Informed Trading of Options, Option Expiration Risk, and Stock Return Predictability

Martijn Cremers[a],

Ruslan Goyenko[b],

Paul Schultz[a],

and

Stephen Szaura[b]

Abstract

The percentage difference between implied stock prices from options and actual stock prices, which we term the implied price difference ('IPD'), predicts stock returns up to ten months into the future. IPD predicts stock returns even after including a number of other option-based variables, and among both the largest stocks as well as small cap stocks. Information about future stock returns comes from long-term, not short-term options. This suggests that informed investors who do not know when their longer-term information will be impounded in prices tend to trade long-term options to avoid expiration risk.

March, 2019.

[a]Mendoza College of Business, University of Notre Dame.
[b]Desautels Faculty of Management, McGill University.
The authors acknowledge financial support of CDI (Canadian Derivatives Institute).
Corresponding author: Paul Schultz, pschultz@nd.edu.

1. Introduction

An investor with private information about a firm's value can profit on it by trading the firm's stock, options on the stock, or both. Trading options provides two significant advantages to informed investors. First, options provide leverage or relax the investor's borrowing constraints. When buying stock on margin, an investor cannot borrow more than 50% of the stock's cost. A basis of option pricing theory is that a call option can be replicated by borrowing money and buying shares. The implicit borrowing in an option purchase, particularly if the option is out-of-the-money, can be much greater than 50%. The implicit leverage means that for an informed investor, returns are much higher for an option position than a stock position. The second advantage of trading options for an informed trader is that it is often easier and cheaper to buy put options or sell call options than it is to short-sell shares, especially if the short position may need to be held for a longer period of time. Short-selling can be expensive, and it may be difficult to find shares to borrow. In those cases, buying puts or writing calls provide attractive alternatives for investors who believe that a stock is overvalued.

Trading options also present two significant disadvantages to informed investors relative to trading stocks. First, shares are typically much more liquid than options. Bid-ask spreads for options are typically much wider than spreads for stocks and putting together a large options position can move option prices significantly. A second disadvantage of options, that has not received much attention in the literature, is that informed investors seeking to profit from their information by trading options face expiration risk. Even if an investor knows with certainty that a stock is mispriced, she may not be able to predict when the mispricing will be corrected or when the necessary information will become public. If options expire before the stock price reflects the investor's information, she will lose the value of the options and must acquire new options and pay high trading costs a second time. We provide evidence that expiration risk is important in practice, and that informed traders appear to forego the greater leverage of short-term options and instead trade options with more time to expiration to mitigate expiration risk.

To examine the information in option prices, we follow Manaster and Rendleman (1982) and calculate implied stock prices jointly with implied volatilities. The implied stock price is a more direct and intuitive measure of option traders' assessment of the value of the underlying stock than are implied volatilities (see An, Ang, Bali, and Cakici (2014), Cremers and Weinbaum (2010) and Xing, Zhang, and Zhao (2010)) or relative option trading volumes (see Ge, Lin, and Pearson (2016), Johnson and So (2012) and Pan and Poteshman (2006)). And, as we will show, it forecasts stock returns more accurately than measures based on implied volatilities or the trading volume of calls relative to puts. The percentage difference between the implied and actual stock price, which we refer to as the implied price difference or

IPD, predicts stock returns for up to ten months into the future. IPD predicts stock returns for both the largest 500 stocks and smaller stocks.

When we form equally-weighted portfolios of stocks with low and high values of IPD, we find that the low IPD portfolios, in which implied stock prices are lower than actual prices, earn negative, statistically significant four-factor abnormal returns in each of the next three months. The equally-weighted high IPD portfolios do not, however, earn statistically significant positive abnormal returns. For the small stocks that make up a large proportion of the equally-weighted portfolios, negative information is impounded in option prices before stock prices. This is consistent with informed investors using options to get around short sale costs and restrictions for small stocks. In contrast, value-weighted portfolios of high IPD stocks earn positive abnormal returns over the next three months while low IPD stock portfolios earn negative returns. Both of these results are strongly statistically and economically significant. For the three months after portfolio formation, the cumulative four-factor alpha is -1.28% for the value-weighted portfolio of low IPD stocks and 1.21% for the portfolio of high IPD stocks. For the large stocks that dominate value-weighted portfolios, short-sale restrictions appear to be a less important reason for informed investors to trade options. For large stocks, the leverage provided by options remains a reason for informed investors to trade them.

A number of papers, including Cremers and Weinbaum (2010), Ge, Lin and Pearson (2016), Johnson and So (2012), Pan and Poteshman (2006), and Xing, Zhang, and Zhao (2010) use variables derived from option volume or implied volatilities in options to predict stock returns. We estimate Fama-Macbeth regressions with month-ahead stock returns as the dependent variable and IPD, other option-based variables, and predictive variables like book-to-market, size, and momentum as explanatory variables. When all stocks are included in the regressions, IPD remains a significant predictor of returns when any one of the individual alternative option-based measures are added to the regression. It does not, however, predict returns as well as the Cremers and Weinbaum measure. On the other hand, when the regressions include just the 500 largest stocks, IPD is a highly significant predictor of stock returns while all of the other option-based variables, including the Cremers and Weinbaum measure, are insignificant. When all variables are included in the regression, the coefficient on IPD is 0.5489 with a t-statistic of 4.09. The standard deviation of IPD for large stocks is 1.058%, so, after including all other option-based variables, a one standard deviation increase in IPD increases stock returns the next month by 58 basis points.

Perhaps our most significant finding is that implied stock prices estimated using only long term options are much better predictors of stock returns than implied stock prices estimated from short term options. When we form portfolios based on IPD calculated with options with 30 – 60 days to expiration, the four-factor alphas of the high IPD portfolio exceed the alphas of the low IPD portfolio by an insignificant 4.6 basis points over the next month. In contrast, the four-factor alphas of high IPD portfolios exceed the alphas of low IPD portfolios by 69 basis points per month when options with 90 - 180 days to expiration are used to estimate IPD. When options with more than 180 days to expiration are used to estimate IPD, the alphas of the high IPD portfolios are larger than the low IPD portfolio alphas by 82 basis points.

The predictive information contained in long-term options can take a long time before it is finally reflected in stock prices. For the sample of all stocks, we find stock return predictability up to 10 months, and for the sample of the largest 500 stocks, this predictability is economically and statistically significant for up to 12 months. The sum of the IPD coefficients for the first 12 months for the largest 500 stocks is 1.85 when IPD is estimated using options with at least 90 days to expiration. The standard deviation of IPD for the largest 500 stocks is 1.33 when estimated using long-term options. Hence, a one standard deviation change in IPD leads to a 1.33 x 1.85 = 2.46% difference in returns over 12 months. It is important to note though that this is the difference when long-term options are used to estimate IPD. When short-term options are used to calculate IPD, IPD is not a significant predictor of stock returns in any of the following 12 months.

This long time over which information contained in options predict stock returns indicates that informed investors face substantial expiration risk when trading options. Because they do not know when information will be incorporated in stock prices, they appear to prefer trading longer term options to minimize their expiration risk. We show analytically that the costs of rolling over expiring options until information in incorporated in stock prices can be much greater than the extra cost of longer-term options. Using Black-Scholes prices, we show that a one-year option on a stock with a standard deviation of 0.4 costs 18% of the stock price. On the other hand, if information is incorporated more slowly than the informed investor expects, and the investor rolls over a series of four three month options, the total cost is 33.6% of the stock price. This cost does not include the large bid-ask spreads that are paid each time a position is rolled over.

Not all informed investors can trade options. Many institutional investors are constrained to trading shares only. Long-term investors can profitably buy stock on the basis of information that may not be revealed for several months while institutional investors with short horizons may not find it worthwhile exploit such information. Consistent with this, we find that IPD calculated from long term options does a far better job of predicting returns for stocks that are not held heavily by long-term investors.

If informed investors are trading options, we would expect IPD to have the most predictive power when there is a lot of uncertainty. We estimate implied stock prices and implied volatilities jointly, so there is no mechanical relation between them. Each month, we first sort stocks into three terciles based on implied volatility from options with more than 90 days to expiration. Within each tercile, we then sort stocks into five quintiles based on IPD from long-term options. When all stocks are considered, the difference between high and low IPD portfolios' month-ahead four-factor alphas is 34 basis points for the low implied volatility tercile, 51 basis points for the middle tercile, and 103 basis points for the high implied volatility tercile. For the largest 500 stocks, the differences are 49, 77 and 90 basis points. IPD has the most predictive power when there is a lot of uncertainty.

The remainder of the paper is organized as follows. Section 2 reviews related literature. In Section 3 we discuss the advantages to informed traders of buying specific types of options. In Section 4 we discuss how we estimate IPD, the average percentage difference between implied and actual stock prices. In Section 5 we show how IPD predicts stock returns for the three months following its estimation, even when other options-based variables are considered. We discuss expiration risk and show that the information in IPD comes mainly from long-term options in Section 6. Section 7 offers conclusions.

