes: At the end of each month, all stocks with available data are sorted independently into quintile portfolios based on net arbitrage trading (AHFSR) from the most recent quarter-end. Stocks are also sorted independently into quintile portfolios based on the call–put implied volatility spread (CPVOL). The exhibit reports the next-month Fama and French (2015) five-factor alphas for the 25 portfolios created by the intersection of the independent sorts. The row labeled "CPVOL(5–1)" represents the difference in risk-adjusted return between the highest and lowest CPVOL portfolios within each AHFSR portfolio. The column labeled "AHFSR(5–1)" represents the difference in risk-adjusted return between the highest and lowest AHFSR portfolios within each CPVOL portfolio. Alphas are reported in monthly percentage terms. Newey and West (1987) t*-statistics and significance levels are given for the 5–1 portfolios. The sample period is from January 1996 to December 2012.*

*** and *** indicate significance at the 5% and 1% levels, respectively.*

of the value-weighted portfolios are generally lower relative to those of the equal-weighted portfolios, indicating the impact of smaller stocks in the equal-weighted portfolios. Nevertheless, the average return spreads are comparable in magnitude and statistical significance. The results of univariate portfolio sorting are consistent with prior findings in the related literature and demonstrate the economically and statistically significant predictive power of each measure of informed trading. The Sharpe ratios of the 5–1 portfolios are presented at the bottom of each panel. Based on this measure, the 5–1 portfolio formed on CPVOL appears to have better performance than that of the 5–1 portfolio formed on AHFSR in general.

Next, we examine the predictive ability of each measure of informed trading through an independent sorting approach. This procedure can determine if and how the explanatory power of informed trading in one market varies across values of informed trading in the other market. For this test, stocks are sorted independently at the end of each month into quintiles based on CPVOL and the most recent quarter-end value of AHFSR. The intersection of these two independent sorts creates 25 portfolios of stocks. Exhibit 3 reports FF5 risk-adjusted returns for each of the portfolios.

We find that AHFSR has significant explanatory power within each CPVOL quintile. In the column of AHFSR(5–1) portfolio returns, the return spread is smallest within the middle CPVOL quintile, but all spreads are statistically significant at the 5% level or less. Looking at the row of CPVOL(5–1) portfolio returns, CPVOL also has significant explanatory power within each AHFSR quintile. All risk-adjusted spreads are significant at the 1% level. These results are qualitatively similar when computing value-weighted rather than equal-weighted average returns, analyzing average monthly returns unadjusted for risk, and performing sequential double sorting (e.g., sort into quintiles based on CPVOL, and then within each quintile, sort based on AHFSR to create 25 portfolios). Thus, we do not find evidence of a distinct relationship between the cross-sectional explanatory power of these two measures of informed trading.

Fama–MacBeth Cross-Sectional Regressions

In this section, we use Fama and MacBeth (1973) regressions to investigate the explanatory power of net arbitrage trading and the call–put implied volatility spread while simultaneously controlling for asset pricing model factor loadings and stock characteristics associated with the cross section of future stock returns. In the first stage, we estimate monthly cross-sectional regressions of realized returns in month $t + 1$ on values of AHFSR, CPVOL, and a vector of control variables ($\tilde{X}_t$) in month t. Specifically, the main cross-sectional model estimated at the end of each month is as follows:

$$r_{i,t+1} = a_t + b_{1,t} AHFSR_{i,t} + b_{2,t} CPVOL_{i,t} + b'_{3,t} \tilde{X}_{i,t} + e_{i,t+1} \quad (1)$$

In the second stage, we calculate the time-series average of the cross-sectional regression coefficients. Newey and West (1987) standard errors are computed with six lags.

Exhibit 4 presents cross-sectional regression results using the monthly FF5 alpha as the dependent variable. We first estimate regressions with only one measure of informed trading (and controls for size, B/M, market beta, illiquidity, realized volatility, return reversal, momentum, and mispricing) to illustrate how the coefficient estimates on AHFSR and CPVOL change when both are included together. In the presence of controls for a number of other characteristics, both measures of informed trading have statistically significant cross-sectional explanatory power when analyzed separately in Columns 1 and 2. In Column 3, the inclusion of both AHFSR and CPVOL simultaneously has almost no impact on the coefficient estimate and t-statistic of each variable. In terms of economic significance, one standard deviation increase in AHFSR and CPVOL increases the expected monthly risk-adjusted return by approximately 0.23% and 0.25%, respectively (using the coefficient estimates from Column 3 and the average cross-sectional standard deviations presented in Exhibit 1).

