Journal of Financial Economics 94 (2009) 310–326

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

# Cross-section of option returns and volatility ☆

Amit Goyal [a], Alessio Saretto [b,*]

[a] Goizueta Business School, Emory University, Atlanta, GA 30322, USA
[b] Krannert School of Management, Purdue University, West Lafayette, IN 47907, USA

A R T I C L E I N F O

*Article history:*
Received 10 March 2008
Received in revised form
29 September 2008
Accepted 22 January 2009
Available online 21 July 2009

*JEL classification:*
G11
G13
G14

*Keywords:*
Option returns
Historical volatility
Implied volatility
Overreaction

A B S T R A C T

We study the cross-section of stock option returns by sorting stocks on the difference between historical realized volatility and at-the-money implied volatility. We find that a zero-cost trading strategy that is long (short) in the portfolio with a large positive (negative) difference between these two volatility measures produces an economically and statistically significant average monthly return. The results are robust to different market conditions, to stock risks-characteristics, to various industry groupings, to option liquidity characteristics, and are not explained by usual risk factor models.



## 1. Introduction

Options allow an investor to trade on a view about the underlying security price and volatility. A successful option trading strategy must rely on a signal about at least one of these inputs. In the vernacular of option traders, at the heart of every volatility trade lies the trader's conviction that the market expectation about future volatility, which is implied by the option price, is somehow not correct.[1] Since all the option pricing models require at least an estimate of the parameters that characterize the probability distribution of future volatility, volatility misestimation is the most obvious source of options mispricing.

A common finding reported by studies of measurement and forecast of volatility is mean-reversion.[2] One forecast of volatility is embedded in the implied volatility (IV) of the stock, which can be obtained by inverting the Black

☆ We thank Tarun Chordia, Robert Dittmar, Darrell Duffie, Bernard Dumas, Andrea Frazzini, Laura Frieder, Robert Geske, Raffaella Giacomini, Mark Grinblatt, Narasimhan Jegadeesh, John McConnell, Richard Roll, Pedro Santa-Clara, Elena Sernova, Jay Shanken, Walter Torous, the seminar participants at Barclays Global Investors, Emory University, Erasmus University, HEC Lausanne, Purdue University, Rice University, Simon Fraser University, Tilburg University, UCLA, University of Illinois Urbana-Champaign, University of Kansas, University of Michigan, Virginia Tech, Swiss Finance Institute Conference on Portfolio Management and Derivatives, and especially Stephen Figlewski (the referee) for valuable suggestions. Alessio Saretto gratefully acknowledges the Purdue CIBER Center for financial support. The usual disclaimer applies.

* Corresponding author.
*E-mail address:* asaretto@purdue.edu (A. Saretto).

0304-405X/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jfineco.2009.01.001

[1] Options are also, of course, a very important risk management tool for hedging. A successful options trading strategy for these purposes can be implemented without any assumption that the market prices are incorrect.

[2] The volatility forecasting literature is extensive and too voluminous to cite in detail here. The interested reader is referred to the recent surveys in Granger and Poon (2003) and Andersen, Bollerslev, Christoffersen, and Diebold (2006).

and Scholes (1973) model.[3] IV from an option on a stock should, therefore, reflect the fact that future volatility will, on average, be closer to its long-run average historical volatility (HV) than to its current volatility. We do not suggest that IV should be the same as realized (historical or current) volatility. Our conjecture is only that large deviations of IV from HV are indicative of option mispricing.

Therefore, we sort stocks based on the difference between HV and IV. HV is calculated using the standard deviation of daily realized stock returns over the most recent 12 months and IV is computed by taking the average of the implied volatilities of the call and put contracts which are closest to at the money (ATM) and are one month to maturity. These selection criteria ensure that we construct a homogeneous sample with respect to the option contract characteristics across stocks, and that we consider the most liquid options contracts. Since we want to choose option positions which take advantage of volatility mispricing and have the least directional exposure to the underlying stocks, we compute returns of straddles and delta-hedged call portfolios.

We find that a zero-cost trading strategy involving a long position in an option portfolio of stocks with a large positive difference between HV and IV and a short position in an option portfolio of stocks with a large negative difference generates statistically and economically significant returns. For example, a long-short decile portfolio of straddles yields a monthly average return of 22.7% and a Sharpe ratio of 0.710. Similarly, we find statistically and economically significant positive returns for high decile portfolios and negative returns for low decile portfolios of delta-hedged calls.[4] These returns are comparable in magnitude to those in Coval and Shumway (2001), who report absolute returns of around 3% per week for zero-beta straddles on the Standard and Poor's 500 (S&P 500).

We then examine whether option returns to the long-short strategy are related to aggregate risk and/or characteristics. The alphas from standard risk-factor models with standard equity-risk and option-risk factors are very high and close to the raw returns.[5] Using cross-sectional regressions as well as via double sorted portfolios, we find that, while option returns covary with some of the stock characteristics that are found to be important for stock returns, this covariance is not enough to explain the high realized returns to our strategy.

Our results are robust to the choice of sample periods as well as volatility measures. Consistent with the literature on transaction costs on options markets, we find that trading frictions reduce the profitability of the option portfolio strategy. For instance, the long-short decile straddle portfolio returns are reduced to 3.9% per month if we consider trading options at an effective spread equal to the quoted spread.[6] We also find that the before-cost profits are higher for illiquid options than for liquid options. Finally, we show that margin requirements for short positions are roughly equal to one and a half times the cost of written options, which drives another wedge into the profitability of our strategies. Our analysis, therefore, shows that liquidity considerations reduce, but do not eliminate, the economically important profits of our portfolios.

To further understand the underlying reasons of the empirical regularity that we observe, we study the behavior of volatility (IV and HV) around the portfolio formation date. We find that the deviations between HV and IV are transitory—there is very little difference between the volatility measures one year before and after the portfolio formation, while the portfolio formation month, by construction, has large deviations. At the same time, in the month preceding portfolio formation the stocks in decile one have negative returns and the stocks in decile 10 have positive returns. In the months other than the portfolio formation month there is no appreciable pattern in stock returns. The fact that temporary deviations of IV from HV are accompanied by extreme patterns in stock returns suggests that investors overreact to current events, increasing (decreasing) their estimate of future volatility after large negative (positive) stock returns.

The conjecture that overreaction to current stock returns leads to misestimation of future volatility is consistent with the behavioral model of Barberis and Huang (2001) (henceforth BH).[7] Investors in the BH model have non-standard preferences with loss aversion and also engage in mental accounting. They get utility from gains and losses in wealth, rather than absolute levels of their wealth. In particular, they are more sensitive to losses than to gains (loss aversion). In addition, they apply this loss aversion to gains and losses defined narrowly over individual stocks (narrow framing in mental accounting). BH use their model to explain many empirical features of the data, such as excess volatility, time-series predictability, and cross-sectional value premium. The driving force for BH's results is the fact that the individual stock riskiness (discount rate) changes with the stock's past performance. If a stock has had good recent performance,

[3] Strictly speaking, IV is only a rough estimate of the market's risk-neutral estimate of future volatility of the underlying asset. Britten-Jones and Neuberger (2000) derive a procedure that gives the correct estimate of the option-implied (i.e., risk-neutral) integrated variance over the life of the option contract when prices are continuous but volatility is stochastic. Jiang and Tian (2005) improve upon this procedure and also show its validity in a jump-diffusion setting.

[4] We also sort stocks into two groups based on the sign of difference between HV and IV. Returns on these portfolios (also reported in the paper) are qualitatively similar to those for deciles.

[5] Although these regressions are linear factor models, we find that non-linear adjustments make virtually no difference. For instance, models with betas conditional on option greeks and models with Leland (1999) utility have very similar alphas as those from standard models.

[6] De Fontnouvelle, Fisher, and Harris (2003) and Mayhew (2002) show that typically the ratio of effective to quoted spread is less than 0.5. On the other hand, Battalio, Hatch, and Jennings (2004) study two periods in the later part of the sample, January 2000 and June 2002, and find that for a small sample of stocks the ratio of effective spread to quoted spread is around 0.8.

[7] While there are many models in behavioral finance that endogenously generate under- and overreaction, very few of them are concerned with volatility mispricing (since, by design, the objective is to explain empirical patterns in the first moments, viz. returns).

then the investors become less concerned about future losses on this stock, as future losses are cushioned by prior gains. This stock is, therefore, perceived to be less risky than before. Conversely, poor performance of a stock makes the investor more sensitive to the possibility of future losses, raising its riskiness.

Our empirical findings are in line with the predictions of the BH model. Stocks in decile one experience negative returns in the portfolio formation month, and have higher IV (perceived future volatility) than before portfolio formation; stocks in decile 10 have positive returns in the portfolio formation month, and have lower IV than before portfolio formation.

These expectations of changing volatility are, however, not justified by future events. We verify this by running a standard rationality test of forecasting future volatility using IV. We find that the forecast error in predicting future volatility using current IV is negative for decile one and is positive for decile 10. To summarize, investors overestimate future volatility of stocks following negative stock returns and under estimate future volatility following positive returns. This overreaction is precisely what leads to abnormal option returns. Our paper, thus, contributes to the debate about efficient markets, by identifying a trading strategy that produces abnormal returns and relating it to a behavioral model in the relatively unexplored area of options markets.

Our empirical tests and results are related to those of Black and Scholes (1972). Using data from an over-the-counter options market maker's book, they find that returns on options on high volatility (measured by HV) stocks are higher than those on low HV stocks. While Black and Scholes' motivation was to test the empirical validity of their pricing model, it is interesting to note the similarity between their and our results. On a more current note, our evidence of investor irrationality is analogous to Stein (1989). Stein studies the term structure of the implied volatility of index options and finds that investors overreact to the current information. They ignore the long-run mean reversion in implied volatility and instead overweight the current short-term implied volatility in their estimates of long-term implied volatility. Our findings are also consistent with those of Poteshman (2001), who finds that investors in the options market overreact, particularly to periods of increasing or decreasing changes in volatility. Finally, our paper is related to the growing recent literature that analyzes trading in options. Coval and Shumway (2001) and Bakshi and Kapadia (2003) study trading in index options. Chava and Tookes (2006), Ni, Pan, and Poteshman (2008), and Ni (2006) study the impact of news/information on trading in individual equity options. To the best of our knowledge, we are the first to study the economic impact of volatility mispricing through individual equity option trading strategies.

The rest of the paper is organized as follows. The next section discusses the data. Section 3 presents the main results of the paper by studying option portfolio strategies. Whether returns to option portfolios are related to fundamental risks and/or characteristics is investigated in Section 4. We discuss robustness checks as well as the impact of trading frictions on portfolio profitability in Section 5. Section 6 presents evidence consistent with investor overreaction. We conclude in Section 7.

## 2. Data

The data on options are from the OptionMetrics Ivy DB database. The data contain information on the entire U.S. equity option market and includes daily closing bid and ask quotes on American options as well as their IV and greeks (deltas, gammas, vegas) for the period from January 1996 through December 2006.[8] The IV and greeks are calculated using a binomial tree model using Cox, Ross, and Rubinstein (1979).[9]

We apply a series of data filters to minimize the impact of recording errors. First we eliminate prices that violate arbitrage bounds. Second we eliminate all observations for which the ask price is lower than the bid price, the bid price is equal to zero, or the bid–ask spread is lower than the minimum tick size (equal to \$0.05 for option trading below \$3 and \$0.10 in any other cases). Finally, following Driessen, Maenhout, and Vilkov (2009), we remove all observations for which the option open interest is equal to zero, in order to eliminate options with no liquidity.