2. Related Literature

When considering whether to trade options or the underlying stock, informed investors weigh the increased leverage of options against their lower liquidity and higher trading costs. This tradeoff is formally modeled by Easley, O'Hara and Srinivas (1998). They construct an asymmetric information model in which informed investors choose between trading options or stocks. In their model, at least some informed trading takes place in the options market, if options provide sufficient leverage, if the liquidity of the stock market is low, or if there are many informed traders. On the other hand, if the options market is not sufficiently liquid, all of the informed trading takes place in the stock market.

In recent years, a number of researchers have considered the implications from Easley, O'Hara and Srinivas (1998) and studied whether option prices or, equivalently, implied volatilities, include information that is not yet incorporated in stock prices. Cremers and Weinbaum (2010) show that deviations from put-call parity can predict stock returns. Using all pairs of calls and puts on a stock with the same strike price and expiration date, and weighting each pair by its open interest, they calculate weighted average differences between call and put implied volatilities. A higher (lower) implied volatility for calls than puts means that the call prices are high (low) relative to the put prices. They divide stocks into five quintiles based on these differences and calculate abnormal returns over the next week and next four weeks for these portfolios. Portfolios with call implied volatilities that exceed put implied volatilities earn positive abnormal returns while those with larger put implied volatilities earn negative abnormal returns. The finding that higher call implied volatilities imply positive abnormal returns suggests that option prices' ability to predict stock returns is not solely a result of short sale restrictions. Cremers and Weinbaum show further that the predictive power of option prices increases as options become more liquid or underlying shares become less liquid.

An, Ang, Bali, and Cakici (2014) examine the power of changes in implied volatilities to forecast stock returns. Implied volatilities are obtained from the daily implied volatility surface calculated by OptionMetrics. Their empirical analysis uses end-of-month call and put implied volatilities for options with a delta of 0.5 and 30 days to maturity. An et al. sort stocks into decile portfolios each month based on the change in the stock's call and put implied volatilities. Larger increases in call volatilities are associated with larger stock returns the next month while larger increases in put volatilities are associated with lower stock returns in the following month. The difference in returns between the portfolio with the largest (typically positive) change in implied call volatility and the portfolio with the smallest (typically negative) change in implied volatility is about 1% over the next month. Differences in the next months' abnormal returns, calculated with either the CAPM or the Fama-French three factor model, are also about 1%. Differences in returns and abnormal returns across portfolios of stock with the largest and smallest changes in put implied volatilities are about 0.5%.

Volatility smirks, defined as the difference between the implied volatility of out-of-the-money puts and at-the-money calls, are examined by Xing, Zhang, and Zhao (2010). Out-of-the-money puts are a natural investment for investors with negative news or bearish opinions about a stock. Strong demand for out-of-the-money puts will push up their prices and thus their implied volatilities. At-the-money calls are typically the most liquid options, hence subtracting the implied volatility of at-the-money-calls from the implied volatility of out-of-the-money puts provides a measure of the excess demand for puts.

Xing, Zhang, and Zhao (2010) refer to this measure as skewness. Using closing prices from 1996 – 2005, they calculate weekly skewness for individual stocks by averaging daily skewness. They sort stocks into quintiles based on skewness and show that stocks with high skewness (large out-of-the-money put implied volatilities) underperform stocks with low skewness (small out-of-the-money put implied volatilities) by 15 to 20 basis points. Stocks with low skewness continue to outperform stocks with high skewness for up to six months after the portfolio formation.

A complimentary line of research studies how option order flow, rather than option prices, can predict stock prices. Roll, Schwartz, and Subrahmanyam (2010) examine the determinants of the natural logarithm of the ratio of option to share volume, or O/S. They show that O/S increases cross-sectionally with implied volatility, which may proxy for the value of information. O/S decreases with the percentage spread of the options. They present strong evidence that the absolute value of abnormal returns around earnings announcements is increasing in O/S measured before the earnings announcement. This suggests that option investors are particularly active relative to stock investors before large price moves.

Johnson and So (2012) observe that short-sale restrictions make it more likely for an investor with negative information, rather than positive information, to trade options rather than shares. They propose that O/S may contain information about future stock returns. A high value of O/S is likely to reflect a large volume of trading by pessimistic investors who trade options rather than shares because of short-sale restrictions. They calculate O/S using weekly volume shares of put and call options that expire between five and 35 days after the trade. They form portfolios based on O/S ratios and show that higher values of option to stock volume predict lower returns in the four weeks following portfolio formation.

Ge, Lin, and Pearson (2016) confirm that high O/S ratios predict negative stock returns, but also provide evidence that implicit leverage, not short-sale restrictions, is behind option trades' power to predict stock returns. They break down volume for puts and calls into buy volume and sell volume, and into trades that open and close positions. Both bullish and bearish volume predict stock returns with the strongest predictions coming from volume that opens call positions. Ge, Lin, and Pearson state that high values of O/S are associated with negative future returns because "more components of options volume negatively predict returns than positively predict returns, due to the fact that trading volume stemming from the unwinding of bought call positions negatively predicts returns."

Hu (2014) notes that option market makers hedge their option positions by taking offsetting positions in the underlying stock. The number of shares in the offsetting position is the option's delta times the options net volume. Hu breaks down the order imbalance in shares into the imbalance caused by option transactions and the imbalance that is independent of options trading. The imbalance attributed to market maker hedging of options trades predicts future stock returns, implying that option order flow contains information about the stock value.

Pan and Poteshman (2006) show that the daily ratio of the number of put contracts purchased to the sum of put and call contracts purchased predicts stock returns. A large ratio, reflecting greater public purchases of puts than calls, implies negative private information while a small ratio, reflecting greater public purchases of calls than puts, implies that traders have bullish private information. Slope coefficient estimates obtained by regressing the next-day four-factor adjusted stock return on the ratio indicate that buying stocks with all volume coming from buys of calls and selling stocks with all volume coming from put buys yields a highly significant average return of over 50 basis points over the next day. Larger excess returns are produced when they use options that are out-of-the-money or close to expiration as these options provide more leverage. Pan and Poteshman note that their results are not incompatible with market efficiency. In their tests, returns are predicted using information that is not available to the public. Investors are not able to observe whether option trades, and thus option volume, is buyer initiated.

An, Ang, Bali, and Cakici (2014), Xing, Zhang, and Zhao (2010), Cremers and Weinbaum (2010) all predict stock returns based on implied volatilities. Implied volatilities from calls can be high if option traders believe the stock's value is higher than the current stock price. Similarly, implied volatilities from puts can be high if option traders believe that the current stock price is too high. High implied volatilities can, however, simply reflect option traders' beliefs that volatilities will be high. Implied volatilities therefore provide indirect and noisy ways to estimate option traders' beliefs about future stock prices.[1]

Some recent papers have made use of stock prices implied from options. Chakravarty, Gulen, and Mayhew (2004) use vector autoregressions of stock prices and implied stock prices from options at one second intervals to estimate the proportion of information impounded in prices by option and stock prices. Across 60 stocks, they find that the information share of options varies from 11.8% to 23.5%. The information share is higher for out-of-the money options than for in-the-money or at-the-money options. The information share of options increases with the ratio of option volume to stock volume, and decreases with the ratio of option effective spreads to stock effective spreads. These results are consistent with the Easley, O'Hara and Srinivas model.

Muravyev, Pearson, and Broussard (2013) use option prices and put-call parity to estimate implied bid and ask prices for 36 liquid stocks and three ETFs for the period from April 17, 2003 through October 18, 2006. In contrast to papers that find options prices predict stock prices, they present evidence that over short intervals, almost all price discovery occurs in the stock market. Using tick-by-tick trade and quote data, they find many instances in which the implied ask price from options was less than the stock's bid price, or when the implied bid price from options exceeded the stock's ask price. When these price discrepancies occurred option prices typically changed to eliminate the arbitrage opportunity while stock prices typically moved further away from option implied prices. Their results are confirmed in a much larger sample using changes in option bid and ask prices rather than implied stock prices. Muravyev, Pearson, and Broussard use options with 10 to 70 to expiration to calculate implied prices. We find in this paper that options with longer times to expiration have more information about future stock returns.

3. Estimates of the advantages and disadvantages of trading options rather than stock

Options data is from CBOE/LiveVol and includes intraday option trades and quotes for 2004-2013. We apply the following standard microstructure data filters to our options trades data: option trades must be between 9:30 am and 4:00 pm, cancelled trades are removed, all option best bid and best ask quotes must satisfy the relationship 0 < Best Bid < Best Ask < 5·Best Bid, and option trades are deleted if the best bid, best ask, or trade price is missing. We require that the option trade price be less than or equal to twice the option midpoint at the time of trade. Option trades are also deleted if contemporaneous option quote midpoints are less than ten cents, the options have zero trade volume, or have ten or fewer days to maturity.

[1] See Ni, Pan, and Poteshman (2008) for evidence that investors with information on volatility trade options, and that volatility trades affect option prices.