In untabulated analyses, we find that the estimates and corresponding t-statistics on the two measures of informed trading are generally robust to various specifications for the control variables. In particular, the results are qualitatively unchanged when the mispricing measure is excluded. One standard deviation increase in mispricing decreases the expected risk-adjusted return by approximately 0.18% per month using the coefficient estimates from Column 3. Therefore, even though the economic importance of this variable in explaining returns is comparable with that of AHFSR and CPVOL, the information conveyed in each market appears to be unrelated to this measure of mispricing. The coefficient estimate on mispricing, the lagged return, and firm size are statistically significant in all specifications.

Column 4 includes four additional measures of informed trading in the options market: the put–call ratio, realized–implied volatility spread, spread in call and put implied volatility innovations, and the implied volatility skew. Out of these five option market measures, only the coefficient estimates on VSKEW and CPVOL are statistically significant. Compared with the prior specifications, the coefficient on CPVOL is slightly diminished but still statistically significant at the 1% level. Importantly, even when all five measures of informed options trading are included in the regression, the coefficient estimate and t-statistic on AHFSR are relatively unaffected. The results suggest not only that the cross-sectional explanatory power of informed trading in each market cannot be explained by other stock characteristics or by a measure of mispricing, but also that the explanatory power of informed trading in the equity market is distinct from that of informed trading in the options market.[5]

INFORMED TRADING MEASURES IN THE CROSS SECTION

This section examines differences between the cross-sectional explanatory power of informed trading in the equity market and in the options market. Given that all of the prior analyses focus on returns over one month, we first investigate the return predictability of AHFSR and CPVOL over longer horizons. Exhibit 5 presents quarterly and cumulative returns of bivariate portfolios formed on AHFSR and CPVOL. Stocks are sorted at the end of each month into quintiles based on CPVOL. Then, within each CPVOL quintile, stocks are sorted at the end of each month based on the value of AHFSR from the most recent quarter-end. Each AHFSR subquintile is combined across CPVOL quintiles into a single quintile, resulting in quintile portfolios of stocks with differences in AHFSR but nearly identical

[5] In untabulated analyses, we find that these conclusions are robust to the use of monthly excess returns (unadjusted for risk) or Carhart (1997) four-factor alphas as the dependent variable as well as to the inclusion of controls for idiosyncratic return volatility, historical return skewness, stock turnover, total stock volume, total option volume, the call–put open interest ratio, stock price, total institutional ownership, and the degree of asymmetric information (measured as the probability of informed trading using data from Stephen Brown's website—http://scholar.rhsmith.umd.edu/sbrown/pin-data).

EXHIBIT 4

Fama–MacBeth Regressions of Risk-Adjusted Returns on Measures of Informed Trading

	Dependent Variable: Fama and French (2015) Five-Factor Alpha			
	(1)	**(2)**	**(3)**	**(4)**
AHFSR	0.071***		0.070***	0.069***
	(4.90)		(4.83)	(4.78)
CPVOL		0.047***	0.046***	0.035***
		(5.26)	(5.21)	(3.16)
PCR				–0.096
				(–0.78)
RVOL – IVOL				–0.005
				(–0.91)
ΔCVOL – ΔPVOL				–0.002
				(–0.19)
VSKEW				–0.016**
				(–2.20)
Mispricing	–0.014***	–0.014***	–0.014***	–0.014***
	(–3.20)	(–3.13)	(–3.12)	(–3.24)
Log(Size)	–0.089**	–0.084**	–0.093**	–0.071**
	(–2.39)	(–2.25)	(–2.47)	(–2.07)
Log(B/M)	0.008	0.012	0.004	0.027
	(0.12)	(0.19)	(0.05)	(0.41)
Beta	0.009	0.000	0.010	0.008
	(0.13)	(0.00)	(0.14)	(0.11)
Illiquidity	–0.170	–0.158	–0.170	–0.157
	(–0.38)	(–0.36)	(–0.39)	(–0.36)
RVOL	0.002	0.002	0.002	0.004
	(0.43)	(0.39)	(0.44)	(0.67)
Lag Return	–0.028***	–0.026***	–0.026***	–0.026***
	(–3.80)	(–3.55)	(–3.59)	(–3.55)
Momentum	0.001	0.002	0.001	0.001
	(0.48)	(0.54)	(0.49)	(0.51)
Adj. R^2	0.028	0.028	0.029	0.032