We construct portfolios of options and their underlying stocks. These portfolios are formed based on information available on the first trading day (usually a Monday) immediately following the expiration Saturday of the month (all the options expire on the Saturday immediately following the third Friday of the expiration month). To have a continuous time-series with constant maturity, we consider only options that mature in the next month. Among these options with one month to maturity, we then select the contracts which are closest to ATM. Since it is not always possible to select options with moneyness (defined as the ratio of strike to stock price) exactly equal to one, we keep options with moneyness between 0.975 and 1.025. Thus, for each stock and for each month in the sample, we select the call and the put contract that are closest to ATM and expire next month. After next month expiration, we select a new pair of call and put contracts that are at that time closest to ATM and have one month to expiration. Our final sample includes 4,344 stocks and is composed of 75,627 monthly pairs of call and put contracts.

For each stock and for each month in the sample, we calculate two different measures of volatility: historical volatility (HV) and implied volatility (IV). HV is calculated using the standard deviation of realized daily stock

[8] Battalio and Schultz (2006) note that option and underlying prices are recorded at different times in the Ivy DB database, creating problems when an arbitrage relation such as the put–call parity is examined. This property of the data is not an issue for us because the tests that we conduct do not require perfectly coordinated trading data in the two markets.

[9] Interest rates used by the Ivy DB option models are derived from British Banker's Association LIBOR rates and settlement prices of Chicago Mercantile Exchange Eurodollar futures. The current dividend yield is assumed to remain constant over the life of the option and the security is assumed to pay dividends at specific predetermined times. We refer the reader to the Ivy DB reference manual for further details.

**Table 1**
Summary statistics.

The data on options are from the OptionMetrics Ivy DB database. We select one call and one put for each stock in each month of the sample period. All options have expirations of one month and moneyness closest to one. We first compute the time-series average of these volatilities for each stock and then report the cross-sectional average of these average volatilities. The other statistics are computed in a similar fashion. For each month and for each stock, historical volatility (HV) is calculated using the standard deviation of daily realized stock returns over the most recent 12 months. For each month and for each stock, implied volatility (IV) is the average of the implied volatilities of the call and put contracts that are closer to being at the money and have one month to maturity. The volatilities are annualized. The sample includes 4,344 stocks and is composed of 75,627 monthly pairs of call and put contracts. The sample period is 1996 to 2006.

| | Mean | Median | StDev | Min | Max | Skew | Kurt |
|---|---|---|---|---|---|---|---|
| IV | 0.544 | 0.528 | 0.118 | 0.397 | 0.771 | 0.482 | 2.712 |
| HV | 0.563 | 0.554 | 0.101 | 0.437 | 0.731 | 0.288 | 2.269 |
| ΔIV | −0.012 | −0.013 | 0.140 | −0.217 | 0.197 | 0.045 | 2.343 |
| ΔHV | −0.001 | −0.001 | 0.018 | −0.028 | 0.026 | 0.008 | 3.173 |

returns over the most recent 12 months. IV is computed by taking the average of the ATM call and put implied volatilities. We compute the time-series average of these volatilities for each stock and then report the cross-sectional average of these average volatilities in Table 1. The other statistics (median, standard deviation, minimum, maximum, skewness, and kurtosis) are computed in a similar fashion so that the numbers reported in the table are the cross-sectional averages of the time-series statistics and can be interpreted as the summary statistics of an average stock.

HV and IV are close to each other, with values of 56.3% and 54.4%, respectively. Similar to the finding of Driessen, Maenhout, and Vilkov (2009), HV is slightly higher than IV by about 2% for individual stocks. The overall distribution of IV is, however, more volatile and more positively skewed than that of HV. The average monthly change in both measures of volatility is very close to zero. Changes in IV can be quite drastic and usually correspond to events of critical importance for the survival of a firm. For example, UICI, a health insurance company, has a ΔIV of 86% which corresponds to the release of a particularly negative quarterly loss for the fourth quarter of 1999. During the month of December, UICI options went from trading at an ATM IV of 31% to an IV of 117%. The stock price lost 56% of its value in the same month. Many of the other large spikes in volatility happen during months of large declines in stock prices. For example, the IV of the stocks in the technology sector jumped over 150% during the burst of the Nasdaq bubble in the spring of 2000. Spikes in individual stock IV also happen on earnings announcements (Dubinsky and Johannes, 2005).

## 3. Option portfolio strategies

All option pricing models require, at least, an estimate of the parameters that characterize the probability distribution of future volatility. It is well-known that volatility is highly mean-reverting—the average autocorrelation for individual stock volatility in our sample is 0.7. This implies that large deviations of current volatility from its long-term average are temporary in nature and are likely to reduce in magnitude at a quick rate (determined by the mean-reversion parameter). Any forecast of future volatility must account for this mean-reversion. One such forecast is embedded in the IV of the stock. Therefore, IV from an option on a stock should reflect the fact that future volatility will, on average, be closer to its HV than to its current volatility.

Note that we are not suggesting that mean-reversion implies that IV should be the same as realized (historical or current) volatility. Indeed the stochastic nature of volatility and the existence of a volatility risk premium necessarily results in differences between IV and HV.[10] However, high autocorrelation of volatility implies that large deviations between IV and HV are unlikely to persist. Therefore, we speculate that, if there is volatility mispricing, it is more likely to manifest itself in extreme temporary deviations between HV and IV. Stocks for which IV is much lower than HV have cheap options, and stocks for which IV is much higher than HV have expensive options.

To further motivate volatility trading strategy based on both HV and IV, we conduct a standard rationality test of predicting future realized volatility from HV and IV. In particular, following Christensen and Prabhala (1998), for each month in the sample we run the following cross-sectional regression[11]:

$$fv_{i,t+1} = \alpha_t + \beta_{1t} iv_{i,t} + \beta_{2t} hv_{i,t} + \varepsilon_{i,t+1}, \tag{1}$$

where $fv$ is the (log of) the future realized volatility over the life of the option, where the future volatility is calculated as the standard deviation of daily returns, and $iv$ and $hv$ are the logs of IV and HV, respectively. The underlying hypothesis is that, if IV is an unbiased forecast of FV, then the parameters $\alpha$, $\beta_1$, and $\beta_2$ should be equal to zero, one, and zero, respectively.

We run regression (1) each month and calculate the time-series average of the regression coefficients, together with their standard error, corrected for autocorrelation in time-series following Newey and West (1987). We find that $\beta_1$ ($\beta_2$) is 0.967 (0.877) in a univariate regression, with very high statistical significance. In a bivariate regression, $\beta_1$ and $\beta_2$ are 0.695 and 0.286, respectively ($t$-statistics are 21.8 and 10.2, respectively).[12] Furthermore, out-of-sample mean square error and mean absolute square error decrease, from 0.032 to 0.030 and

[10] Deviations of IV from HV will be more pronounced for stocks with higher volatility of volatility than for stocks with lower volatility of volatility. Stocks with high prices of volatility risk (positive or negative) will also exhibit differences between IV and (future) realized volatility: IV is only a risk-neutral expectation of future volatility.

[11] The results from a time-series regression for each stock are very similar to the ones reported here. However, it is a non-trivial exercise to calculate the standard error of the cross-sectional average of regression coefficients.

[12] Christensen and Prabhala (1998) point out that measurement error in IV introduces a bias in estimated coefficients. We obtain similar results by instrumenting current IV with lagged IV and with lagged measures of one-month and 12-month realized volatility.

from 0.110 to 0.105, respectively, when HV is included in the regression. Taken together these pieces of evidence suggest that predicting future volatility using HV and IV reduces the forecast error by approximately 5% (versus a prediction using only IV). Thus, both IV and HV contain valuable information about future volatility, and IV does not subsume the information contained in HV. These results hint at the possibility that a trading strategy based on HV (in addition to IV) might be profitable.

### 3.1. Portfolio formation

We construct two types of portfolios. First, we sort stocks into deciles based on the log difference between HV and IV. Decile one consists of stocks with the lowest (negative) difference while decile 10 consists of stocks with the highest (positive) difference between these two volatility measures. The decile portfolios are equal-weighted. Second, we sort stocks into two groups depending on the sign of the difference between HV and IV. We label these groups P (for positive, HV higher than IV) and N (for negative, HV lower than IV). These two groups are relative value-weighted—weights in each of the two groups are proportional to the (absolute) deviation between HV and IV. On average, the equal-weighted decile portfolios contain 53 stocks in each month and the relative-weighted P (N) portfolio contains 279 (252) stocks.

Descriptive statistics of the equal-weighted decile portfolios and the relative value-weighted P and N portfolios are reported in Table 2. These are calculated by first computing averages (equal-or value-weighted depending on the portfolio) across stocks for each month in the sample and for each portfolio. This procedure gives us a continuous time-series of portfolio statistics. The table then reports the time-series averages of these portfolio statistics.

While HV increases as one proceeds from decile one to decile 10, IV follows the opposite pattern generally decreasing from decile one to 10. We note that the spread in HV between portfolio one and 10 is much larger than that in IV. This shows that our sort is not just on the levels on IV but on richer dynamics of the difference between HV and IV.

There is not much variation (not accounted for by differences in IV and underlying prices) in option greeks across portfolios. For instance, deltas of calls in all deciles are close to 0.54 while the deltas of puts in all deciles are close to −0.47. The gammas (second derivative with respect to underlying price) and vegas (first derivative with respect to volatility) are of similar magnitude across deciles. The P and N portfolios also have similar options greeks.

The difference HV–IV between P and N portfolios ($15\% = 8.2\% - (-6.8\%)$) is less than half as much as that between decile 10 and one portfolios ($34.5\% = 19.7\% - (-14.8\%)$). This implies that the extreme decile portfolios have options with larger deviations of HV and IV than do the P and N portfolios.

In unreported results, we find that there is no industry over-concentration in any of our portfolios. We do find that stocks in our sample are typically large and belong to the top two deciles of market capitalization by NYSE breakpoints. For example, the average market capitalization of a stock in decile one (10) is \$8.1 (\$8.9) billion, although there is no pattern in size across any of the portfolios. Finally, we find that the skewness and kurtosis of stock returns is higher for higher numbered deciles than it is for lower numbered deciles.