An advantage of buying options rather than shares is the higher implicit leverage in the options. To see the leverage in call options, consider the Black-Scholes model. The value of a European call is given by

$$C = \left(S - \sum_{i=1}^{I} e^{-rt_i} D_i\right) N(d1) - e^{-rT} K N(d2), \quad (1)$$

where $d1 = \frac{\log\left(\frac{(S-\sum e^{-rt} D)}{K}\right) + (r+\sigma^2/2)*T}{\sigma\sqrt{T}}$ $and\ d2 = d1 - \sigma\sqrt{T},$ S is the stock price, $\sum_{i=1}^{I} e^{-rt_i} D$ is the present value of dividends paid over the life of the option, K is the strike price, r is the risk-free interest rate, σ is the stock volatility, T is the time to expiration, and N(.) is the cumulative normal distribution. Note that Black-Scholes option pricing is based on replicating a call option using a portfolio of shares of stock combined with riskless borrowing. In the Black-Scholes formula, N(d1) is the number of shares in the portfolio, and $(S\text{-}\Sigma e^{-rT}D) \cdot N(d1)$ is the total value of the shares. The second term in the formula is the amount borrowed. Hence, the leverage implicit in a call option, which is the amount borrowed divided by the value of the shares, can be calculated as

$$Leverage\ = [e^{-rT} K N(d2)] / \left[\left(S - \sum_{i=1}^{I} e^{-rt_i} D_i\right) N(d1)\right]. \quad (2)$$

Panel A of Table 1 shows the mean implicit leverage of in-the-money, at-the-money, and out-of-the-money calls in our sample for various times to expiration. We define in-the-money options as those with an absolute value for their delta of 0.625 to 0.875, at-the-money options as those with absolute values of deltas of 0.375 to 0.625, while out-of-the money options have absolute values of deltas in the range of 0.125 to 0.375[2]. For at-the-money calls with less than 30 days to expiration, the mean implicit leverage is 0.8835. That is, buying these calls is like buying shares and borrowing 88.35% of the cost of the shares. This leverage is greater than the maximum margin of 50% available for buying shares. As Panel B shows, leverage increases as options move further out of the money and as the time to expiration decreases. For example, for at-the-money options the implicit leverage decreases from 88.35% for options with less than 30 days to expiration to 83.1% for options with 30 to 179 days to expiration. An informed investor, in choosing the time-to-expiration of options, balances the greater leverage of short-term options against the greater likelihood that the options will expire before the information is impounded in prices.[3]

[2] This is definition of moneyness is also used by Bollen and Whaley (2004), among others.

[3] Johnson and So (2012) use N(d1)S/C as a measure of leverage. Their measure can be thought of as the multiple of the stock return earned by investing in the option rather than the stock. So, if N(d1)S/C is 2, the option return is twice the share return. Our measure, on the other hand, gives the proportion of the position that is financed by borrowing. We use it because it makes for an easy comparison with margin requirements.

For puts, we calculate the proportion of money from the implicit short position that is invested at the riskless rate. When a stock is shorted, the short-seller has to put up 150% of the short-sale proceeds as collateral. The Black-Scholes model for puts is given by

$$P = e^{-rT} K N(-d2) - \left(S - \sum_{i=1}^{I} e^{-rt_i} D_i\right) N(-d1). \quad (3)$$

The put is priced as a short position in N(-d1) shares and an investment of e^{-rT} K N(-d2) in the riskless security. The proportion of money that the investor implicitly puts up when he purchases a put is the value of the put divided by the (S – Σ e^{-rt}D) · N(-d1) proceeds from the short position. That is,

$$Percent\ Collateral = \frac{e^{-rT} KN(-d2) - (S - \sum_{i=1}^{I} e^{-rt_i} D_i) N(-d1)}{(S - \sum_{i=1}^{I} e^{-rt_i} D_i) N(-d1)} \quad (4)$$

The implicit collateral requirements for puts are shown in Panel B of Table 1. For at-the-money options, the average of the implicit collateral requirement is 0.3933 if the option has more than 180 days to expiration, 0.2363 for options with 30 to 180 days to expiration and 0.1416 for options with 10 to 30 days to expiration. Accordingly, buying puts allows investors to have far lower collateral requirements than in a short sale of stock.

Shorter-term options have greater implicit leverage and lower implicit collateral requirements than longer term options. This suggests that, in addition to preferring to trade options over shares, informed investors prefer to trade shorter maturity options over longer maturity options. A disadvantage of short-term options though is expiration risk. If the option expires before the information is impounded in the stock price, the informed investor will need to buy new options. If the informed investor is unsure when information will arrive or when it will be impounded in the share price, expiration risk can be significant.

To demonstrate the potential costs of expiration risk, we assume that investors pay prices given by the Black-Scholes model for calls and puts when they initiate positions and when they roll over into new options. For the sake of these examples, we assume that the options are at-the-money, the annual riskless rate is 5%, and that the underlying stock does not pay dividends. We calculate option values for annual return standard deviations of 30%, 40%, and 50%. We then use the Black-Scholes prices to compare the costs of buying a one year option with the costs of rolling over one, two, three, four, and six month options for a total holding period of one year. In comparing the costs of long-term options with the cost of rolling over short-term options, we discount the costs of future option expenditures at the riskless rate of interest.

Panel C of Table 1 compares the costs of buying long-term options with the costs of rolling over shorter-term options. Total costs are expressed as a fraction of the price of the underlying stock. The first row presents results for call options when the standard deviation is 0.3. The first number in the row, 0.142, is the cost, as a fraction of the underlying stock price, of buying a 12-month call option. So, if the underlying stock price is \$50, the Black-Scholes value for the call option is 0.142·\$50 = \$7.10. The next column shows the total costs of buying a six month option, and buying a second six month option when the first expires. It is 0.190 or 19% of the share price. Buying a six month call is cheaper than buying a 12 month call, but buying a six month call and then rolling over into a second is more expensive than buying the 12 month call.

Other values in that row indicate that as an investor buys shorter term options and rolls them over more times, the total cost of having a call option for a year increases. If an investor buys options with one month to expiration and rolls them over each month, the total cost is 42.8% of the stock price. That is a little more than three times the cost of buying a 12 month call option. The next two rows present analogous results for options on stocks with higher volatility. Results are similar but, with higher volatility, the costs of rolling over a succession of options is even more expensive relative to the costs of buying a long-term option.

Results for puts are shown in the next three rows. As with calls, rolling over a position multiple times is more expensive than buying a longer term option. For example, if the standard deviation of returns is 0.3 annually, buying a 12 month put will cost 9.4% of the stock price and buying 12 one month puts in succession will cost 38.1% of the stock price, or four times as much.

So, if an option expires before information is impounded in the stock price, an informed investor will need to purchase new options. As we have seen, the cost of buying a series of options can be much greater than the cost of buying one long-term option. Up to this point though, we have only considered the option premiums. Investors also have to pay trading costs.

For each option, we calculate the percentage effective spread on each trade as

$$Effective\ Spread = 2 \cdot \frac{\left| Trade\ Price - \frac{Bid\ Price + Ask\ Price}{2} \right|}{\frac{Bid\ Price + Ask\ Price}{2}}. \quad (5)$$

For each option each day, a weighted average percentage effective spread is calculated using the dollar volume of each trade as its weight. We then calculate an equal-weighted average across contract each day. Panel D shows the means and medians of these averages across stocks and over time.

Three things stand out in Panel D. First, trading costs are high for options. The median effective spread ranges from 4.6% of the bid-ask midpoint for options on large stocks with more than 90 days to expiration to 11.4% for options on all stocks with 10 to 30 days to expiration. If an investor rolls over option positions, he will pay half of the effective spread each time. Second, spreads are lower for options on large stocks. Third, effective spreads are lower for options with greater times to expiration. Bid-ask spreads alone can make it much more expensive to roll over option positions than to buy long-term options.

The implications of Table 1 can be summarized as follows. Informed investors who know when their information will be impounded in prices should buy options with short-times to expiration. Investors get the most leverage with these options. On the other hand, if the informed investor does not know when his information will be reflected in stock prices, there is a strong incentive to buy longer-term options. The leverage is somewhat lower, but still much greater than can be obtained by buying stocks on margin. The cost of purchasing new options if options expire before information is impounded in prices is, however, quite high.

4. Estimating the standardized difference between implied and actual stock prices

Following Manaster and Rendlemann (1982), we jointly estimate the implied stock price and implied volatility using the extended Black Scholes model for a European option on dividend paying stocks. For each option contract (identified by expiration date, strike price, and whether it's a call or put), over each 30 minute interval, we use an iterative process to solve for the values of $S_{Implied}$ and $\sigma_{Implied}$ that minimize the sum of the squared differences between option trade prices (or quote midpoints) and Black-Scholes prices. Specifically, for call options we solve

$$\min_{S_{Implied},\sigma_{Implied}} \sum_{k=1}^{N} \left(\left(S_{Implied} - \sum\nolimits_{i=1}^{I} e^{-rt_i} D_i \right) N(d1) - e^{-rT} K N(d2) - C_k \right)^2, \quad (6)$$

where C_k is the observed call option trade price for option trade k from LiveVol and N is the number of option trades in 30 minute interval. $S_{Implied}$ and $\sigma_{Implied}$ are calculated for put options in an analogous manner. We require that there be at least 3 different option trade prices within a 30 minute interval (in order to uniquely identify the implied stock and volatility) otherwise the interval is discarded.[4]

We then calculate daily relative implied prices ('RIP') using the mean difference between implied prices and stock midpoints in each half hour interval. That is

[4] As a starting point for the iterative process, we use the stock quote midpoint (average of the stock bid and ask) and the option implied volatility (calculated in LiveVol using the CRR tree method adjusted for dividends) of the last trade of the prior 30 minute interval. When solving for $S_{Implied}$ and $\sigma_{Implied}$: we perform the iteration over the log starting values to avoid the situation where the implied stock and implied volatility become negative and we impose the condition in the joint estimation of : $S_{Implied} - \sum_{i=1}^{I} e^{-rt_i} D_i > 0$ since in the expression $d1$ the term $\log\left(\frac{(S_{Implied} - \sum_{i=1}^{I} e^{-rt_i} D_i)}{K}\right)$ results in an error when $S_{Implied} - \sum_{i=1}^{I} e^{-rt_i} D_i < 0$.

$$RIP = \sum_i \frac{(S_{Implied,i} - S_{Midpoint,i})}{S_{Midpoint,i}} / N, \quad (7)$$

where i denotes the half-hour interval, N is the number of 30-minute intervals for which implied prices can be calculated, $S_{Midoint,I}$ is the stock quoted midpoint in the end of 30 min interval, and RIP is the relative implied price measure based on estimation using option trade prices.