Notes: This exhibit presents results from Fama–MacBeth cross-sectional regressions. The dependent variable is the risk-adjusted return or alpha in month t + 1*. Monthly alphas are measured using a regression of daily excess returns within a given month on the Fama and French (2015) five-factor model. The monthly alpha is defined as the sum of the estimated intercept and residual across all days within the month. The primary explanatory variables are net arbitrage trading (AHFSR) and the call–put implied volatility spread (CPVOL). Control variables include the log of market capitalization in billions of dollars (Size), log of B/M, market beta (Beta), Amihud (2002) illiquidity, realized volatility (RVOL) in month t, realized return in month t (Lag Return), cumulative return from month* t − 12 *to month* t − 1 *(Momentum), the Stambaugh, Yu, and Yuan (2015) mispricing measure, put-call volume ratio (PCR), realized–implied volatility spread (RVOL–IVOL), spread in call and put implied volatility innovations (ΔCVOL − ΔPVOL), and the implied volatility skew (VSKEW). The average adjusted R^2 is reported in the last row. Newey and West (1987)* t*-statistics are given in parentheses. The sample period is from January 1996 to December 2012.*

*** and *** indicate significance at the 5% and 1% levels, respectively.*

distribution of values of CPVOL. For example, the lowest AHFSR quintile portfolio is composed of stocks with the lowest values of AHFSR within each of the five CPVOL ranked quintiles. This approach allows us to isolate the explanatory power of each informed trading measure.

Panel A for Exhibit 5 shows that stocks with the highest values of AHFSR outperform stocks with the lowest values of AHFSR in the following

Exhibit 5

Long-Term Return Predictability of Informed Trading in the Equity and Options Markets

	Quarterly Return				Cumulative Return		
Quintile	***t* + 1 to *t* + 3**	***t* + 4 to *t* + 6**	***t* + 7 to *t* + 9**	***t* + 10 to *t* + 12**	***t* + 1 to *t* + 6**	***t* + 1 to *t* + 9**	***t* + 1 to *t* + 12**
Panel A: Portfolios Formed on AHFSR Controlling for CPVOL							
1	2.323	2.751	3.355	3.573	4.915	8.061	8.022
2	2.692	3.073	3.193	3.397	5.630	8.707	8.612
3	2.768	2.953	3.174	3.212	5.625	8.666	8.500
4	3.456	3.355	3.497	3.490	6.743	10.219	10.100
5	3.965	3.675	3.757	3.638	7.381	11.115	10.980
5–1	1.642***	0.923***	0.402*	0.065	2.466***	3.055***	2.957***
t-Stat	(5.45)	(4.35)	(1.92)	(0.32)	(5.36)	(4.76)	(4.53)
Panel B: Portfolios Formed on CPVOL Controlling for AHFSR							
1	2.678	3.422	3.622	3.596	5.915	9.388	9.400
2	2.744	2.971	3.213	3.205	5.584	8.703	8.584
3	3.070	2.972	3.131	3.180	5.933	8.972	8.842
4	3.066	3.072	3.281	3.479	6.073	9.344	9.204
5	3.615	3.373	3.728	3.860	6.745	10.297	10.128
5–1	0.937***	−0.048	0.106	0.264	0.830**	0.909**	0.728
t-Stat	(4.84)	(–0.23)	(0.55)	(1.61)	(2.39)	(2.07)	(1.63)