### 3.2. Portfolio returns

Since our interest is in studying returns on options based only on their volatility characteristics, we want to neutralize the impact of movements in the underlying stocks as much as possible. We accomplish this task by forming straddle portfolios and delta-hedged call portfolios (results for delta-hedged put portfolios are very similar to those for delta-hedged calls, and are omitted for brevity). Straddles are formed as a combination of one

**Table 2**
Formation-period statistics of portfolios sorted on the difference between HV and IV.

HV and IV are calculated as in Table 1. We sort stocks into portfolios based on the difference between historical volatility (HV) and implied volatility (IV). Portfolios 1 through 10 are obtained by sorting stocks into deciles based on the log difference between HV and IV. These 10 portfolios are equally weighted. Portfolios N and P are obtained by sorting stocks into two groups depending on the sign of the difference between HV and IV. We label these groups P, for positive difference, and N, for negative difference. These two portfolios are relative value-weighted—weights in each of the two groups are proportional to the (absolute) deviation between HV and IV. $\Delta$, $\Gamma$, and $\mathscr{V}$ are the delta, gamma, and vega, respectively, of the options. All statistics are first averaged across stocks in each portfolio and then averaged across time. The sample includes 4,344 stocks and is composed of 75,627 monthly pairs of call and put contracts. The sample period is 1996 to 2006.

| | Decile portfolios | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | N | P |
| $HV_t - IV_t$ | −0.148 | −0.076 | −0.046 | −0.023 | −0.004 | 0.016 | 0.036 | 0.060 | 0.097 | 0.197 | −0.068 | 0.082 |
| $HV_t$ | 0.389 | 0.410 | 0.421 | 0.429 | 0.446 | 0.463 | 0.485 | 0.509 | 0.550 | 0.651 | 0.419 | 0.532 |
| $IV_t$ | 0.537 | 0.486 | 0.467 | 0.452 | 0.450 | 0.447 | 0.449 | 0.449 | 0.453 | 0.455 | 0.487 | 0.450 |
| $\Delta^c$ | 0.544 | 0.541 | 0.540 | 0.537 | 0.536 | 0.534 | 0.532 | 0.528 | 0.531 | 0.531 | 0.540 | 0.532 |
| $\Delta^p$ | −0.460 | −0.464 | −0.465 | −0.468 | −0.469 | −0.471 | −0.475 | −0.478 | −0.474 | −0.476 | −0.465 | −0.474 |
| $\Gamma$ | 0.133 | 0.124 | 0.124 | 0.125 | 0.128 | 0.128 | 0.133 | 0.135 | 0.142 | 0.167 | 0.126 | 0.142 |
| $\mathscr{V}$ | 3.967 | 4.168 | 4.330 | 4.350 | 4.263 | 4.316 | 4.160 | 4.189 | 3.997 | 3.727 | 4.110 | 4.110 |

**Table 3**
Post-formation returns of portfolios sorted on the difference between HV and IV.

Portfolios are formed as in Table 2. The returns on options are constructed using, as a reference beginning price, the average of the closing bid and ask quotes and, as the closing price, the terminal payoff of the option depending on the stock price and the strike price of the option. The hedge ratio for the delta-hedged calls is calculated using the current IV estimate. The options monthly returns are equal-weighted (for deciles) or value-weighted (for P and N portfolios) across all the stocks in the portfolio. SR is the Sharpe ratio and CE is the certainty equivalent. CE is computed from a utility function with constant relative risk-aversion parameters of three and seven. The sample includes 4,344 stocks and is composed of 75,627 monthly pairs of call and put contracts. The sample period is 1996 to 2006.

| | Decile portfolios | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 10–1 | N | P | P–N |
| Panel A: Straddle returns | | | | | | | | | | | | | | |
| mean | −0.128 | −0.078 | −0.057 | −0.033 | −0.018 | −0.010 | −0.002 | 0.007 | 0.027 | 0.099 | 0.227 | −0.094 | 0.044 | 0.138 |
| std | 0.182 | 0.178 | 0.220 | 0.225 | 0.227 | 0.255 | 0.227 | 0.233 | 0.239 | 0.291 | 0.251 | 0.188 | 0.241 | 0.189 |
| min | −0.572 | −0.411 | −0.414 | −0.406 | −0.353 | −0.486 | −0.540 | −0.440 | −0.368 | −0.321 | −0.271 | −0.400 | −0.347 | −0.572 |
| max | 0.724 | 0.667 | 0.832 | 1.129 | 0.877 | 1.413 | 1.197 | 1.187 | 1.133 | 1.501 | 1.492 | 0.829 | 1.194 | 1.016 |
| SR | −0.722 | −0.455 | −0.272 | −0.162 | −0.092 | −0.051 | −0.020 | 0.017 | 0.101 | 0.329 | 0.903 | −0.518 | 0.168 | 0.730 |
| CE ($\gamma = 3$) | −0.177 | −0.125 | −0.120 | −0.092 | −0.082 | −0.081 | −0.061 | −0.056 | −0.037 | 0.017 | 0.156 | −0.138 | −0.016 | 0.080 |
| CE ($\gamma = 7$) | −0.245 | −0.183 | −0.183 | −0.149 | −0.144 | −0.155 | −0.152 | −0.123 | −0.102 | −0.058 | 0.069 | −0.181 | −0.075 | −0.109 |
| Panel B: Delta-hedged call returns | | | | | | | | | | | | | | |
| mean | −0.017 | −0.011 | −0.008 | −0.006 | −0.004 | −0.002 | −0.002 | 0.001 | 0.003 | 0.010 | 0.027 | −0.013 | 0.005 | 0.018 |
| std | 0.024 | 0.022 | 0.025 | 0.024 | 0.024 | 0.027 | 0.025 | 0.027 | 0.030 | 0.034 | 0.033 | 0.025 | 0.028 | 0.027 |
| min | −0.088 | −0.073 | −0.079 | −0.062 | −0.053 | −0.056 | −0.047 | −0.048 | −0.059 | −0.050 | −0.045 | −0.077 | −0.042 | −0.086 |
| max | 0.097 | 0.067 | 0.103 | 0.122 | 0.099 | 0.142 | 0.120 | 0.147 | 0.142 | 0.172 | 0.194 | 0.132 | 0.131 | 0.141 |
| SR | −0.856 | −0.628 | −0.458 | −0.399 | −0.297 | −0.198 | −0.200 | −0.090 | 0.007 | 0.195 | 0.809 | −0.660 | 0.063 | 0.674 |
| CE ($\gamma = 3$) | −0.018 | −0.012 | −0.009 | −0.007 | −0.005 | −0.003 | −0.003 | −0.000 | 0.002 | 0.008 | 0.025 | −0.014 | 0.004 | 0.017 |
| CE ($\gamma = 7$) | −0.019 | −0.013 | −0.010 | −0.008 | −0.006 | −0.005 | −0.004 | −0.002 | 0.001 | 0.006 | 0.024 | −0.015 | 0.002 | 0.016 |

call and one put with the same underlying, strike price, and maturity.[13] Delta-hedged call positions are obtained by buying one call contract and short-selling delta shares of the underlying stock. Stock prices are taken from the Center for Research in Security Prices (CRSP) database. We use the delta (based on the current IV) provided by the Ivy DB database.[14,15]

For each stock and for each month in the sample we select a call and a put contract that is approximately ATM and has one month to maturity. We then construct time-series of straddle and delta-hedged call returns. The straddle returns are constructed using, as a reference beginning price, the sum of the average of the closing bid and ask quotes of the call and put, and, as the closing price, the terminal payoff of the option that expires in the money. The terminal payoff depends on the stock price at expiration and the strike price of the option. The delta-hedged call returns are constructed using, as a reference beginning price, the average of the closing bid and ask quotes of the call minus delta shares of the stock valued at the closing stock price on the strategy initiation day, and, as the closing price, the terminal payoff of the call minus delta shares of the stock valued at the closing stock price on the day the option expires.[16]

To ameliorate microstructure biases, we start trading a day after the day on which we obtain the signal (the difference between HV and IV). Specifically, we form portfolios on the first trading day (typically a Monday) after the expiration Friday of the month and we initiate option portfolio strategies on the second trading day (typically a Tuesday) after the expiration Friday of the month. The returns are equal-weighted for decile portfolios and relative value-weighted for P and N portfolios, as described earlier.

Table 3 reports summary statistics of the returns of the 10 decile portfolios, P and N portfolios, as well as of two spread portfolios: 10–1 portfolio formed by taking a long

[13] In addition to the simple straddle returns, we also considered zero-delta and zero-beta straddles. Zero-delta straddles were formed using the delta provided by the Ivy DB database, while zero-beta straddles were constructed following the procedure in Coval and Shumway (2001). The returns on these portfolios were very similar to the ones reported in the paper for the plain vanilla straddles.

[14] If there is volatility mispricing in options, a more powerful and profitable approach is to recalculate delta based on an implied volatility estimate. However, we do not attempt to estimate a new delta because we do not have an alternative estimate of implied volatility (only a signal that IV is higher/lower than HV). Green and Figlewski (1999) note that a delta-hedged strategy based on incorrect delta entails risk and does not provide a riskless rate of return. This means that we are conservative in our construction of delta-hedged portfolios—we earn lower returns and have higher risk.

[15] Our delta-hedged portfolios are held until expiration and not rebalanced during the holding period (similar to the straddle portfolios). This is a conservative approach as our buy-and-hold strategy entails higher risk than a frequently/daily rebalanced strategy. Note, however, that rebalancing the portfolio would involve transaction costs of trading the underlying stock to adjust the delta.

[16] The options are American. However, we ignore the possibility of early exercise in our analysis for simplicity. Optimal early exercise decisions would bias our results downwards for the long positions in portfolios and upwards for the short positions in portfolios. The net effect is not clear. See Poteshman and Serbin (2003) for a discussion of early exercise behavior.

position in the options in decile 10 and by writing the options in decile one; and P–N portfolio formed by taking a long position in the options in portfolio P and by writing the options in portfolio N.

In addition to standard descriptive statistics, we also calculate two measures related to the risk-return trade-off: Sharpe ratio (SR) and certainty equivalent (CE). SR is the most commonly used measure of risk-return trade-off; nonetheless, CE is potentially a better measure than SR because it takes into account all the moments of the return distribution. We compute CE for a long position in the portfolio and we use a power utility function with a coefficient of relative risk aversion, $\gamma$, equal to three and seven.

Panel A of Table 3 shows that straddle decile portfolios exhibit a striking pattern with average returns that go from $-12.8\%$ to 9.9%, respectively. The volatility of these returns is also low at between 18% and 30% per month. The 10–1 straddle strategy has an average return of 22.7% with a 25.1% monthly standard deviation (the minimum monthly return in the sample is $-32.1\%$), leading to a monthly SR of 0.903 and a CE($\gamma = 3$) of 15.6% per month. The P–N straddle portfolio has lower average return of 13.8% but lower volatility at 18.9% than the 10–1 portfolio. P–N portfolio has lower returns than 10–1 portfolio, because the range of the sorting variable (HV–IV) is higher for the 10–1 portfolio than it is for the P–N portfolio, as described in the previous section. To put all these numbers in perspective, the value-weighted CRSP portfolio has a monthly SR of 0.131 and a monthly CE of 0.582% ($\gamma = 3$) for our sample period.

The magnitude of returns for delta-hedged calls in Panel B is lower than that for straddles. This is to be expected for two reasons. First, straddles benefit from mispricing of both calls and puts, while delta-hedged calls benefit only from mispricing of calls. Second, delta-hedged call position involves part of the portfolio weight in stock, for which we make no claim of mispricing. In spite of this, we see that our sorting criterion still lends itself to positive returns for high decile portfolios and negative returns for low decile portfolios. The average return of the delta-hedged 10–1 portfolio is 2.7% with standard deviations of 3.3%. The low standard deviation of these portfolios leads to high SR (0.809), and the absence of huge positive and negative returns also leads to positive CE (2.5%).

Note that these option returns are not driven by directional exposure to the underlying asset. When underlying stocks are sorted according to the same portfolio classification, the returns of the stock portfolios in the month after portfolio formation exhibit no pattern across deciles, or between P and N portfolios.

Note also, that the spread in option returns across portfolios is not simply a result of the fact that stocks in decile one (10) have high (low) IV. It is true that stocks in decile one have high IV and low HV, and stocks in decile 10 have low IV and high HV (see Table 2). However, we find that the average difference in the return of the extreme decile portfolios obtained by sorting only on the levels of IV or only on the levels of HV is often economically small and not statistically significant.

## 4. Controls for risk and characteristics

Our next task is to establish whether the large portfolio returns are systematic compensation for risk or abnormal. Since options are derivative securities, it is reasonable to assume that option returns depend, at least, on the same sources of risks or characteristics that explain individual stock returns. The absence of a general formal theoretical model for the cross-section of option returns, however, makes our endeavor non-trivial.[17] We approach our problem from several different perspectives. We start by running factor-model regressions with the standard equity-risk factors augmented with risk factors for options. We then explore whether stock/option characteristics are related to the variation in our portfolio returns. This second analysis is done on individual options via cross-sectional regressions, as well as via double sorted portfolios.