Our measure of the implied price difference ('IPD') of stocks, based on option trades and actual stock prices, is obtained by averaging RIP for each day across options on a stock using open interest as weights, and then averaging across all days in a month. IPD is calculated using only 30-minute intervals for which at least three option trade prices are available. Hence, it is only calculated during periods of price discovery. As such, we can think of it as the average percentage difference between option-implied stock prices and actual stock prices during periods of price discovery.

We also calculate RIP and IPD using option bid-ask midpoints at the time of trades rather than option trade prices. IPD calculated from bid-ask midpoints is highly correlated with IPD estimated from trades, but IPD estimated from trade prices is a somewhat stronger predictor of stock prices. Hence we focus on that measure in the empirical work to follow. We will use IPD from bid-ask midpoints to demonstrate the robustness of our results.

Panel A of Table 2 provides summary statistics for IPD across all months and stocks. The first five rows describe IPD for all stocks. When options with all maturities are used to calculate implied prices, the mean IPD is -0.135, indicating that on average, implied stock prices from option trades are 0.135% lower than actual stock prices. The median difference is -0.035. The next four rows describe the distribution if IPD when options with different times to expiration are used to calculate implied stock prices. The mean IPD is -0.075% and the median is -0.016% when options with 10 to 30 days to expiration are used to estimate implied stock prices. As the time to expiration of options used to calculate IPD is increased, the mean IPD decreases steadily and the standard deviation of IPD increases. Summary statistics for IPD for the 500 largest stocks are shown in the next five rows. The standard deviation of IPD is smaller for the 500 largest stocks than for all stocks regardless of the maturities of options used to calculate IPD. IPD is closer to zero for large stocks than for all stocks together.

Panel B of Table 2 reports the distribution of the number of contracts traded per month across stocks and months, for options with different times until expiration. The distribution is right-skewed with very large maximums relative to the means, and means roughly twice the size of medians. There is very little trading in options during some months, while volume is very heavy in others. Mean and median option volume is greater for the 500 largest stocks than for all stocks together. For all stocks, and for the 500 largest stocks, mean and median volume decrease monotonically with the time to expiration.

5. Predicting stock returns with IPD

a. Differences in returns of high and low IPD portfolios

Recall that IPD is calculated using option and stock prices during 30-minute periods when both options and stocks are actively traded. A positive IPD means that implied prices exceed actual prices, suggesting that option traders are bullish relative to stock traders. Each month, we sort all stocks into quintile portfolios based on their IPD. In the low IPD portfolio, implied stock prices are low relative to actual stock prices during the formation month. In the high IPD portfolio, implied stock prices are high relative to actual stock prices during the portfolio formation month. Figure 1 depicts values of IPD for each of the five portfolios, from ten months before through ten months after the portfolio formation. The average IPD of the stocks in the high IPD portfolio increases from around 0% six months before the portfolio formation to a little over 1% in the formation month. That is, implied stock prices exceed actual stock prices by about 1% on average in the formation month for the high IPD portfolio. It drops sharply the next month but remains positive for several months afterwards. The average IPD values of the low IPD portfolio are the mirror image of the values of the high IPD portfolio. They decline over several months before the portfolio formation month and drop sharply to about -1.5% in the formation month. The values then increase slowly over the succeeding months.

Returns and four-factor alphas (Fama-French factors plus momentum) are calculated for each of the five portfolios for each of the three months following portfolio formation. Returns for the five portfolios around the formation month are shown in Figure 2. Of particular interest is that returns for the high IPD portfolio are negative in the formation month, but are large in the months following portfolio formation. For the low IPD portfolio, stock returns are high in the portfolio formation month and low thereafter. Returns and alphas, along with t-statistics that test whether the alphas are different from zero, are reported in Table 3.

Panel A provides results for value-weighted IPD portfolios formed from all stocks. For the first month after portfolio formation, the return for the portfolio with low IPD and hence low implied stock prices relative to actual prices is 0.3543% and the portfolio's alpha is -0.3810%. The t-statistic, based on Newey West adjusted standard errors with 3 lags, for the alpha is -2.33, indicating that the alpha is significantly less than zero. The portfolio with high IPD and hence high implied stock prices earns returns of 0.8748% in the month after portfolio formation. The portfolio's alpha is 0.2553% with a t-statistic of 2.95. So, portfolios with high option-implied stock prices relative to actual prices earn larger returns over the next month. The difference in returns is 0.5205% while the difference in alphas is 0.6364%. The t-statistic for the difference in alphas is a highly significant 3.31.

These results suggest that options contain information that allow investors to earn abnormal returns in the stock market over the next calendar month. It is interesting that both positive and negative abnormal returns are predicted by option implied stock prices. This suggests that short-sale restrictions are not the only reason that informed investors trade in the options market. The difference in returns and difference in alphas is larger in the second and third month after portfolio formation than in the first month afterwards, suggesting that the information in option prices is incorporated into stock prices slowly.

Panel B provides analogous results when equal-weighted portfolios are used. The differences between returns of high and low IPD portfolios and alphas of high and low IPD portfolios are greater than when value-weighted portfolios are used. Likewise, t-statistics for differences in alphas of high and low IPD portfolios are greater than the t-statistics for value-weighted portfolios in Panel A. An important difference between results for equal and value-weighted portfolios is that when equal-weighted portfolios are used, four-factor alphas are significantly negative for the low IPD portfolio, but alphas are not significantly positive for the high IPD portfolio. With equal-weighted portfolios, more weight is given to small stocks. Short-selling is typically more difficult and costly for these stocks than for large stocks. Hence, low equally-weighted IPD portfolios may be especially likely to reflect negative information that is not incorporated into stock prices.

In the remainder of the paper, we focus on value-weighted portfolios. Results from these portfolios are more likely to represent implementable strategies that do not rely on trading of illiquid securities.

In Panel C, value-weighted portfolios are formed from differences between implied and actual stock prices, but now the implied prices are calculated using bid-ask midpoint prices for the options. When the difference between implied and actual stock price is based on trades, a high IPD can be the result of a disproportionate number of trades at the ask price and a low IPD can occur if most trades occur at bid prices. When midpoints are used, the difference between implied and actual stock prices is less clearly a function of order flow. Despite this, returns and alphas for portfolios in Panel B are little changed from those in Panel A.

Panels D & E replicate Panels A & C using the 500 largest stocks based on market capitalization at the beginning of the year. In general, results remain similar when the portfolios are restricted to large stocks, with the results still indicating that the difference between implied and actual stock prices can be used to forecast returns up to three months into the future. T-statistics are slightly lower, but this is not surprising as there are fewer observations. With these large stocks, illiquidity is unlikely to be a problem. Investors are more likely to be able to take advantage of trading opportunities presented by differences between implied and actual stock prices.

Panel F shows cumulative stock returns over the three months since the formation of IPD portfolios. As before, the portfolios are value-weighted. The first three columns of the table report cumulative returns, alphas, and t-statistics for IPD portfolios formed from all stocks and with implied stock prices estimated from trade prices. For the portfolio with the lowest IPD, that is the smallest implied stock prices relative to the actual prices, the return over the next three months is 1.65% and the four-factor alpha is -1.28%. The t-statistic for the alpha is -3.70. In contrast, the portfolio with the highest IPD has a return over the next three months of 3.22% and an alpha of 1.21%. The alpha's t-statistic is 5.34. The difference in returns between the high and low IPD portfolios is 1.56%, and the difference in the four-factor alphas is 2.49%. It appears then, that a strategy of going long a portfolio of stocks with high IPDs and short a portfolio of low IPD stocks produces returns over the next three months of 1.56% and four-factor excess returns of 2.49%.

The next three columns of Panel F show cumulative returns, alphas, and t-statistics for IPD portfolios based on option bid-ask midpoints for the three months after portfolio formation. The last three columns duplicate the first three but form portfolios using only the largest 500 stocks. In both cases, results are similar to the original results. Portfolios of low IPD stocks have low returns and negative alphas in the three months after portfolio formation. Portfolios of high IPD stocks, that is stocks for which the implied price is high relative to the actual price, have high returns and positive alphas in the three months after portfolio formation.