Notes: At the end of each month, all stocks with available data are sequentially sorted into quintile portfolios based on measures of informed trading. In Panel A, stocks are sorted based on the call–put implied volatility spread (CPVOL). Then, within each CPVOL quintile, stocks are sorted into subquintiles based on net arbitrage trading (AHFSR) from the most recent quarter-end. Each AHFSR subquintile is combined across CPVOL quintiles into a single quintile. This approach creates quintile portfolios of stocks with differences in AHFSR but similar distributions of CPVOL. A similar procedure is performed in Panel B, except stocks are sorted first by AHFSR and then by CPVOL. For each quintile portfolio, the exhibit reports quarterly and cumulative returns over the next four quarters. The row labeled "5–1" represents the difference in return between the highest and lowest quintile portfolios. Returns are reported in monthly percentage terms. Newey and West (1987) t*-statistics and significance levels are given for the 5–1 portfolio. The sample period is from January 1996 to December 2012.*

, **, and * indicate significance at the 10%, 5%, and 1% levels, respectively.*

quarter. The return spread from month $t + 1$ to $t + 3$ is 1.642% and is statistically significant at the 1% level. The quarterly return spread decreases in the next three quarters but remains positive within each quarter and statistically significant through the third quarter when the return spread is 0.402% and significant at the 10% level. Thus, it appears that the information underlying informed trading in the equity market is slowly incorporated into prices over several months. The cumulative return spread also increases from 1.642% in the first quarter to 2.466% over the first two quarters and 3.055% over three quarters. This return spread remains significant for at least one year, indicating that the predictability is not due to temporary mispricing and does not reverse in the future.

We repeat the bivariate portfolio sorting procedure in Panel B of Exhibit 5 to create portfolios based on CPVOL while controlling for AHFSR. In Panel B, stocks with the highest values of CPVOL outperform stocks with the lowest values of CPVOL over the next three months. The return spread from month $t + 1$ to $t + 3$ is 0.937%. However, the quarterly return spread is no longer significant within each of the following three quarters. Although the cumulative return spread is statistically significant for up to nine months, it is smaller relative to the immediate three-month return. These results indicate that the information conveyed through the options market is reflected in stock prices relatively quickly. Thus, while AHFSR and CPVOL both contain information about the cross section of returns in the following month, the nature of the return predictive power of the two measures seems to differ over longer time horizons.

Finally, we use portfolio sorting to evaluate the patterns of average stock characteristics within quintiles in order to evaluate how these attributes are related to informed trading in each market. These results are left untabulated for brevity. We first perform a univariate sort based on one informed trading measure and examine the average value of the other informed trading measure within each quintile. When sorting stocks into quintiles according to AHFSR, each quintile appears to have similar values of CPVOL on average, and the distribution of average values is very small relative to the sample cross-sectional standard deviation of CPVOL. We reach similar conclusions when sorting stocks into quintiles according to CPVOL and evaluating average values of AHFSR within each quintile. These results further reinforce the inference that the measures of informed trading are not strongly related in the cross section, and they are consistent with the previous results demonstrating little overlap in the explanatory power of each measure for future returns.

Next, we evaluate the contemporaneous relationship between informed trading in one market and other stock characteristics by employing the bivariate portfolio sorting approach used in Exhibit 5 in order to control for the impact of informed trading activity in the other market. Informed equity and options demand exhibit similar relationships to some of the variables examined. For both AHFSR and CPVOL, the mispricing measure of Stambaugh, Yu, and Yuan (2015) follows a U-shaped pattern moving from the lowest to highest quintile. Because the mispricing measure is supposed to capture the extent of overpricing according to various return anomalies, the fact that the mispricing measure does not monotonically increase with either measure of informed trading would suggest that the return predictability of informed trading activity is not directly related to the return anomalies underlying the mispricing measure. This is consistent with our findings in Exhibit 4. Firm size and stock price follow inverted U-shaped patterns across each set of quintiles, and the realized volatility over the past month follows a U-shaped pattern across each set of quintiles. This indicates that more extreme values of AHFSR and CPVOL are associated with firms with lower market capitalization, lower stock price, and higher volatility.