We acknowledge that we (like others) are subject to a joint hypothesis problem—the estimated alphas are derived from models and, therefore, rejection of the null of zero alpha is a joint rejection of zero alpha and the model. Our hope is that these experiments taken together lend credence to our belief that the portfolio returns from the previous section are not related to obvious sources of risk and characteristics.

### 4.1. Risk-adjusted returns

We regress the returns of spread (10–1 and P–N) straddle and delta-hedged call portfolios on various specifications of a linear pricing model consisting of the Fama and French (1993) three factors, the Carhart (1997) momentum factor, and an aggregate volatility factor. For the straddle portfolio, the aggregate volatility factor is the excess return of a zero-beta S&P 500 index ATM straddle, as in Coval and Shumway (2001). For the delta-hedged call portfolio, the aggregate volatility factor is the excess return of a delta-hedged S&P 500 index call.[18] Since all the factors are spread traded portfolios, the intercept from the following regression can be interpreted in the usual sense of mispricing relative to the factor model:

$$R_{pt} = \alpha_p + \beta_p' F_t + e_{pt}, \tag{2}$$

where $R_p$ is the return on portfolio and $F$'s are factors. Although, any linear factor model is unlikely to characterize the cross-section of option returns over any discrete time interval, we use it merely to illustrate that the option returns described in this paper are not related to aggregate sources of risk in an obvious way.

[17] The only notable exception is Duarte and Jones (2007). We calculate alphas based on their model and find that these are very close to the raw returns. Additional details are available upon request.

[18] We obtain data on the first four factors from Ken French's Web site, while we construct the volatility factors ourselves following the procedure described in Coval and Shumway (2001). During our sample period, the average return of the zero-beta S&P 500 index ATM straddle is $-11.50\%$ per month while the average return of the delta-hedged S&P 500 index call is $-0.46\%$.

**Table 4**
Risk-adjusted option returns.

Portfolios are formed as in Table 2. We present results from the following time-series regression of 10–1 and P–N portfolio returns:

$$R_{pt} = \alpha_p + \beta_p F_t + e_{pt}.$$

The risk factors are the Fama and French (1993) three factors (MKT–Rf, SMB, and HML), the Carhart (1997) momentum factor (MOM), and the Coval and Shumway (2001) excess zero-beta S&P 500 straddle factor (ZB-STRAD–Rf). DH-CALL–Rf is the S&P 500 delta-hedged call factor return. The first row gives the coefficients while the second row gives the *t*-statistics in parentheses. The sample includes 4,344 stocks and is composed of 75,627 monthly pairs of call and put contracts. The sample period is 1996 to 2006.

| | Straddles | | | | Delta-hedged calls | | | |
|---|---|---|---|---|---|---|---|---|
| | 10–1 | | P–N | | 10–1 | | P–N | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Alpha | 0.242 | 0.244 | 0.148 | 0.150 | 0.026 | 0.025 | 0.017 | 0.016 |
| | (9.96) | (9.73) | (8.03) | (7.76) | (8.22) | (7.85) | (6.39) | (5.61) |
| MKT–Rf | −0.357 | −0.578 | −0.312 | −0.418 | 0.031 | −0.020 | 0.031 | 0.002 |
| | (−0.51) | (−0.85) | (−0.60) | (−0.80) | (0.29) | (−0.21) | (0.37) | (0.03) |
| SMB | | −0.460 | | −0.946 | | 0.081 | | 0.029 |
| | | (−0.53) | | (−1.76) | | (0.72) | | (0.31) |
| HML | | −0.952 | | −0.799 | | −0.214 | | −0.107 |
| | | (−1.05) | | (−1.36) | | (−1.51) | | (−1.02) |
| MOM | | 0.378 | | 0.285 | | 0.098 | | 0.032 |
| | | (0.59) | | (0.60) | | (0.99) | | (0.38) |
| ZB-STRAD–Rf | 0.110 | 0.099 | 0.070 | 0.058 | | | | |
| | (2.83) | (2.99) | (2.25) | (2.11) | | | | |
| DH-CALL–Rf | | | | | 0.233 | 0.205 | 0.098 | 0.086 |
| | | | | | (1.82) | (1.89) | (0.89) | (0.87) |
| $\bar{R}^2$ | 0.074 | 0.070 | 0.053 | 0.067 | 0.032 | 0.075 | −0.002 | −0.009 |

Estimated parameters for these factor regressions are reported in Table 4. The market model regressions show that all spread portfolios have insignificant loadings on the market factor. The other regressions show that the loadings on Fama and French and momentum factors are also statistically insignificant. More interesting is the fact that the spread straddle portfolios load positively on the zero-beta straddle portfolio (regressions (3) and (4)). This implies that our portfolios earn abnormal returns (positive alphas) even with positive exposure to volatility risk. Regressions (5)–(8) show a similar pattern for delta-hedged calls, although the loadings on the volatility factor are significant only at the 10% level for only the 10–1 portfolio. The magnitude of the alphas in all regressions is very close to that of the average raw returns.[19]

We also make efforts to ameliorate the problem associated with linear factor models in two ways. First, we estimate the following factor-model regressions with conditional betas:

$$R_{pt} = \alpha_p + (\beta_{0p} + \beta_{1p}' \Theta_{pt-1})' F_t + e_{pt}, \tag{3}$$

where $\Theta$'s are either portfolio option greeks (delta, gamma, and vega), portfolio implied volatility, or aggregate implied volatility measured by the Chicago Board Options Exchange Volatility Index, VIX. Conditional betas are used to proxy for the time-variation (over the life of the option) in expected returns of options. The alphas from this model are very similar to those reported in Table 4.

Second, we estimate Leland (1999) alpha. Leland proposes a correction to the linear factor models that allows the computation of a robust risk measure for assets with arbitrary return distributions. This measure is based on the equilibrium model of Rubinstein (1976) in which a CRRA (constant relative risk aversion) investor holds the market. Our estimates of Leland's alpha are also very close to the ones reported in Table 4. For instance, Leland alpha for the 10–1 straddle portfolio is 23.2%.

We have remarked earlier that the levels of skewness and kurtosis in stock returns are higher for decile 10 than those for decile one. Are there differences in the sensitivity of portfolio returns to the *risk* of, rather than

[19] We also tried to investigate whether liquidity risk factor can explain the option returns. However, we were unable to reconstruct the Pástor and Stambaugh (2003) or Sadka (2006) liquidity factors to exactly match our holding period (from Tuesday of the fourth week to the third Friday of the next month). When we used the misaligned factors in factor-model regressions, we found that the loadings on these liquidity factors were not significant.

the *levels* of, higher moments? To explore this, we first run Eq. (2) including factors for skewness and kurtosis, such as the square and the cube of the market return, for each portfolio.[20] We then take the betas from these time-series regressions and run a second-stage cross-sectional regression:

$$\overline{R_p} - \widehat{\beta_{1p}}' \overline{F_1} = \widehat{\beta_{2p}}' \lambda_2 + \alpha_p, \tag{4}$$

where the bars denote the time-series sample averages, $\widehat{\beta_1}$'s are loadings on traded factors, $\widehat{\beta_2}$'s are loadings on non-traded factors, $\lambda_2$'s are the prices of risks of non-traded factors, and the residuals from the second-stage regression are the pricing errors. As a practical matter, we stack the first- and the second-stage regressions together in a generalized method of moments (GMM) framework that allows us to account for the estimation error in betas in calculating the standard errors from the second-stage regression (see Cochrane, 2001, for details on this procedure). We find that the betas on the non-traded factors are insignificant and the prices of risk ($\lambda_2$) of these factors are also insignificant. Most important from our perspective, we find that alphas from this experiment are very close to the ones in Table 4.

### 4.2. Stock characteristics

We now investigate how the long-short straddle portfolio returns are related to equity characteristics. We first run cross-sectional regressions of risk-adjusted individual option returns on lagged characteristics. Specifically, our regressions specification is similar to that in Brennan, Chordia, and Subrahmanyam (1998):

$$R_{it} - \widehat{\beta}_i' F_t = \gamma_{0t} + \gamma_{1t}' Z_{it-1} + e_{it}, \tag{5}$$

where $R$ is the return on individual options (in excess of risk-free rate), $F$'s are factors, and $Z$'s are characteristics. The $\widehat{\beta}$'s on the left-hand side of the equation are estimated via a first-pass time-series regression using the entire sample. The factors are the same as in Section 4.1. Besides the primary variable of interest (HV–IV), the other characteristics chosen are: size, book-to-market, past six-month return, skewness, and kurtosis. All characteristics are lagged by one month in these regressions.

We run these regressions every month and report the time-series averages of $\gamma$ coefficients and their $t$-statistics, corrected for autocorrelation following Newey and West (1987) in Table 5. Consistent with results in prior sections, the difference between HV and IV is strongly statistically significant in explaining the pattern of subsequent returns. Kurtosis (and to some extent, skewness) is the only stock characteristic that seems to have some predictive power for option returns, albeit with relatively small economic magnitude.

We also investigated a bigger set of stock/option variables by including proportion of systematic risk ($R^2$), dispersion in analyst forecasts, and credit rating of the company's bonds. We select $R^2$ since Duan and Wei (2009) find that systematic risk proportion is useful for cross-sectionally explaining the prices of equity options, analyst dispersion because of the evidence in Diether, Malloy, and Scherbina (2002), and the credit rating to check if our option returns are related to default risk. Finally, two option characteristics (gamma and vega) are chosen to reflect information that is not directly contained in equities. The unreported results of these extensive experiments show that, first, none of these variables has predictive power for option returns, and, second, the difference between HV and IV remains strongly significant in each case.

To provide yet another perspective of whether characteristics subsume our effect, we consider two-way sorts—one based on the volatility signal (HV–IV) and the

**Table 5**
Option returns controlling for stock characteristics (cross-sectional regressions).

Options are selected as in Table 1. We estimate the following cross-sectional regression for individual option returns:

$$R_{it} - \widehat{\beta}_i' F_t = \gamma_{0t} + \gamma_{1t}' Z_{it-1} + e_{it},$$

where the $\widehat{\beta}$'s are calculated using a first-pass regression using the entire sample, $F$'s are the factors, and $Z$'s are characteristics. The factors used in risk-adjustment are the Fama and French (1993) factors, momentum factor, and an option factor. The option factors are ZB-STRAD–Rf and DH-CALL–Rf for straddles and delta-hedged calls, respectively. The characteristics are $hv - iv$ (log difference of HV and IV), Size (market capitalization), BtoM (book-to-market), Mom (last six-month cumulative stock return), and skewness and kurtosis of stock returns (calculated using the last year's daily data). The table reports the averages of $\gamma$ coefficients and the associated $t$-statistics, corrected for autocorrelation following Newey and West (1987) in parentheses. The last row gives the average $\overline{R}^2$ from the monthly regressions. The sample includes 4,344 stocks and is composed of 75,627 monthly pairs of call and put contracts. The sample period is 1996 to 2006.

| | Straddles | | | Delta-hedged calls |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| const | −0.006 | 0.044 | 0.096 | 0.004 |
| | (−0.50) | (0.83) | (1.83) | (0.70) |
| $hv - iv$ | 0.225 | 0.224 | 0.258 | 0.029 |
| | (8.89) | (8.51) | (9.48) | (9.35) |
| Size | | −0.004 | −0.006 | −0.001 |
| | | (−1.04) | (−1.72) | (−1.47) |
| BtoM | | 0.003 | 0.003 | −0.000 |
| | | (0.13) | (0.11) | (−0.06) |
| Mom | | 0.003 | 0.005 | 0.001 |
| | | (0.21) | (0.35) | (0.66) |
| Skew | | | −0.007 | −0.001 |
| | | | (−1.87) | (−1.15) |
| Kurt | | | −0.002 | −0.000 |
| | | | (−3.95) | (−3.66) |
| $\overline{R}^2$ | 0.006 | 0.018 | 0.020 | 0.023 |

[20] See Kraus and Litzenberger (1976), Friend and Westerfield (1980), and Harvey and Siddique (2000) for studies relating skewness to stock returns.