Table 3 shows that both the average returns and average alphas of the high minus low IPD portfolios are positive and significant. They are also, however, highly variable and are negative in many months. Figure 3 shows month-ahead alphas and risk premia, defined as the difference in returns, for the high minus low IPD portfolio each month over 2004 to 2013. Both are variable, particularly over 2008 and 2009. There are months when the risk premia reaches 18% and other months where it falls to -12%. Alphas are less variable but still reach as low as -7%. The strategy of buying high IPD stocks and shorting low IPD stocks is by no means riskless.

b. Predicting stock returns with IPD after adjustment for other factors

We are not the first to show that option prices can predict future stock returns. The literature on predicting stock returns is voluminous, and researchers have identified a number of variables that appear to have predictive power. It is reasonable to ask whether IPD contributes anything to the prediction of stock prices beyond that of other variables, particularly those derived from options. Each month, we regress stock returns on IPD from the previous month and other variables estimated from past data. Following Fama and MacBeth (1973), we calculate the time series average of the coefficients and their t-statistics based on the monthly coefficient estimates. The other variables included in the regression are estimated as follows.

ESS

ESS is the Effective Spread for the Stock. All trades for the stock for the previous month are obtained from TAQ. For each trade, the effective spread is |Trade Price – Bid-Ask Midpoint|/(Bid-Ask Midpoint).
Each day, for each stock, effective spreads for individual trades are averaged, using the dollar value of the trades as weights. A simple average of daily effective spreads is used for the monthly effective spread.

ESO

ESO is the Effective Spread for the Options. For each option contract each day, a weighted-average effective spread is calculated, with the dollar value of the trade serving as its weight. A simple average effective spread is calculated across options each day, and a simple average spread is calculated across days in the month.

Size

Size is the natural logarithm of the market capitalization of the stock. It is estimated as of the end of the previous month.

SOI

SOI is the Stock Order Imbalance, defined as the stock buy volume minus the stock sell volume, divided by the total volume. Individual trades, obtained from TAQ, are signed using the Lee and Ready (1991) algorithm. The SOI is calculated daily and averaged over the month. The SOI from the previous month is used in regressions.

OOI

OOI is the Options Order Imbalance calculated as the difference between the synthetic positive and negative exposure to the underlying stock using the signed CBOE/ISE option volume weighted by the absolute value of the option's delta and scaled by total option volume. It is calculated daily, averaged monthly, and the previous month's OOI is used in regressions.

STD

STD is the Standard Deviation of daily stock returns estimated over the prior month.

O/S

O/S is the natural logarithm of the option to stock volume ratio as in Roll, Schwartz, and Subrahmanyam (2010), Johnson and So (2012), and Ge, Lin and Pearson (2016) where option volume is the total option volume across all strikes and maturities and stock volume is in round lots of 100 to make it comparable to option contracts on 100 shares. With short sale restrictions, investors with negative information are more likely to trade options than stock. Hence a large value of OS may predict low returns. The O/S used in regressions is calculated over the previous month.

BM

BM is the book-to market ratio. Equity book value is obtained from Compustat. It is defined as Total Common Equity (CEQ) plus Deferred Taxes & Invest Tax Credit (TXDITC) (if available) minus Preferred Stock (Redemption (PSTKRV), Liquidating (PSTKL) or Carrying Value (UPSTK), used in that order), divided by the market value of equity at the end of the fiscal year (PRCCF×CSHO).

PIN

PIN is the Probability of Informed trading measure of Easley, Kiefer, O'Hara, and Paperman (1996).
It is computed using the intraday stock trading data from TAQ. We first aggregate the number of market buy and sell orders each day using the Lee and Ready (1991) algorithm to determine the trade direction for each stock trade. The resulting daily buy and sell order counts are used to estimate the probability of information-based trading (PIN). PIN is calculated quarterly using calendar quarters. PIN estimated in the quarter prior to the beginning of the return month is used in regressions.

Turnover

Turnover is the stock trading volume divided by shares outstanding. It is calculated daily and averaged across days of the month prior to the return month.

CW

CW is the difference between call and put implied volatilities as in Cremers and Weinbaum (2012). This is calculated daily for pairs of puts and calls with the same strike price and expiration date. A daily weighted average difference is calculated across option pairs using the open interest as weights. That is, for day t,

$$CW_t = \sum_i w_{i,t}\left(IV_{i,t}^{Call} - IV_{i,t}^{Put}\right), \qquad (8)$$

Where i is the expiration date and strike price combination, the weight $w_{i,t}$ is the average of the put and call open interest, IV^{Call} is the implied volatility of the call and IV^{Put} is the implied volatility of the put. In calculating daily averages, option pairs are omitted if the implied volatility, option delta, open interest or option trade volume are missing, if the best bid or ask quotes are less than or equal to zero or where the ask quote is less than or equal to the bid. We omit options with absolute values of delta greater than 0.98 or less than 0.02. The average daily CW is averaged across days of the month weighting each day equally. CW from the previous month is used in regressions.

CVOL and **PVOL**

CVOL and PVOL are monthly differences in implied volatilities (IVOL) for calls and puts respectively as in An et al JF 2014). The data are from OptionMetrics Volatility surface. We remove observations with missing IVOLs, and deltas. Only ATM series (abs(delta)=0.5) with 30 days to maturity are retained. We use the last available daily observation of the month for each call/put on a share class level to compute the monthly changes.

Skewness

Skewness is the risk neutral skewness computed from the OptionMetrics volatility surface for options with 30 days to maturity. It is the difference between implied volatilities of puts with an absolute value of delta of .2, and the average implied volatility from put and call contracts with an absolute value of delta of 0.5. Xing, Zhang, and Zhao (2010) also refer to this as the option smirk. Skewness is calculated daily and daily values are averaged to calculate a skewness for the month prior to the regression return month.

IVOL

IVOL (defined as in An et al JF 2014), is the average of at the money (|delta|=0.5) call and put Implied Volatilities from the OptionMetrics volatility surface data for series with 30 days to maturity. It is calculated daily and daily values are averaged to produce a monthly value. IVOL used in regressions is estimated over the month prior to the return.

PP

PP is the put-to-call volume ratio as calculated by Pan and Poteshman (2006). That is,

$$PP = \frac{\sum Open\ Buy\ Put}{\sum Open\ Buy\ Put + \sum Open\ Buy\ Call}, \qquad (9)$$

where Open Buy Put is the volume from purchases of puts that open put positions for customers while Open Buy Call is the volume from purchases that open call positions for customers. Where Buy and Sell orders are Open Buy/Sell trading volumes obtained from CBOE/ISE exchanges. We calculate it daily and use the daily average over the month before the return month. PP is only available for 2005-2013.

Correlations between the variables are shown in Table 4. Correlations are estimated using all months. Panel A reports correlations across all stocks while Panel B provides correlations for the 500 largest stocks. It is not surprising that some of these variables are highly correlated. The correlation between PIN and the effective spread for the stock (ESS) is 0.588 when all stocks are used and 0.558 for the largest 500 stocks. The correlation between the implied volatility for calls (CVOL) and the implied volatility for puts (PVOL) is 0.548 for all stocks and 0.835 for the 500 largest stocks. Correlations between IPD and other variables are small for the most part. The exception is the Cremers and Weinbaum (2012) difference between call and put implied volatilities (CW). The correlation between IPD and CW is 0.399 for all stocks and 0.262 for the 500 largest stocks.

Cross-sectional regressions of stock returns on IPD and other explanatory variables from the previous month are run for each month over 2004-2013. Average coefficients and t-statistics based on Newey West adjusted standard errors with three lags are shown in Table 5.

Panel A shows results for all stocks. Given the number of explanatory variables in each regression, it is not surprising that most are insignificant most of the time. Regression (1) includes IPD but no other variables that are derived from option prices. The coefficient on IPD is positive and highly significant, with a t-statistic of 5.58. High implied stock prices relative to actual stock prices are associated with higher stock returns the next month. The coefficient on the effective stock spread is negative and significant. The coefficients on size and the stock order imbalance are respectively negative and positive and only statistically significant at the 10% level. Clearly, IPD has explanatory power for stock returns even after adjusting for these other variables.

The second regression introduces the call-put implied volatility spread (CW). The coefficient on IPD falls to half of its previous value and the coefficient's t-statistic declines to 2.14. The coefficient on the call-put implied volatility spread is 0.0727 with a t-statistic of 5.14. As in Cremers and Weinbaum (2010), portfolios of stocks with large differences between call and put implied volatilities outperform portfolios with smaller or negative differences.

In the following regressions, option-derived additional explanatory variables are added one-by-one. Coefficients on PVOL, Skewness, and OOI are statistically significant and have signs consistent with previous research, but have little effect on the magnitude of the IPD coefficient or its statistical significance. In the sixth regression, all variables are included. In this case, the coefficient on IPD falls to 0.0392, with a t-statistic of only 0.88. Size, skewness, options order imbalance, and the Cremers and Weinbaum difference between call and put volatilities remain statistically significant predictors of returns. Recall that the correlation between the CW measure and IPD is high at 0.4. In the last regression, the CW measure is dropped. The coefficient on IPD increases to 0.1273 and is again statistically significant. Looking just at the regressions in Panel A, it appears that CW is the best and most robust predictor of stock returns. When we look at large stocks though, that is not true.