Despite these similarities, the two measures of informed trading are related differently to certain stock characteristics. Stock illiquidity is lower on average within the extreme AHFSR quintiles, while the reverse is true for the CPVOL quintiles, indicating that informed traders prefer to use options when the stock is relatively illiquid. Monthly raw returns, risk-adjusted returns, and momentum returns follow U-shaped patterns across the informed equity demand quintiles, with the highest quintile having the highest values for each return type. This suggests that informed traders in the equity market take larger directional bets (both long and short) on stocks with better past performance. Across the informed options demand quintiles, however, the risk-adjusted returns and monthly returns all decrease monotonically with CPVOL, indicating a negative relationship between informed options trading and past monthly returns. Conversely, momentum follows a U-shaped pattern across the CPVOL quintiles, so this relationship between informed options trading and past returns appears to differ depending on the recency of the return information. There do not appear to be any distinct patterns for B/M or market beta across the portfolios. Finally, institutional ownership tends to be greater for stocks with extreme values of AHFSR, while the reverse is true for stocks with extreme values of CPVOL.

To summarize, while informed equity trading and informed options trading tend to be similarly related in the cross section to contemporaneous measures of mispricing, size, price, and volatility, the two measures exhibit significantly different relationships to institutional ownership and past performance. Stocks with the greatest demand from informed traders in the equity market tend to have better past performance in the cross section, while stocks with the greatest demand from informed traders in the options market tend to have worse past performance in the cross section.

TIME-SERIES PREDICTABILITY

This section explores the information content of informed equity trading and informed options trading in a time-series setting. Given the ability of the informed trading measures to predict cross-sectional differences in returns, we test whether the informed trading in the equity and options markets have any ability to predict equity returns through time. Specifically, we evaluate the ability of aggregate measures of informed demand in each of these markets to predict future returns of well-diversified portfolios. Because our measure of informed equity trading depends on holdings data that are updated quarterly, our time-series analyses are performed at a quarterly frequency. Aggregate measures of informed equity and informed options trading are computed as

EXHIBIT 6

Forecasting Quarterly Market Returns Using Aggregate Measures of Informed Trading

	Dependent Variable: Sample Average Return			
	(1)	(2)	(3)	(4)
AHFSR	3.322			2.698
	(0.91)			(0.79)
CPVOL		4.077***		3.771***
		(3.11)		(2.96)
DivYld			1.665*	0.932
			(1.83)	(1.52)
Adj. R^2	0.004	0.084	0.035	0.084

Notes: This exhibit presents results from time-series regressions of the excess market return in quarter t + 1 on aggregate measures of informed trading in the equity market (AHFSR), informed trading in the options market (CPVOL), and the annualized dividend yield in quarter t. The excess market return is measured as the value-weighted average quarterly excess return across all stocks in the sample. Aggregate informed equity and options trading are measured as the value-weighted averages of each variable across all stocks in the sample. The adjusted R^2 is reported in the last row. Newey and West (1987) t-*statistics are given in parentheses. The sample period is from January 1996 to December 2012.*

, and * indicate significance at the 10% and 1% levels, respectively.*

the value-weighted average of AHFSR and CPVOL, respectively, across all stocks in the sample at the end of each quarter.

In Exhibit 6, we use aggregate AHFSR and aggregate CPVOL to forecast excess market returns, measured as the value-weighted average quarterly return across all stocks in our sample. As an additional explanatory variable, we also include the S&P 500 Index dividend yield, a well-known predictor of market returns. We first estimate a regression of the quarterly excess market return on each aggregate measure of informed trading separately. Both aggregate measures of informed demand are positively related to the market return in the following quarter. However, the coefficient estimate for AHFSR is insignificant, whereas the estimates for CPVOL are significant at the 1% level. Consistent with the prior literature, we also find a positive and significant relationship between the lagged dividend yield and the quarterly market return. In the last column of each panel, we include all three of these explanatory variables simultaneously. Compared with the univariate regressions, the magnitude and *t*-statistic of each coefficient are smaller. Nevertheless, aggregate CPVOL continues to have significant predictive power at the 1% level, whereas the coefficients on aggregate AHFSR and the dividend yield are both insignificant. We find qualitatively similar results using the CRSP Value-Weighted Index and the S&P 500 as alternative market proxies. Our results indicate that the predictive power of informed options trading dominates that of informed equity trading and the dividend yield.