**Table 6**
Option returns controlling for stock characteristics (double-sorted portfolios).

Each month, we first sort stocks into quintiles based on stock characteristics and then, within each quintile, we sort stocks based on the difference between the historical HV and the current IV. The second sort is a sort into quintiles or into P/N portfolios, as in Table 2. The volatility portfolios are then averaged over each of the five characteristic portfolios. Beta is the stock beta calculated from the market model using last 60 months, Size is the market capitalization, BtoM is the book-to-market, Mom is the last six-month cumulative return, and skewness and kurtosis of stock returns are calculated using the last year's daily data. Breakpoints for all stock characteristics are calculated each month based only on stocks in our sample. The table reports the average return and the associated *t*-statistics, in parentheses, of this continuous time-series of monthly returns. The sample includes 4,344 stocks and is composed of 75,627 monthly pairs of call and put contracts. The sample period is 1996 to 2006.

| | Quintile portfolios | | | | | | |
|---|---|---|---|---|---|---|---|
| Control | 1 | 2 | 3 | 4 | 5 | 5–1 | P–N |
| Panel A: Straddle returns | | | | | | | |
| Beta | −0.108 | −0.050 | −0.019 | −0.007 | 0.043 | 0.151 | 0.087 |
| | (−7.34) | (−2.66) | (−1.00) | (−0.35) | (2.03) | (8.89) | (7.00) |
| Size | −0.101 | −0.048 | −0.018 | 0.008 | 0.069 | 0.170 | 0.095 |
| | (−6.94) | (−2.77) | (−0.96) | (0.44) | (3.15) | (9.88) | (7.82) |
| BtoM | −0.113 | −0.041 | −0.019 | −0.000 | 0.064 | 0.177 | 0.103 |
| | (−7.64) | (−2.32) | (−0.91) | (−0.02) | (2.94) | (9.96) | (7.73) |
| Mom | −0.101 | −0.040 | −0.019 | −0.002 | 0.066 | 0.166 | 0.101 |
| | (−6.82) | (−2.35) | (−0.98) | (−0.11) | (3.09) | (10.22) | (8.54) |
| Skew | −0.103 | −0.048 | −0.005 | 0.001 | 0.059 | 0.162 | 0.101 |
| | (−7.31) | (−2.64) | (−0.29) | (0.05) | (2.78) | (10.39) | (8.23) |
| Kurt | −0.105 | −0.050 | −0.010 | 0.008 | 0.061 | 0.166 | 0.093 |
| | (−7.53) | (−2.86) | (−0.52) | (0.39) | (2.86) | (10.61) | (7.56) |
| Panel B: Delta-hedged call returns | | | | | | | |
| Beta | −0.014 | −0.008 | −0.003 | −0.001 | 0.003 | 0.017 | 0.011 |
| | (−8.16) | (−4.11) | (−1.71) | (−0.67) | (1.39) | (9.51) | (7.57) |
| Size | −0.015 | −0.007 | −0.005 | 0.000 | 0.007 | 0.022 | 0.012 |
| | (−7.85) | (−3.51) | (−2.29) | (0.09) | (2.64) | (8.88) | (7.39) |
| BtoM | −0.015 | −0.006 | −0.004 | −0.002 | 0.007 | 0.022 | 0.013 |
| | (−8.38) | (−3.18) | (−1.75) | (−1.02) | (2.85) | (9.99) | (7.54) |
| Mom | −0.013 | −0.007 | −0.003 | −0.001 | 0.007 | 0.020 | 0.013 |
| | (−6.95) | (−3.88) | (−1.43) | (−0.40) | (2.69) | (8.94) | (7.51) |
| Skew | −0.014 | −0.008 | −0.002 | −0.000 | 0.006 | 0.020 | 0.013 |
| | (−8.20) | (−3.79) | (−1.14) | (−0.03) | (2.28) | (9.15) | (7.81) |
| Kurt | −0.015 | −0.008 | −0.002 | −0.000 | 0.006 | 0.021 | 0.013 |
| | (−8.36) | (−4.09) | (−1.05) | (−0.05) | (2.34) | (9.21) | (7.28) |

second based on characteristics. The advantage of this approach over the cross-sectional regressions is that it does not impose any linear structure of returns (the disadvantage is that we can only control for one characteristic at one time).

Our sorts are conditional. We first sort stocks into quintiles based on stock characteristics. We sort stocks into quintile portfolios, as opposed to deciles, to keep the portfolios well populated. Second, within each quintile, we sort stocks based on the difference between HV and IV into quintiles or into the P and N portfolios. The five volatility quintiles or the two P and N portfolios are then averaged over each of the five characteristic portfolios. They, thus, represent volatility portfolios controlling for characteristics. Breakpoints for all stock characteristics are calculated each month based only on the stocks in our sample.

We report the average return and the associated *t*-statistic of this continuous time-series of monthly option portfolio returns in Table 6. In Panel A we report results for straddle portfolios and in Panel B we report results for delta-hedged call portfolios. In both panels, we find similar magnitudes of average returns across all controls. Average return range from 15.1% to 17.7% for the 5–1 straddle portfolio, from 8.7% to 10.3% for the P–N straddle portfolio, from 1.7% to 2.2% for the 5–1 delta-hedged call portfolio, and from 1.1% to 1.3% for the P–N delta-hedged call portfolio. These numbers are also comparable to those reported in Table 3, albeit a bit lower as expected (since we sort into quintiles in Table 6 as opposed to deciles in Table 3).

We conclude that, while the option returns covary with some of the stock characteristics that are found to be important for stock returns, this covariance is not enough to subsume the predictive power of the difference between HV and IV in explaining option returns.

## 5. Robustness and trading execution

### 5.1. Robustness

The results in the previous sections are presented after we have made many choices about key variables and sample periods. In this section, we check whether our results are robust to these decisions. We only present the salient features of these tests to not overwhelm the readers with numbers (complete set of results can be obtained from us upon request).

*Moneyness*: We select options close to ATM with moneyness in the range 0.975–1.025. Our reasons for choosing a narrow range of moneyness are twofold. One, we do not want the option returns to be driven by the smile in the volatility surface. Second, we want to use options which have a high sensitivity to volatility changes (vega) since our strategy is essentially a volatility trade. The vega is the highest around ATM and decreases quickly moving in or out of the money. However, we check the sensitivity of our results by increasing the moneyness range to 0.95–1.05. This leads to an increase in the number of overall stocks to 4,587 (from 4,344) and the average number of stocks in each decile portfolio to 97 (from 53). The volatility of the returns on the option portfolios also decreases as a consequence of the greater number of stock options in each portfolio. However, the magnitude of raw returns and the alphas is very similar to that reported in the paper.

*Subsample returns*: We replicate the analysis of Table 3 by dividing the data into two subsamples. The subsamples are formed by considering two different periods based on the sign of the changes in the VIX index and the sign of the market value-weighted CRSP portfolio returns. The option portfolio returns are higher in months in which VIX is increasing. For instance, the 10–1 straddle portfolio has an average return of 29.3% in months of positive changes in VIX and 18.0% in months of negative changes in VIX. We obtain essentially the same result when we sort the sample based on market returns—option returns are higher in months of negative market returns. These two results are not completely independent since market returns and changes in VIX are negatively correlated.

When the sample is divided in the two subsamples 1996–2000 and 2001–2006, we observe that the average returns are statistically significant in both subsamples (slightly higher for the period 1996–2000).

Since the options market is particularly active during months in which the S&P 500 futures options expire ("triple witching Friday") we also compute the average return for the strategies in only those particular months and compare these to the returns in other months. We find that there is no statistically meaningful difference in portfolio returns across these two sets of months.

The equity option market was particularly active during the years of the technology bubble. It is, therefore, useful to establish if portfolio returns are high only in the technology industry. We find this not to be the case. The long-short straddle portfolio is quite profitable in each industry. The highest average return (23.9% per month) is in the finance sector while the lowest return (18.9%) is in the utilities industry. We also check if the distribution of industries is uniform across our volatility sorted deciles and find this to be the case.

*Volatility measures*: Our basic measure of IV is the average of one-month ATM call and put implied volatility. We redo our analysis with two modifications. First, we calculate the IV using only the call or the put. Second, our options are American—this implies that the early exercise premium embedded in IV could make the IV measure not strictly comparable to HV. We check for this by removing all observations in which stock pays a dividend during the holding period.[21] The results of both these experiments are virtually identical to those reported in the paper.

An alternative to the Black and Scholes implied volatility provided by Ivy DB database is a model-free implied volatility (Jiang and Tian, 2005). This computation requires a large number of strikes for each stock at any point in time. The median number of strikes for options in our database is three, which implies that we can construct reliable estimates of model-free IV for a very small subset of stocks (for which there are at least 10 strikes for each option). Our results are qualitatively similar for this restricted sample.

We calculate HV from daily stock return data. We do not use GARCH (or any versions, thereof) to estimate volatility as our purpose is not to forecast future volatility from calibrated models. One can use high-frequency intra-day data to potentially improve our measure of HV. However, unavailability of this data to us precludes us from doing this. Our hope is that there is no systematic bias in our use of daily data vis-à-vis intra-day data, especially since we calculate HV from a long time period of one year.

*Earnings announcements*: Dubinsky and Johannes (2005) find spikes in IV around earnings announcements. We check whether this influences our results again by running two tests. First, we remove observations where our trade dates coincide with earnings announcement dates (approximately 5% of observations). Second, we remove all observations where a company announces earnings during the month prior to portfolio formation date or during the holding period month. Removing these observations has no material impact on our results. In addition to the above tests, we find that the earnings announcements are uniformly distributed in number across portfolios. Moreover, none of the portfolios shows abnormally positive or negative earnings around these announcements—the standardized unexpected earnings measure shows no pattern across decile portfolios.

### 5.2. Transaction costs

There is a large body of literature that documents that transaction costs in the options market are quite large and are in part responsible for some pricing anomalies, such as violations of the put–call parity relation.[22] It is essential to understand to what degree these frictions prevent an investor from exploiting the profits on the portfolio strategies studied in this paper. We accordingly discuss the impact of transaction costs, measured by the bid–ask spread and margin requirements, on the feasibility of the long-short strategy in this section.

#### 5.2.1. Bid–ask spreads

The results reported so far are based on returns computed using the mid-point price as a reference; however, it might not be possible to trade at that price in every circumstance. De Fontnouvelle, Fisher, and Harris (2003) and Mayhew (2002) show that the effective spreads for equity options are large in absolute terms but small relative to the quoted spreads. Typically the ratio of effective to quoted spread is less than 0.5. On the other hand, Battalio, Hatch, and Jennings (2004) study a period in the later part of our sample (January 2000 to June 2002) and find that for a small sample of large stocks the ratio of effective spread to quoted spread fluctuates between 0.8 and 1. Since transactions data are not

[21] We acknowledge the fact that while this controls for early exercise option of calls, American puts might still have a premium.