Regressions in Panel A are replicated in Panel B using just the largest 500 stocks. This is where IPD shines. It is positive and significant in all of the regressions. The effective stock spread coefficient is also positive and significant across regressions, but other variables, including the CW measure, are insignificant. IPD is unique in its ability to forecast returns of large stocks. This is important because it is with the largest stocks that trading rules based on IPD are most likely to be implementable. The coefficient on IPD is 0.5489 in the regression that includes all option-based variables. The standard deviation of IPD is 1.06% for the 500 largest stocks, so a one standard deviation change in IPD leads to a 58 basis point difference in returns the following month.

Implied stock prices, as characterized by our IPD measure, have power to predict stock returns even after including a large number of variables that have previously been shown to have predictive power. What sets IPD apart from other variables derived from option prices is that only IPD predicts returns for large stocks, i.e., that when we consider only the 500 largest stocks, IPD is able to predict returns while other measures do not.

6. Expiration Risk and Informed Trading in Long and Short-term Options

For investors with information, an advantage of trading options rather than stock is that the implicit leverage in options enables the investor to earn larger returns. This leverage is greatest for out-of-the money options and options with short times to expiration. An informed investor who buys short-term options, however, faces expiration risk. Even if the investor is correct in his assessment of a stock's value, his options can expire before the information is impounded in the stock price. As we have seen, IPD predicts returns several months ahead, suggesting that the option traders' information is incorporated in stock prices over a long period of time, and therefore that expiration risk is likely to be important in many cases.

In this section, we study the trade-off between leverage and expiration risk in informed investors' option purchases. While longer-term options have somewhat lower implied leverage, they mitigate expiration risk and allow the investor to avoid paying for additional options if the information is not incorporated in the stock price by the time the short-term option expires. Therefore, if expiration risk is important to informed investors, we should find that IPD calculated from long-term options better predicts long-term returns than IPD from shorter-term options.

6.1 Portfolio returns when IPD is estimated with long-term or short-term options

In Table 6, as in Panel A of Table 3, we form value-weighted portfolios of stocks based on IPD and examine the returns of these portfolios over the next three months. In Table 6 though, we calculate implied stock prices separately from options with various times to expiration. Panel A of Table 6 shows returns and four-factor alphas for portfolios when implied stock prices are estimated from options with ten to 29 days to expiration.[5] In this case, the difference between returns of high and low IPD portfolios over the next month is small and negative rather than positive. The difference between returns of high and low IPD portfolios is also small for the second and third month following portfolio formation. Similarly, the difference between the high and low IPD portfolios' four-factor alphas is statistically insignificant for each of the three months following portfolio formation. Panel B reports returns for portfolios formed using IPD estimated from options with 30 to 59 days to expiration. Here again there is no evidence that portfolios of high IPD stocks outperform portfolios of low IPD stocks. Implied stock prices from short-term options are not useful for predicting stock returns one to three months later.

Panel C shows returns for IPD portfolios formed from implied price estimates from options with 60 to 89 days to expiration. Now four-factor alphas are larger for high IPD portfolios than low IPD portfolios for each of the two months following portfolio formation. The difference in the first month is 40 basis points with a marginally significant t-statistic of 1.78. For the second month, the difference in four-factor alphas is 51 basis points with a t-statistic of 2.25. It is interesting that the difference in the third month is only two basis points. The options used to estimate implied stock prices expire in the third month, so it is possible that investors who believe their information will not be impounded in stock prices over the next two months buy longer-term options.

Panels D and E report returns for IPD portfolios when the implied prices are calculated using options with 90 to 179 days to expiration, and with more than 180 days to expiration. In each case, high IPD portfolios have larger returns and larger four-factor alphas than low IPD portfolios for each of the three months after portfolio formation. When options with at least 180 days to expiration are used to form portfolios, the difference in raw returns between the high and low IPD portfolios is 75 basis points in the first month, 112 basis points in the second month, and 77 basis points in the third month following portfolio formation. The difference in four-factor alphas is 82 basis points in the first month, 129 basis points in the second month, and 101 basis points in the third month. Each of these differences is statistically significant, as are all the differences in alphas when options with 90 to 179 days are used to calculate implied stock prices.

[5] We exclude options with less than 10 days to expiration to avoid roll-over trades.

The lesson of Table 6 is that implied stock prices from longer-term options contain information about stock returns over the next one to three months, but implied stock prices from options with short times to expiration do not. In one way, this seems surprising, as options with shorter times to expiration have, all else equal, more implicit leverage and would thus be more attractive to informed investors. On the other hand, if there is substantial uncertainty about when information will be impounded in stock prices, informed investors can avoid expiration risk with longer-term options. Therefore, the results in Table 6 indicate that expiration risk is sufficiently important to overcome the disadvantage of lower leverage of longer-term options vis-à-vis short-term options.

Predictive regression estimates with the next month's return as the dependent variable and the full range of explanatory variables are presented in Table 7. These regressions are the same as the ones reported in Table 5, but here IPD is calculated from options with maturities that are greater than 90 days. Panel A describes the results for predictive regressions using all stocks with options. In each specification, regardless of which option-based variables are included, the coefficient on IPD is positive and significant. With the restriction on the time to expiration of options used to calculate IPD, the number of stocks in the monthly regressions decreases by about 10%. Despite this, the t-statistics for IPD are higher than the t-statistics from the equivalent regressions in Table 5. Implied prices from longer-term options appear to contain more information about future stock returns than implied stock prices from short-term options.

The CW and options order imbalance (OOI) measures are also statistically significant predictors of the next month's stock returns. These two variables, and IPD, all have similar t-statistics when options with times to expiration of more than 90 days are used to calculate IPD. When long-term options are used to calculate IPD, the CW measure is no longer the best predictor of stock returns for all stocks.

Panel B of Table 7 presents regression estimates for the next months' returns of the 500 largest stocks when IPD is estimated using only long-term options. Coefficients on IPD are now more significant than in Table 5 where all options are used to estimate IPD. Regressions in Panel B again show that IPD predicts returns on large stocks while other options based measures, including CW, do not. This is a significant advantage of the IPD measure.

We have shown thus far that IPD estimated from options of all expiration dates predicted returns up to three months in the future. If investors are concerned that options will expire before information is incorporated in stock prices, they may be more inclined to buy longer-term options. It seems that this would be especially true of information that was, ex-post, incorporated in stock prices long after the informed invested. Hence, it seems likely that stock returns more than three months out are much more likely to be predicted by longer term options.

We rerun regressions of stock returns on IPD and other variables using IPD estimated from options with more than 90 days to expiration and options with between 10 and 60 days to expiration. These regressions are run for each of the 14 months after estimation of IPD. Results are presented in Table 8.

To save space, we report just the coefficients on IPD and their t-statistics. All regressions include the full set of standard controls, i.e., ESS, Size, SOI, R_t, R_{t-1}, MOM, σ_{stock}, ESO, O/S, B/M, turnover, PIN, CW, CVOL, PVOL, Skewness, IVOL, PP, and OOI. The first two columns report regression results when IPD is estimated from options with more than 90 days to expiration. In the first column, when all stocks are included, the coefficient on IPD is positive for all months except t+11 and t+12, and significant for 1, 2, 3, 5, 7, 8, 9 and 10 months after IPD is estimated. None of the coefficients are significant after 10 months. The second column reports results when options with more than 90 days to expiration are used to calculate IPD, but only the 500 largest stocks are included in the regression. In the regressions with the largest stocks, coefficients are always positive and are generally larger and more significant than when all stocks are used. The last statistically significant coefficient is at t+12.

The last two regressions report results when IPD is estimated using options with 10 to 60 days to expiration. When all stocks are included, there are a couple of months when the coefficient on IPD is positive and significant, but results in general are weak. In the last regression, where IPD is calculated using options with 10 to 60 days to expiration and only the 500 largest stocks are included, the coefficient is only marginally significant in one month, t+13, and it has the wrong sign.

To summarize, even when other option-based variables are included, IPD calculated from long-term options can predict returns as far as 10 months in the future. This seems like a long time, but Xing, Zhang, and Zhao (2010) find that skewness predicts returns for up to six months. On the other hand, IPD calculated from short-term options seems to have little power to predict future stock returns. Investors who trade long-term options appear to have information that is only incorporated in stock prices four or more months in the future. If informed investors know that it may take that long for information to be impounded in prices, they will prefer to trade long-term options.

The excess returns to be earned from using IPD are modest. As Table 2 shows, the standard deviation of IPD for the 500 largest stocks is 1.33% when calculated from options with at least 90 days to expiration. The sum of the coefficients on IPD for all 14 months is 1.9635. So, this suggests that for the largest stocks, a one standard deviation increase in IPD increases returns by 2.61% over the next 14 months. And, as Figure 3 shows, the difference in returns between high and low IPD portfolios varies significantly month to month.

6.2 Shareholder horizons and predictability of stock returns with IPD

Informed investors may trade long-term options when there is uncertainty as to when information will be incorporated into stock prices. This is reflected in the power of IPD calculated from long-term options to predict stock returns up to ten months in the future, while IPD from short-term options has little predictive power. Some investors who are uncertain as to when information will be impounded in stock prices may instead buy or sell the underlying shares. These investors include financial institutions that are constrained to holding equity rather than options.