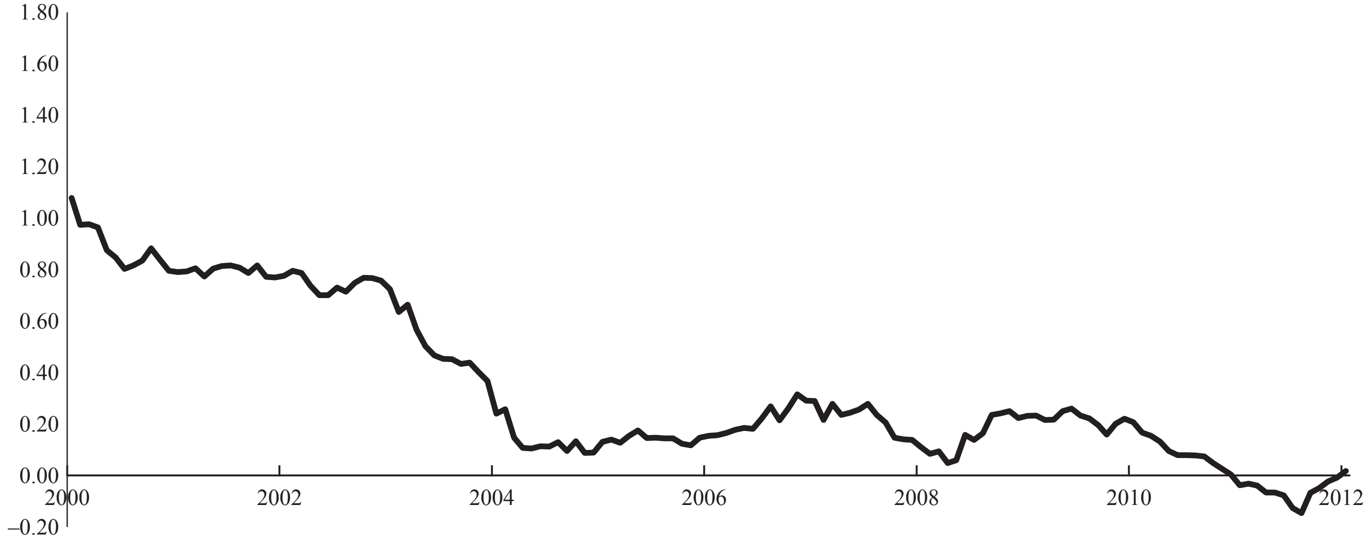
In untabulated analyses, we also examine the predictive power of the aggregate informed trading measures for the quarterly returns of quintile portfolios formed based on a measure of informed trading. This analysis can provide insight into whether the predictive power of aggregate informed trading differs across the various levels of demand from these traders. First, we sort stocks into quintiles at the end of each quarter based on AHFSR and compute the value-weighted quarterly return time series for each portfolio. Then, we regress each quintile portfolio excess return on aggregate informed options demand measured at the end of the prior quarter. The coefficient estimate is positive and significant at the 1% level in all cases. The magnitude of the coefficients is reduced somewhat as we move from the lowest to the highest quintile, but this pattern is not particularly strong. When we add aggregate informed equity demand as an additional control, the coefficient is positive but is much closer to zero and always statistically insignificant.

We repeat this analysis using quintile portfolios formed based on CPVOL. When we regress these portfolio excess returns on aggregate informed equity demand, the coefficient estimates are monotonically decreasing as we move from the lowest to the highest quintile, but none of these estimates are statistically significant. When aggregate CPVOL is included, we continue to find significant predictive power across all quintile portfolios. Overall, the results in this section suggest that the aggregate demand from informed traders in the options market can predict average returns both at the market level as well as the portfolio level (regardless of which informed trading measure is used to construct these portfolios). We find no evidence that the aggregate demand from informed equity traders has significant time-series predictive power.[6]