[22] See for example Figlewski (1989), George and Longstaff (1993), Gould and Galai (1974), Ho and Macris (1984), Ofek, Richardson, and Whitelaw (2004), and Swidler and Diltz (1992).

available to us, we consider three effective spread measures equal to 50%, 75%, and 100% of the quoted spread. For example if the bid price of an option is $3 and the ask price is $4, we consider the following scenarios: buy at $3.75 and sell at $3.25 (50% effective to quoted spread); buy at $3.87 and sell at $3.12 (75% effective to quoted spread); buy at $4 and sell at $3 (100% effective to quoted spread). This is done only at the initiation of the strategy since we terminate the portfolio at the expiration of the option.

Since the settlement of individual equity options requires delivery of the underlying, we also include the transaction costs of trading the underlying stocks. In the case of the straddles, the cost is incurred only at expiration and is relative to the shares that need to be bought or delivered as a consequence of the exercise of one of the two options. In the case of the delta-hedged strategies, the cost is incurred both at the initiation and at the expiration of the strategy as shares of the underlying (for a quantity equal to the option delta) are bought or sold on the first and last trading day.

The stock trading costs are computed from two different sources. First, the effective spreads for each stock are computed using the intra-day transactions and quotations (TAQ) data.[23] In all cases where we are unable to obtain data from TAQ, we calculate effective spreads using the method of Hanna and Ready (2005). Hanna and Ready estimate effective bid–ask spreads for stocks using transactions data. Then, they fit a regression model for the spreads using market capitalization, share price, monthly turnover, and monthly volatility as independent variables. We use the same characteristics and their estimated parameters (from their Table 2) to obtain an estimate of the bid–ask spreads. We refer the interested reader to their paper for further details on the computation of stock spreads.

Finally, to address the concern that the results might be driven by options that are thinly traded, we repeat the analysis by splitting the sample into two different liquidity groups. For each stock we compute the average quoted bid–ask spread and the daily average dollar volume of all the option contracts traded on that stock during the previous month. We then sort stocks into low and high option liquidity groups, based on these characteristics, and calculate average returns for the long-short straddle portfolios for these two groups of stocks.

We report the results of these computations for straddle portfolios in Panel A of Table 7. The 10–1 portfolio returns (left-hand side of Panel A) decrease substantially, as expected, after taking transaction costs into account: the average return decreases from 22.7% per month when trading at mid-point prices to 3.9% (statistically significant only at the 10% level) per month if we consider trading options at an effective spread equal to the quoted spread.

The liquidity of options also has an impact on returns as returns are higher for thinly traded stocks. Consider, as

[23] See Chordia, Roll, and Subrahmanyam (2000) for details on the construction of these data. We thank Tarun Chordia for making these data available to us.

**Table 7**
Impact of liquidity and transaction costs.

Portfolios are formed as in Table 2. They are further sorted into two groups based on stock options liquidity characteristics. We consider groups based on the average quoted bid–ask spread of all the options series traded in the previous month, as well as based on daily average dollar volume of all the options series traded in the previous month. The returns on options are computed from the mid-point opening price (MidP) and from the effective bid–ask spread (ESPR), estimated to be equal to 50%, 75%, and 100% of the quoted spread (QSPR). The closing price of options is equal to the terminal payoff of the option depending on the stock price and the strike price of the option. If the option expires in the money, exercising the option incurs stock transaction costs too. The table reports the average returns and the associated *t*-statistics, in parentheses, of this continuous time-series of monthly returns. The sample includes 4,344 stocks and is composed of 75,627 monthly pairs of call and put contracts. The sample period is 1996 to 2006.

| | 10–1 | | | | P–N | | | |
|---|---|---|---|---|---|---|---|---|
| | | ESPR/QSPR | | | | ESPR/QSPR | | |
| | MidP | 50% | 75% | 100% | MidP | 50% | 75% | 100% |
| | Panel A: Straddle returns | | | | | | | |
| All | 0.227 | 0.126 | 0.083 | 0.039 | 0.138 | 0.040 | −0.002 | −0.045 |
| | (10.41) | (5.98) | (3.94) | (1.84) | (8.42) | (2.48) | (−0.12) | (−2.73) |
| Alpha | 0.244 | 0.140 | 0.096 | 0.051 | 0.150 | 0.049 | 0.005 | −0.039 |
| | (9.73) | (5.81) | (3.98) | (2.10) | (7.76) | (2.54) | (0.28) | (−1.94) |
| | Based on average bid–ask spread of options | | | | | | | |
| Low | 0.195 | 0.130 | 0.105 | 0.082 | 0.144 | 0.084 | 0.060 | 0.038 |
| | (7.27) | (4.99) | (4.07) | (3.17) | (6.55) | (3.92) | (2.84) | (1.76) |
| High | 0.239 | 0.115 | 0.065 | 0.014 | 0.140 | 0.022 | −0.026 | −0.078 |
| | (8.95) | (4.51) | (2.57) | (0.56) | (6.19) | (0.99) | (−1.15) | (−3.27) |
| | Based on average trading volume of options | | | | | | | |
| Low | 0.227 | 0.110 | 0.064 | 0.017 | 0.134 | 0.023 | −0.022 | −0.070 |
| | (8.83) | (4.50) | (2.62) | (0.69) | (8.14) | (1.46) | (−1.41) | (−4.34) |
| High | 0.207 | 0.135 | 0.107 | 0.081 | 0.145 | 0.077 | 0.049 | 0.022 |
| | (6.95) | (4.65) | (3.70) | (2.77) | (5.68) | (3.05) | (1.95) | (0.86) |
| | Panel B: Delta-hedged call returns | | | | | | | |
| All | 0.027 | 0.010 | 0.007 | 0.005 | 0.018 | 0.002 | −0.000 | −0.002 |
| | (9.33) | (3.49) | (2.66) | (1.82) | (7.77) | (0.82) | (−0.14) | (−1.11) |
| | Based on average bid–ask spread of options | | | | | | | |
| Low | 0.024 | 0.013 | 0.012 | 0.010 | 0.018 | 0.008 | 0.007 | 0.006 |
| | (6.51) | (3.61) | (3.23) | (2.88) | (5.47) | (2.45) | (2.07) | (1.70) |
| High | 0.028 | 0.007 | 0.004 | 0.001 | 0.019 | −0.001 | −0.003 | −0.006 |
| | (9.30) | (2.46) | (1.49) | (0.52) | (6.91) | (−0.23) | (−1.19) | (−2.19) |
| | Based on average trading volume of options | | | | | | | |
| Low | 0.027 | 0.008 | 0.005 | 0.002 | 0.017 | −0.001 | −0.004 | −0.006 |
| | (9.53) | (2.88) | (1.90) | (0.92) | (8.21) | (−0.68) | (−1.93) | (−3.21) |
| High | 0.024 | 0.012 | 0.011 | 0.010 | 0.019 | 0.008 | 0.007 | 0.006 |
| | (5.61) | (2.90) | (2.57) | (2.27) | (5.15) | (2.21) | (1.85) | (1.51) |

an illustration, the results obtained by sorting on the average bid–ask spread of options. The 10–1 straddle portfolio has mid-point returns that are 19.5% on average for stocks with more liquid options (low bid–ask spreads) and 23.9% for stocks with less liquid options (high bid–ask spreads). These returns decline further with transaction costs. If the effective bid–ask spreads are the same as quoted spreads, the returns are still significantly positive at 8.2% for more liquid options and insignificant at 1.4% for less liquid options. This pattern arises because, by construction, the impact of transaction costs (as measured by bid–ask spreads) is higher for stocks with less liquid options. The results are qualitatively the same when we sort stocks based on the options average daily trading volume. A similar deterioration in performance is observed for the P–N straddle portfolio (right-hand side of Panel A). Since the before-costs returns of the P–N portfolio are lower than those of the 10–1 portfolio, the after-costs returns are positive and significant only for effective spread equal to half the quoted spread.

The cost-adjusted performance of delta-hedged call portfolios is reported in Panel B of Table 7. The pattern of higher returns for less liquid options found in Panel A for straddles is repeated in Panel B for delta-hedged calls. For instance, the average return (calculated using mid-points) of the 10–1 delta-hedged call portfolio increases from 2.4% to 2.8% per month, as one goes from the less liquid options to the more liquid options (liquidity as measured by bid–ask spreads). Bid–ask spreads also decrease the portfolio returns. For effective bid–ask spreads equal to the quoted spreads, the delta-hedged calls have statistically insignificant average returns of around 0.5% (−0.2%) for the 10–1 (P–N) portfolio.

We conclude that trading costs reduce the profits to our portfolios but do not eliminate them at reasonable estimates of effective spreads.[24] We also find that the profitability of option portfolios is higher for less liquid options.

### *5.2.2. Margin requirements*

Since the spread portfolios involve writing options in each month of the sample, we now investigate the impact of margin requirements on the strategy feasibility. While each brokerage house has its own way of determining margin requirements, most of the brokers adopt a scenario-based system in which the option position is evaluated in a range of scenarios using an option pricing model (typically the Black and Scholes, 1973, model). The most widely used scenario analysis algorithm for the determination of margin requirements is the Standard Portfolio Analysis of Risk (SPAN) system.[25]

[24] Note that we skip an additional day in constructing our portfolio strategies. While our motivation for this procedure is to avoid microstructure issues, the unintended consequence of this approach is that our traders trade based on the closing quotes on Tuesday. In actual practice, the option traders would have the whole day to decide when to optimally trade and minimize the market impact costs.

[25] We refer the interested reader to Santa-Clara and Saretto (2009) or to the SPAN manual (http://www.cme.com/clearing/rmspan/span/compont2480.html) for a detailed description of the algorithm.

In SPAN, the theoretical value of the position in each of the scenarios is compared to the market value of the option and the relative possible losses and gains are determined. The largest loss among those computed in the scenario analysis is called the option risk charge. The initial margin requirement is equal to the sum of the risk charge plus the proceeds obtained from writing the option. This procedure is repeated every day and the maintenance margin is computed as the sum of the option closing price and the risk charge.

The key factors that affect the determination of the margin requirements are the ranges of movement for the underlying stock value and the underlying volatility. In what follows, we use ±15% as the range for the price of the underlying stock, with progressive increments of 3%, and ±10% as the range for level of volatility.

For simplicity, we limit the analysis to straddle portfolio one (the decile that is shorted in the 10–1 strategy). Results of the analysis for the delta-hedged call portfolio are very similar. We calculate two different statistics: haircut ratio and maximum cumulative margin call. We compute the haircut ratio as the ratio of the initial risk charge to the value at which the straddle is written. Thus, the haircut ratio for stock $i$ in month $t$ is given by

$$H_{i,t} = \frac{M_{i,t,1} - S_{i,t,1}}{S_{i,t,1}},$$

where $M_{i,t,1}$ is the initial margin requirement and $S_{i,t,1}$ is the dollar amount received at the initiation of the strategy for writing the straddle. The haircut ratio represents the investor's equity in the option position as a percentage of the value of the straddle. We compute the maximum cumulative margin call as the difference between the highest value of the maintenance margin account throughout the holding period month and the initial margin requirement:

$$MC_{i,t} = \frac{\max\{M_{i,t,s}\}_{s=1}^{N_t} - M_{i,t,1}}{S_{i,t,1}},$$

where $N_t$ is the number of trading days in the holding period month $t$. The maximum cumulative margin call represents the additional capital that the investor might be required to provide to keep the positions open before the expiration of the options.