In this section, we examine the predictive power of IPD for stocks that are held heavily by institutions that invest long-term versus those that are held by institutions that turn over their portfolios frequently. The idea is that if a stock is already being held by investors that purchase for the long-term, information that may not be revealed publicly for several months may still be reflected in prices. On the other hand, short-term investors may avoid purchasing stock's on the basis of information that may not be revealed publicly until after the end of their holding period.

Institutional investor turnover is measured as in Gaspar, Massa, and Matos (2005). Investor holdings come from 13-F filings with the Securities and Exchange Commission which are assembled by CDA/Spectrum. For each month, institutions are sorted into terciles based on the turnover of their portfolio over the previous quarters. Institutional investors in the tercile with the lowest turnover rates are categorized as long-term institutional investors. Each month, the proportion of each stock's shares that are held by long-term institutions is calculated. Stocks are then separated into terciles by the portion of their shares held by long-term institutions.[6]

Table 9 reports month-ahead four-factor alphas across IPD quintiles for each tercile of long-term holdings. IPD is calculated using options with at least 90 days to expiration. Panel A shows results when all stocks are included. For the tercile of stocks with low long-term investment, IPD has a big impact on returns. The difference between the month-ahead four-factor alpha for high IPD stocks and the alpha for low IPD stocks is 1.31%. For the middle and high long-term investment portfolios, the differences in alphas across IPD quintiles are small and statistically insignificant. Panel B reports results for the 500 largest stocks. The difference between alphas of high and low IPD quintiles is 1.22% for the low long-term institutional holdings tercile, 0.77% for the middle tercile, and 0.28% for the high tercile. The difference are statistically significant for the low and middle terciles, but not for the tercile of stocks with high long-term institutional holdings.

[6] We are grateful to Ankur Pareek for providing us with this data.

The results in Table 9 show that IPD calculated from long-term options contains information on stock returns if long-term institutional shareholders do not own very large portions of the shares. Institutional investors with long-horizons will be willing to buy shares even if it takes a long time for their information to be incorporated into stock prices. They may bid up prices long before information about the stock is made public.

We next regress stock returns on dummy variables for above and below average holding of the stock by long-term institutional investors, and interactions between these dummy variables and IPD. We also include the option-based variables and other predictive variables used in previous tables. We run separate regressions for the first and second months after portfolio formation for all stocks, and for the largest 500 stocks. Regression estimates are reported in Table 10.

The first regression has the next month's return as the dependent variable and includes all stocks. The coefficient on the interaction between IPD and the dummy variable for below average long-term institutional holdings is 0.1580 with a t-statistic of 3.96. Even after including all of these other variables, IPD can predict stock returns when long-term institutional holdings are small. On the other hand, the coefficient on the interaction of IPD and above average long-term holdings is only 0.0299 with a t-statistic of 0.64. IPD does not predict stock returns when a large proportion of the shares are held by long-term institutional investors. The next column provides regression results for the second month when all stocks are included. The coefficients on the interaction terms are both positive but statistically insignificant.

The last two columns of the table report regressions using only the 500 largest stocks. For the first month, coefficients are positive and significant for both the interaction between IPD and a dummy for above average long-term holdings and the interaction between IPD and the dummy for below average holdings by long-term investors. The coefficient is larger and more significant though when IPD is interacted with the dummy for below average ownership by long-term institutions. The last regression uses the largest 500 stocks and has the returns in the second month after IPD estimation as the dependent variable. Here, the interaction between IPD and the dummy for below average holdings by long-term investors is 0.4251 with a t-statistic of 4.44. The interactions between IPD and the dummy for above average holdings by long-term investors is insignificant.

After adjusting for numerous other explanatory variables, IPD predicts stock returns for stocks with below average holdings by long-term investors. These findings support our previous evidence that option prices reflect information that is only reflected in stock prices in the long-run. Investment by long-term institutional investors appears, however, to reduce the returns available for options investors.

7. IPD, implied volatilities and stock return predictability

The percentage difference between stock prices implied from options and actual stock prices (IPD) predicts stock returns several months into the future. This is consistent with informed investors trading options rather than stock because of the options' implicit leverage. We expect that the predictive power of IPD should be especially great during periods of uncertainty, or when implied volatilities are high.

Recall that we calculate IPD and implied volatilities jointly. This is important because high option prices relative to stock prices can be due to either high volatility or high implied prices. So, if calculated separately, we would expect them to be related mechanically. Using our jointly estimated implied volatilities and IPDs, we sort stocks into fifteen portfolios for each month over 2004-2013 based on IPD quintiles and implied volatility terciles. We then calculate the next month's value-weighted return for each portfolio.

Table 11 shows mean next month four-factor alphas for these portfolios. Panel A reports results when all stocks are included in the portfolios. Portfolios of high IPD stocks outperform portfolios of low IPD stocks for each of the three implied volatility terciles, and the differences are statistically significant at the 1% level for the middle and high implied volatility terciles. Of particular interest though is that the difference in returns between high and low IPD stocks increases monotonically as we go from low to high implied volatility terciles. The difference in returns between high and low IPD portfolios over the next month is 34.46 basis points when implied volatilities are low, 51.49 basis points for implied volatilities in the middle tercile and 102.89 basis points for high implied volatility stocks. It is when uncertainty is greatest that IPD has the greatest predictive power for risk-adjusted returns. Panel B provides results for portfolios formed from the largest 500 stocks. Results are similar. Differences in four-factor alphas between high and low IPD quintile portfolios increase monotonically with implied volatilities. For the low volatility tercile, the next month's average difference in alphas is 49.39 basis points. The average difference increases to 76.57 basis points for the middle volatility tercile and to 90.39 basis points for the high volatility tercile.

Results in Table 11 indicate that differences in IPD predict especially large differences in four-factor alphas for stocks with high implied volatilities. We next see if this is true for returns after including other variables that predict stock returns. We regress each stock's return each month over 2005-2013 on dummy variables for their implied volatility tercile in the previous month, on interactions between IPD and the volatility dummies, and on the numerous predictive variables used elsewhere in the paper, i.e. ESS, Size, SOI, R_t, R_{t-1}, MOM, σ_{stock}, ESO, O/S, B/M, turnover, PIN, CW, CVOL, PVOL, Skewness, IVOL, PP, and OOI. We would like to determine if the relation between IPD and future returns is strongest when volatility is high after adjustment for these other variables. We note though, that the interaction of implied volatility and IPD is likely to be correlated with CW and other explanatory variables.

The regression estimates are shown in Table 12. In the first regression, all stocks are included and the return one month ahead is the dependent variable. The interaction between IPD and the high implied volatility tercile is 0.0610 with a t-statistic of 1.60. The interaction between IPD and the middle volatility portfolio is 0.1629 with a t-statistic of 2.52. The second regression includes all stocks and has the return for the second month as the dependent variable. All of the interactions between IPD and implied volatility terciles are positive, but none are statistically significant.

The next two columns show results when only the largest 500 stocks are included in the regressions. When the next month's return is the dependent variable, the coefficient on the interaction between IPD and the high implied volatility tercile is 0.3859 with a t-statistics of 3.46. The coefficient on the interaction between IPD and the middle volatility tercile dummy is 0.3439 with a t-statistic of 2.19. For the interaction between IPD and the lowest implied volatility tercile, the coefficient is 0.1881 with a marginally significant t-statistic of 1.84. For the largest 500 stocks, IPD predicts particularly large returns the next month if implied volatilities are large. In the last regression, the dependent variable is the return in the second month after portfolio formation. Coefficients on the interactions between IPD and middle and high implied volatility terciles are positive and significant.

As a whole, the results of Table 12 provide some confirmation of the findings in Table 11 that IPD predicts larger stock returns when volatility is high. In Table 12, in all of the regressions, the coefficient on the interaction between IPD and the lowest volatility tercile is always smaller than the interactions between IPD and the other volatility terciles.

8. **Conclusions**

In this paper, we show that IPD, the percentage difference between implied stock prices from options and actual stock prices, can predict stock returns several months in the future. This is consistent with informed investors trading options rather than stock. We observe that IPD predicts both positive and negative returns. This suggests that informed investors trade options to take advantage of the options' leverage, and not just to circumvent short-sale costs and restrictions. As we would expect if informed traders are trading options, IPD has the most predictive power for returns when implied volatilities are greatest.

A striking finding of this paper is that IPD estimated from options with at least 90 days to expiration predicts stock returns for several months while IPD estimated from options with less than 60 days to expiration has no predictive power. This suggests that informed investors trade long-term options because they do not know when their information will be reflected in stock prices, and bear expiration risk if they buy short-term options.

Recent studies show that several variables derived from implied volatilities of options or from option trading volume can predict stock returns. When calculated from options with at least 90 days to expiration, IPD predicts stock returns for all stocks as well as any of these other measures. When attention is restricted to the largest 500 stocks, IPD easily outperforms other measures in predicting returns. This is important because return predictability for small stocks can be the result of market frictions. Stock return predictions are more likely to lead to implementable trading strategies when the 500 largest stocks are used.