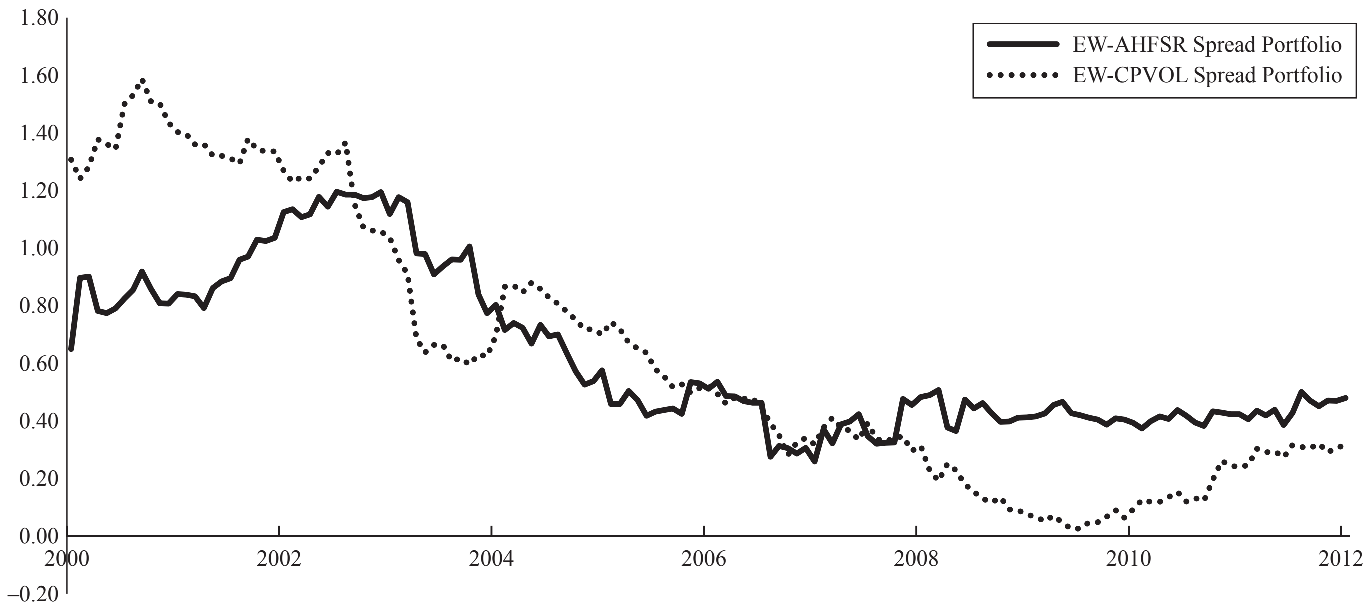
[6] In general, our conclusions from the time-series analyses are qualitatively similar when using equal-weighted averages rather than value-weighted averages, portfolio-level averages rather than aggregate averages, or monthly frequency rather than quarterly frequency.

Exhibit 7
Risk-Adjusted Returns

Panel A: Alpha of HFRI Equity Hedge Total Index



Panel B: Alphas of Informed Equity and Options Trading Portfolios



Notes: Panel A contains the rolling 60-month alpha from regressing the excess return of the HFRI Equity Hedge Total Index on the Fama and French (2015) five-factor model. Panel B contains the rolling 60-month FF5 alphas of the excess returns of the spread portfolios formed by sorting stocks based on informed equity trading or informed options trading. Excess returns are formed using the one-month T-bill rate.

THE VALUE OF INFORMED TRADING OVER TIME

Our previous results highlight the importance of informed trading for generating alpha over our sample period. In this section, we explore the time variation in the importance of informed trading. Panel A of Exhibit 7 contains 60-month rolling alphas for the excess returns of equity hedge funds over our sample period. The obvious downward trend in hedge fund alphas is highly correlated with the downward trend in the 60-month rolling FF5 alphas for the spread portfolios formed by sorting stocks by either measure of informed trading displayed in Panel B. The correlations of the informed equity and options spread portfolio alpha series with the equity hedge fund alphas are 0.79 and 0.85.

In Exhibit 2, the average returns over the full sample for the AHFSR and CPVOL spread portfolios are 0.78% and 0.91%, respectively. When we split the sample into the three periods of 1996 to 2000, 2001 to 2006, and 2007 to 2012, we find a declining pattern in the average returns of both portfolios. For the AHFSR spread portfolio, the returns decrease from 1.19% in the earliest period to 0.73% and then to 0.48% in the following periods. The CPVOL spread portfolio exhibits more dramatic declines, moving from 1.73% to 0.87% and then to 0.27% across the three subperiods. The decline in both raw and risk-adjusted returns of the spread portfolios indicates that the informed trading measures are producing less relevant information for picking stocks that generate a higher return.