Since the portfolio is equal-weighted, we calculate the haircut ratio and the maximum cumulative margin call for the portfolio by taking an equal-weighted average of these statistics for the individual stocks in the portfolio for each month in the sample. We plot the monthly time-series of the portfolio haircut ratios in Panel A, and the monthly time-series of the maximum cumulative margin calls in Panel B of Fig. 1. Panel A shows that the haircut ratio fluctuates between 0.51 and 2.08. The average haircut ratio in our sample is 1.21, indicating that, for each dollar of written options, an investor needs to borrow \$1.21, on average, to satisfy the initial margin requirement. Panel B shows that there are two months in our sample in which the investor does not receive any additional margin call, and the largest additional monthly margin call is \$0.79 for each dollar of written options. The average maximum cumulative margin call is 0.21. Overall the analysis shows

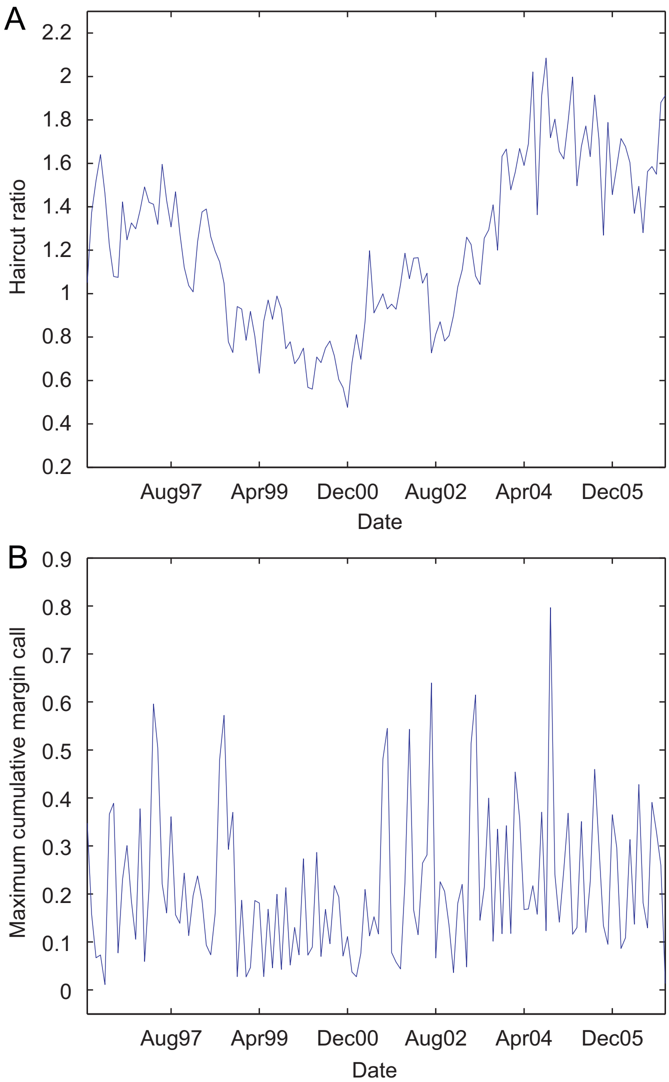


**Fig. 1.** Straddle portfolio margin requirements. Portfolios are formed as in Table 2. This figure plots the time-series of the haircut ratio and of the maximum cumulative margin call for decile one straddle portfolio. (Panel A) Haircut ratio for stock $i$ in month $t$ is given by $H_{i,t} = M_{i,t,1}/S_{i,t,1} - 1$, where $M_{i,t,1}$ is the margin requirement and $S_{i,t,1}$ is the dollar amount received at the initiation of the strategy for writing the straddle. (Panel B) Maximum cumulative margin call is the difference between the highest value of the margin account throughout the holding period month and the initial margin requirement, and is calculated as $MC_{i,t} = (\max\{M_{i,t,s}\}_{s=1}^{N_t} - M_{i,t,1})/S_{i,t,1} - 1$, where $N_t$ is the number of trading days in the holding period month $t$. Decile one straddle portfolio statistics are calculated by taking an equal-weighted average of the statistics for individual stocks in that decile. The sample includes 4,344 stocks and is composed of 75,627 monthly pairs of call and put contracts. The sample period is 1996 to 2006.

that an investor, on average, would need to borrow \$1.42 $(= 1.21 + 0.21)$ for each dollar of written options to satisfy margin requirements in our sample. The worst month from the margin requirement perspective is October 2004, when an investor would have to borrow \$1.71 to cover the initial margin and an additional \$0.79 to cover the margin call.

The evidence, therefore, suggests that margin requirements drive another wedge into the profitability of our trading strategy. An investor would need to put aside as much capital for margins as is invested in trading strategy, since the margin requirements are roughly equal to one and a half times the cost of written options. Since the return earned (if any) on the margin capital is usually small, this would further reduce the effective returns to our trading strategy.

To summarize the transaction cost results, our strategy has after-cost profits which can be much lower than the before-cost profits, and is, therefore, potentially profitable only to funds that can dedicate enough resources to its careful execution.

## 6. Overreaction

We have motivated our trading sorts by the mean-reverting nature of volatility. Large deviations between IV and HV are suggestive of mispricing, and the evidence

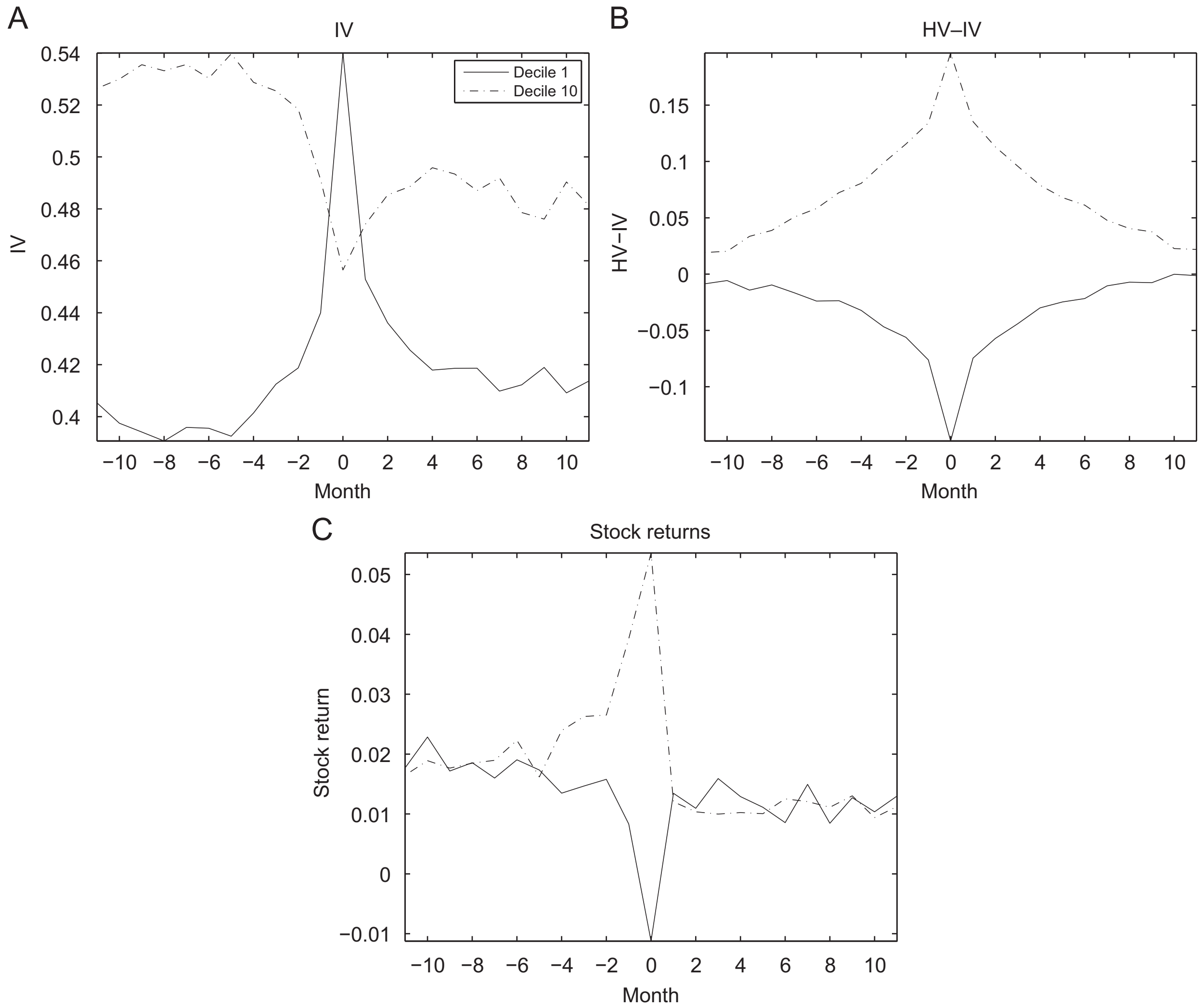


**Fig. 2.** Volatilities and stock returns before and after portfolio formation. This figure plots the average level of implied volatility, IV (Panel A), the average difference of historical volatility (HV) and implied volatility, HV–IV (Panel B), and the average stock returns (Panel C) for decile portfolios one and 10 during the period from 12 months before and 12 months after portfolio formation. Portfolio formation month is month zero. Portfolios are formed as in Table 2. In each plot, the solid line represents the portfolio corresponding to decile one and the dotted line corresponds to decile 10. All statistics are first averaged across stocks in each portfolio and then averaged across time. The sample includes 4,344 stocks and is composed of 75,627 monthly pairs of call and put contracts. The sample period is 1996 to 2006.

presented so far is consistent with this hypothesis. Are the deviations of IV from HV indeed temporary, as conjectured, and if so, what are the determinants of these deviations?

We start answering these questions by analyzing the pattern of volatilities before and after portfolio formation month. In Fig. 2 we plot the level of IV (Panel A) and the difference HV–IV (Panel B) for decile portfolios one and 10. We consider a range of 12 months before and 12 months after portfolio formation (portfolio formation month is month zero). Panel A shows that IV at time zero for decile one (10) is higher (lower) than what it has been in the previous 12 months. The figure also shows a striking pattern of IV *after* portfolio formation. IV for decile one (10) decreases (increases) after portfolio formation almost as quickly as it increases (decreases) in the months preceding the portfolio formation date. The pattern of changes in IV is not accompanied by a similar pattern of changes in HV. Panel B shows that the difference HV–IV is highest at time zero (by construction) and insignificant a year before and after portfolio formation. These figures show that deviations of HV from IV are indeed not persistent.

What leads to these temporary deviations? Recall that IV is an estimate of future volatility. To make the estimate for IV, an investor requires, at the very least, a model of volatility, the parameter estimates for this model, and an estimate of current volatility. Therefore, a misspecification of the model and/or a misestimation of one of the model inputs would lead to a misestimation of IV.

It is difficult to identify characteristics of stocks that would make them susceptible to these misspecification/

misestimation problems. Indeed, we do not find that the stocks in extreme deciles are abnormal in terms of size, book-to-market, or industry. However, the stocks in the extreme deciles do share one characteristic that is different from the stocks in the intermediate deciles—these stocks have large returns in the month immediately preceding the portfolio formation date. Panel C of Fig. 2 shows that stocks in decile one (10) have returns of −1.2% (5.3%) in month zero. Stock returns in the months preceding or following the portfolio formation month show no appreciable pattern (stocks in decile 10, though, have slightly higher than usual returns even in the two months before portfolio formation).