Not only does the value of informed trading fall across time for generating raw or risk- adjusted returns, but the ability of informed trading to explain the cross section of alpha also diminishes through time. Column 3 of Exhibit 4 contains the full-sample average coefficients for AHFSR (0.070) and CPVOL (0.046) in the regression of firm-level alphas on the two measures of informed trading. Across the three subsamples mentioned, the coefficient on AHFSR falls from 0.088 to 0.067 and the coefficient on CPVOL falls from 0.066 to 0.016. The coefficients are significant in all subperiods, but the decrease across the three subperiods indicates that less of the cross-sectional differences in alphas can be explained by the informed trading measures in more recent years.

CONCLUSION

We investigate the informational content of informed equity trading and informed options trading in both a cross-sectional and a time-series setting. After documenting that net arbitrage trading in the equity market and the call–put implied volatility spread from the options market both have predictive power for cross-sectional variation in next month stock returns, we demonstrate that the explanatory power of each variable is essentially unaffected by controlling for the other variable, and that each variable has explanatory power across all values of the other variable. We find no evidence of an economically significant cross-sectional correlation between informed equity trading and informed options trading. Furthermore, we show that informed equity trading and informed options trading differ in their long-term predictive ability as well as their association with past performance. In time series, the correlation between informed trading in each market is more positive. However, informed options trading performs significantly better than informed equity trading when forecasting future market and portfolio excess returns.

These results illustrate important differences in the information conveyed through informed trading activity in the equity market and in the options market. Our subsample analysis indicates that although informed trading in the equity and options markets is significant across time, the value of the information content in informed trading has continually decreased. The correlation between the decline in hedge fund alphas and the decline in alphas of portfolios formed using informed trading measures is noteworthy, but we leave to future research the exploration of the causes of these declines and whether that correlation is causal.

ACKNOWLEDGMENT

We thank Gregory Brown, Jeremiah Green, and Stephen Lenkey for their valuable comments and suggestions.

REFERENCES

Amihud, Y. 2002. "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects." *Journal of Financial Markets* 5: 31–56.

An, B., A. Ang, T. Bali, and N. Cakici. 2014. "The Joint Cross Section of Stocks and Options." *The Journal of Finance* 69: 2279–2337.

Bali, T., and A. Hovakimian. 2009. "Volatility Spreads and Expected Stock Returns." *Management Science* 55: 1797–1812.

Cao, C., B. Liang, A. Lo, and L. Petrasek. 2018. "Hedge Fund Holdings and Stock Market Efficiency." *Review of Asset Pricing Studies* 8 (1): 77–116.

Cao, C., and L. Petrasek. 2014. "Liquidity Risk and Institutional Ownership." *Journal of Financial Markets* 21: 76–97.

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

Chen, Y., Z. Da, and D. Huang. 2016. "Arbitrage Trading: The Long and the Short of It." Working paper, Texas A&M University, University of Notre Dame, University of North Carolina at Greensboro.

Cremers, M., and D. Weinbaum. 2010. "Deviations from Put-Call Parity and Stock Return Predictability." *Journal of Financial and Quantitative Analysis* 45: 335–367.

Fama, E., and K. French. 1993. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics* 33: 3–56.

——. 2015. "A Five-Factor Asset Pricing Model." *Journal of Financial Economics* 116: 1–22.

Fama, E., and J. MacBeth. 1973. "Risk, Return, and Equilibrium: Empirical Tests." *Journal of Political Economy* 81: 607–636.

Newey, W., and K. West. 1987. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55: 703–708.

Pan, J., and A. Poteshman. 2006. "The Information in Option Volume for Future Stock Prices." *Review of Financial Studies* 19: 871–908.

Stambaugh, R., J. Yu, and Y. Yuan. 2015. "Arbitrage Asymmetry and the Idiosyncratic Volatility Puzzle." *The Journal of Finance* 70: 1903–1948.

Xing, Y., X. Zhang, and R. Zhao. 2010. "What Does the Individual Option Volatility Smirk Tell Us about Future Equity Returns?" *Journal of Financial and Quantitative Analysis* 45: 641–662.