The evidence so far, thus, shows that extreme returns to underlying stocks are related to extreme changes in IV (expectation of future riskiness of the stocks). The natural question at this stage is whether the perceived changes in future volatility are rational. We find this not to be the case as future realized volatility does not change by as much as predicted by IV. Indeed, when we run the rationality test of predicting future volatility using current IV (univariate Eq. (1)) separately for stocks in deciles one, 10, and the intermediate deciles, we find that the average $\alpha$ and $\beta_1$ coefficients are −0.275 and 0.975 for decile one, 0.162 and 1.085 for decile 10, and 0.000 and 0.975 for the intermediate deciles, respectively. Recall that, under the null hypothesis that IV is an unbiased forecast of future volatility, the coefficient $\beta_1$ should be equal to one and the coefficient $\alpha$ should be equal to zero. It is easy to see that the average forecast error ($\overline{\text{FV}} - \overline{\text{IV}}$) in these regressions is given by $\alpha + (\beta_1 - 1)\overline{\text{IV}}$, and is thus closely related to $\alpha$ (since $\beta_1$'s are close to one). Negative (positive) $\alpha$ intercepts in these regressions imply, therefore, that IV over (under) predicts future volatility for decile one (10).[26] To summarize, investors overreact to high realizations of stock returns leading them to form an estimate of future volatility (IV) that is not justified by subsequent events.

While overreaction is difficult to reconcile within the traditional finance paradigm of rational investors, there are many behavioral finance models that seek to accommodate empirical challenges to CAPM-like models.[27] The building blocks of these models are either biased beliefs or non-standard preferences. The majority of these models endogenously generate under- and overreaction. However, these studies are typically concerned with reconciling the pattern of returns (first moments) rather than the pattern of volatility of returns (second moments), which is our focus. One exception is the model of Barberis and Huang (2001).

Investors in the BH model are loss averse and do mental accounting. There are three main points of departure in BH from standard rational agents models. First, investors get utility from gains and losses in wealth, rather than from absolute levels of their wealth. Second, they are more sensitive to losses than to gains. This is known as loss aversion and implies a kink in the utility function. Third, investors engage in mental accounting, which answers the question of which gains/losses they pay attention to. BH suggest that when doing mental accounting, investors exhibit narrow framing. In other words, they define the gains and losses narrowly over individual stocks (rather than over total wealth).[28]

The modeling assumptions of BH add an additional dimension of risk to each stock. Investors gain utility from good performance of a stock and become less concerned about future losses on this stock, as future losses are cushioned by prior gains. In other words, they perceive this stock to be less risky than before. Conversely, if one of her stocks performs poorly, the investor finds this painful and becomes more sensitive to the possibility of further losses on this stock. In other words, they treat this stock to be riskier than before.

This pattern of overreaction in the BH model is precisely the one observed in our data. Investors raise IV (increase their estimate of future volatility) after negative returns for stocks in decile one, and reduce IV (decrease their estimate of future volatility) after positive returns for stocks in decile 10. However, the perceived increase/decrease in riskiness of these stocks is not justified by the future and, therefore, leads to high returns on strategies such as ours that seek to exploit volatility mispricing.

In addition to the link with the theoretical model, our findings of investor overreaction in option markets are also related to those of Stein (1989). Stein studies the term structure of the implied volatility of index options and finds that investors ignore the long-run mean reversion in implied volatility and instead overweight the current short-term implied volatility in their estimates of long-term implied volatility. Our evidence is also consistent with Poteshman (2001), who finds that investors in the options market overreact, particularly to periods of increasing or decreasing changes in volatility. This is analogous to our deciles one and 10, which exhibit the greatest changes in volatility, and are also characterized by the most mispricing.

## 7. Conclusion

We emphasize that our results do not depend on the validity of the Black and Scholes (1973) or the Cox, Ross, and Rubinstein (1979) models. Implied volatilities should be interpreted as representation of option prices. Therefore, the reader should view our portfolio sorts as sorts on option prices with decile one (10) representing cheap (expensive) options. This perspective does not require one to take a stand on the correct option pricing model. The objective of our paper is to document the existence of a substantial spread in the cross-section of U.S. equity options sorted on a very simple criterion.

The underlying reason for the empirical regularity that we observe in equity option prices is unclear. While we

[26] Exact calculations yield that the forecast error is −0.254, 0.079, and 0.025 for decile one, decile 10, and intermediate deciles, respectively.

[27] The literature on anomalies and behavioral responses to explain these patterns in the data is too voluminous for us to cite here. See the surveys by Barberis and Thaler (2003) and Daniel and Womack (2001).

[28] BH also model an economy in which investors are loss averse over the fluctuations of their stock portfolio. However, they find that this economy is less successful in explaining the empirical phenomena.

find that our option returns are not related to obvious sources of risk, we cannot conclusively establish that these are true alphas. It is possible that the profits to our volatility portfolios arise as compensation for some unknown aggregate risk. If such is indeed the case, the daunting task of formulating a cross-sectional options return model that accounts for our portfolios' returns is left to future research.

If, instead, these returns are abnormal, the evidence presented in the paper raises the question of what accounts for this volatility mispricing. Barberis and Huang (2001) model an economy with loss averse investors who count gains and losses in individual stocks separately. In the model, investors become excessively optimistic (pessimistic) about the future riskiness of the stock after large positive (negative) returns. Our results are consistent with their model.

## References


Andersen, T., Bollerslev, T., Christoffersen, P., Diebold, F., 2006. Volatility and correlation forecasting. In: Elliott, G., Granger, C., Timmermann, A. (Eds.), Handbook of Economic Forecasting, vol. 1. North-Holland, Amsterdam, pp. 778–878.

Bakshi, G., Kapadia, N., 2003. Delta-hedged gains and the negative market volatility risk premium. Review of Financial Studies 16, 527–566.

Barberis, N., Huang, M., 2001. Mental accounting, loss aversion, and individual stock returns. Journal of Finance 56, 1247–1292.

Barberis, N., Thaler, R., 2003. A survey of behavioral finance. In: Constantinides, G., Harris, M., Stulz, R. (Eds.), Handbook of the Economics of Finance, vol. 1B. North-Holland, Amsterdam, pp. 1053–1123.

Battalio, R., Hatch, B., Jennings, R., 2004. Toward a national market system for U.S. exchange-listed equity options. Journal of Finance 59, 933–962.

Battalio, R., Schultz, P., 2006. Options and the bubble. Journal of Finance 61, 2071–2102.

Black, F., Scholes, M., 1972. The valuation of option contracts and a test of market efficiency. Journal of Finance 27, 399–417.

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of Political Economy 81, 637–654.

Brennan, M., Chordia, T., Subrahmanyam, A., 1998. Alternative factor specifications, security characteristics, and the cross-section of expected returns. Journal of Financial Economics 49, 345–373.

Britten-Jones, M., Neuberger, A., 2000. Option prices, implied price processes, and stochastic volatility. Journal of Finance 55, 839–866.

Carhart, M., 1997. On persistence in mutual fund performance. Journal of Finance 52, 57–82.

Chava, S., Tookes, H., 2006. Where (and how) does news impact trading? Unpublished working paper. Texas A + M University and Yale School of Management.

Chordia, T., Roll, R., Subrahmanyam, A., 2000. Commonality in liquidity. Journal of Financial Economics 56, 3–28.

Christensen, B., Prabhala, N., 1998. The relation between implied and realized volatility. Journal of Financial Economics 50, 125–150.

Cochrane, J., 2001. Asset Pricing. Princeton University Press, Princeton, NJ.

Coval, J., Shumway, T., 2001. Expected option returns. Journal of Finance 56, 983–1009.

Cox, J., Ross, S., Rubinstein, M., 1979. Option pricing: a simplified approach. Journal of Financial Economics 7, 229–263.

Daniel, K., Womack, K., 2001. Behavioral finance. In: Logue, D., Seward, J. (Eds.), Handbook of Modern Finance. Warren Gorham & Lamont, New York, pp. 1–30.

De Fontnouvelle, P., Fisher, R., Harris, J., 2003. The behavior of bid–ask spreads and volume in options markets during the competition for listings in 1999. Journal of Finance 58, 2437–2463.

Diether, K., Malloy, C., Scherbina, A., 2002. Differences of opinion and the cross-section of stock returns. Journal of Finance 56, 2113–2141.

Driessen, J., Maenhout, P., Vilkov, G., 2009. The price of correlation risk: evidence from equity options. Journal of Finance 64, 1377–1406.

Duan, J.-C., Wei, J., 2009. Systematic risk and the price structure of individual equity options. Review of Financial Studies 22, 1981–2006.

Duarte, J., Jones, C., 2007. The price of market volatility risk. Unpublished working paper, University of Southern California.

Dubinsky, A., Johannes, M., 2005. Earnings announcements and equity options. Unpublished working paper, Columbia University.

Fama, E., French, K., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56.

Figlewski, S., 1989. Options arbitrage in imperfect markets. Journal of Finance 44, 1289–1311.

Friend, I., Westerfield, R., 1980. Co-skewness and capital asset pricing. Journal of Finance 35, 897–913.

George, T., Longstaff, F., 1993. Bid–ask spreads and trading activity in the S&P 100 index option market. Journal of Financial and Quantitative Analysis 28, 381–397.

Gould, J., Galai, D., 1974. Transactions costs and the relationship between put and call prices. Journal of Financial Economics 1, 105–129.

Granger, C., Poon, S.-H., 2003. Forecasting volatility in financial markets: a review. Journal of Economic Literature 41, 478–539.

Green, C., Figlewski, S., 1999. Market risk and model risk for financial institution writing options. Journal of Finance 54, 1465–1499.

Hanna, J., Ready, M., 2005. Profitable predictability in the cross section of stock returns. Journal of Financial Economics 78, 463–505.

Harvey, C., Siddique, A., 2000. Conditional skewness in asset pricing tests. Journal of Finance 55, 1263–1295.

Ho, T., Macris, R., 1984. Dealer bid–ask quotes and transaction prices: an empirical study of some AMEX options. Journal of Finance 39, 23–45.

Jiang, G., Tian, Y., 2005. The model-free implied volatility and its information content. Review of Financial Studies 18, 1305–1342.

Kraus, A., Litzenberger, R., 1976. Skewness preference and the valuation of risky assets. Journal of Finance 31, 1085–1100.

Leland, H., 1999. Beyond mean–variance: performance measurement in a nonsymmetrical world. Financial Analysts Journal 55, 27–36.

Mayhew, S., 2002. Competition, market structure, and bid–ask spreads in stock option markets. Journal of Finance 57, 931–958.

Newey, W., West, K., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.

Ni, S., 2006. Stock option return: a puzzle. Unpublished working paper, Hong Kong University of Science and Technology.

Ni, S., Pan, J., Poteshman, A., 2008. Volatility information trading in the option market. Journal of Finance 63, 1059–1091.

Ofek, E., Richardson, M., Whitelaw, R., 2004. Limited arbitrage and short sales restrictions: evidence from the options market. Journal of Financial Economics 74, 305–342.

Pástor, L., Stambaugh, R., 2003. Liquidity risk and expected stock returns. Journal of Political Economy 111, 642–685.

Poteshman, A., 2001. Underreaction, overreaction, and increasing misreaction to information in the options market. Journal of Finance 56, 851–876.

Poteshman, A., Serbin, V., 2003. Clearly irrational financial market behavior: evidence from the early exercise of exchange traded stock options. Journal of Finance 58, 37–70.

Rubinstein, M., 1976. The valuation of uncertain income streams and the pricing of options. Bell Journal of Economics 7, 407–425.

Sadka, R., 2006. Momentum and post-earnings-announcement drift anomalies: the role of liquidity risk. Journal of Financial Economics 80, 309–349.

Santa-Clara, P., Saretto, A., 2009. Option strategies: good deals and margin calls. Journal of Financial Markets 12, 391–417.

Stein, J., 1989. Overreactions in the options market. Journal of Finance 44, 1011–1023.

Swidler, S., Diltz, D., 1992. Implied volatilities and transaction costs. Journal of Financial and Quantitative Analysis 27, 437–447.