Table 1.
Advantages and disadvantages of trading options rather than stock for informed investors

Panel A. Mean leverage for call options. Implicit leverage is calculated from the Black-Scholes model as

$Leverage\ = [e^{-rT}KN(d2)]/[(S-PVDIV)*N(d1)]$ where $d1 = \frac{\log\left(\frac{(S-PVDIV)}{K}\right)+(r+\sigma^2/2)*T}{\sigma\sqrt{T}}$ $and\ d2 = d1-\sigma\sqrt{T}$, S is the stock price, $PVDIV$ is the present value of dividends paid over the life of the option, K is the strike price, r is the risk free interest rate, σ is the stock volatility, T is the time to expiration, and N(.) is the cumulative normal distribution.

	≤ 30 Days to Expiration	30 – 179 Days to Expiration	≥ 180 Days to Expiration
In the Money	0.8450	0.7662	0.6570
At the Money	0.8835	0.8310	0.7678
Out of the Money	0.9154	0.8809	0.8393

Panel B. Mean put collateral requirements. The implicit collateral for puts is

$$Percent\ Collateral = \frac{e^{-rT}KN(-d2)-(S-PVDIV)N(-d1)}{(S-PVDIV)N(-d1)}$$

In the Money	0.1970	0.3782	0.6180
At the Money	0.1416	0.2363	0.3933
Out of the Money	0.1183	0.2037	0.3595

Panel C. Costs of buying 12 month options versus the cost of rolling over a succession of shorter term options. Prices for at-the-money options on non-dividend paying stocks are estimated using the Black-Scholes model with a risk-free interest rate of 5%. To get the costs of purchasing options when rolling over positions in the future, Black-Scholes prices are discounted at a 5% annual rate. Values in the table are the proportion of the stock price paid for a 12 month option or a succession of shorter-term options.

Option Exp.	12 Months	6 Months	4 Months	3 Months	2 Months	1 Month
Purchases	1	2	3	4	6	12
			Calls			
σ = 0.3	0.142	0.190	0.227	0.258	0.312	0.428
σ = 0.4	0.180	0.245	0.294	0.336	0.406	0.563
σ = 0.5	0.218	0.299	0.361	0.413	0.501	0.698
			Puts			
σ = 0.3	0.094	0.142	0.179	0.210	0.262	0.381
σ = 0.4	0.132	0.196	0.246	0.287	0.357	0.516
σ = 0.5	0.169	0.250	0.312	0.364	0.451	0.648

Panel D. Effective Spreads of Options

Using all equity options over 2004-2013, we calculate the effective spread for each trade as

$$Effective\ Spread = 2 \cdot \frac{|Trade\ Price - (Bid\ Price + Ask\ Price)/2|}{(Bid\ Price + Ask\ Price)/2}$$

On every day that an option trades, a weighted average effective spread is calculated weighting each trade by its dollar volume. We then calculate an equal-weighted average across contracts each day. This table shows the means and medians of these averages across stocks and over time.

	All Stocks		Largest 500	
	Mean	Median	Mean	Median
10 to 30 days	0.143	0.114	0.099	0.085
31 to 60 days	0.117	0.088	0.076	0.062
61 to 90 days	0.098	0.069	0.065	0.049
More than 90 days	0.098	0.070	0.061	0.046

Table 2.
Summary Statistics.
Panel A. Monthly summary statistics for percentage differences between implied stock prices from option trade prices and actual stock prices (IPD).

	Mean	Median	Std. Dev.	Minimum	Maximum
All Stocks with Options					
Implied Price Difference All Maturities	-0.135	-0.035	1.952	-96.526	78.817
Implied Price Difference 10-30 Days	-0.075	-0.016	1.042	-49.452	78.817
Implied Price Difference 31-60 Days	-0.076	-0.006	1.201	-52.236	52,762
Implied Price Difference 61-90 Days	-0.082	0.000	1.573	-34.607	56.046
Implied Price Difference > 90 Days	-0.208	-0.074	2.310	-82.231	83.643
500 Largest Stocks					
Implied Price Difference All Maturities	-0.001	-0.000	1.058	-48.857	53.309
Implied Price Difference 10-30 Days	0.010	0.005	0.543	-16.648	45.057
Implied Price Difference 31-60 Days	0.025	0.017	0.644	-28.777	52.762
Implied Price Difference 61-90 Days	0.050	0.032	0.938	-34.607	52.849
Implied Price Difference > 90 Days	-0.034	-0.033	1.330	-35.085	53.445

Panel B. Distribution of number of contracts traded per month.

	Mean	Median	Std. Dev.	Minimum	Maximum
	All Stocks with Options				
≥ 10 and ≤ 30 Days to Expiration	482	226	1,334	1	147,196
> 30 and ≤ 60 Days to Expiration	378	194	893	0	104,670
> 60 and ≤ 90 Days to Expiration	335	154	892	1	100,634
> 90 Days to Expiration	288	142	585	0	43,873
	500 Largest Stocks				
≥ 10 and ≤ 30 Days to Expiration	806	385	2,044	1	127,942
> 30 and ≤ 60 Days to Expiration	552	293	1,369	0	104,670
> 60 and ≤ 90 Days to Expiration	442	212	1,263	1	100,634
> 90 Days to Expiration	366	197	719	3	43,873

Table 3.

Returns for the three months following portfolio formation for IPD (Implied Price Difference from Trades) based portfolios. Returns and α's are in percentage terms. The Fama-French-Carhart four factor model is used to calculate α's. The t-statistics are based on Newey-West adjusted standard errors with three lags.

Panel A. Value-weighted portfolios formed from all stocks with options. The implied price difference is calculated using option trade prices.

Port.	R_{t+1}	α_{t+1}	T Stat.	R_{t+2}	α_{t+2}	T Stat.	R_{t+3}	α_{t+3}	T Stat.
Low	0.3543	-0.3810	-2.33	0.3887	-0.4061	-2.34	0.4275	-0.4094	-1.92
2	0.4851	-0.1244	-1.25	0.4884	-0.1466	-1.20	0.4833	-0.2030	-1.89
3	0.4759	-0.0934	-1.51	0.4651	-0.1472	-2.09	0.5602	-0.1160	-1.76
4	0.7056	0.1628	2.48	0.7303	0.1618	1.82	0.7514	0.1572	1.81
High	0.8748	0.2553	2.95	1.0216	0.3875	3.75	1.0427	0.4299	3.52
H-L	0.5205	0.6364	3.31	0.6329	0.7936	3.59	0.6151	0.8392	2.89

Panel B. Equal-weighted portfolios formed from all stocks with options. The implied price difference is calculated using option trade prices.

Port.	R_{t+1}	α_{t+1}	T Stat.	R_{t+2}	α_{t+2}	T Stat.	R_{t+3}	α_{t+3}	T Stat.
Low	0.2639	-0.6168	-4.86	0.3078	-0.6199	-4.74	0.5320	-0.3921	-2.75
2	0.7817	0.0174	0.24	0.7619	-0.0459	-0.69	0.7586	-0.0637	-0.92
3	0.7289	0.0032	0.04	0.7647	-0.0049	-0.06	0.7220	-0.0391	-0.44
4	0.7757	0.0374	0.40	0.8700	0.0897	0.89	0.8493	0.0740	0.78
High	1.0051	0.1566	1.53	0.9691	0.1068	1.02	0.8967	0.0356	0.39
H-L	0.7412	0.7734	5.91	0.6613	0.7268	5.16	0.3647	0.4277	3.05

Panel C. Value-weighted portfolios formed from all stocks with options. The implied price difference is calculated using option bid-ask midpoint prices.

Port.	R_{t+1}	α_{t+1}	T Stat.	R_{t+2}	α_{t+2}	T Stat.	R_{t+3}	α_{t+3}	T Stat.
Low	0.4487	-0.2701	-1.65	0.3263	-0.4641	-2.97	0.4278	-0.4221	-2.21
2	0.4272	-0.1825	-1.95	0.4946	-0.1254	-1.04	0.5085	-0.1679	-1.61
3	0.5205	-0.0397	-0.72	0.4867	-0.1137	-1.80	0.5233	-0.0828	-1.37
4	0.7052	0.1596	2.17	0.8087	0.2274	2.83	0.7925	0.1967	2.51
High	0.8950	0.2272	1.62	0.9381	0.2825	2.74	1.0115	0.3537	2.44
H-L	0.4463	0.4973	2.23	0.6118	0.7466	3.84	0.5836	0.7758	2.68

Panel D. Value-weighted portfolios formed from the 500 largest stocks. The implied price difference is calculated using option trade prices.

Port.	R_{t+1}	α_{t+1}	T Stat.	R_{t+2}	α_{t+2}	T Stat.	R_{t+3}	α_{t+3}	T Stat.
Low	0.3862	-0.2191	-1.55	0.2575	-0.3685	-2.03	0.3635	-0.3372	-2.03
2	0.5360	-0.0090	-0.09	0.3717	-0.2234	-1.69	0.4408	-0.1900	-1.48
3	0.4428	-0.1133	-1.26	0.5648	-0.0237	-0.28	0.5727	-0.0338	-0.23
4	0.7087	0.1947	2.15	0.7503	0.2169	1.80	0.7929	0.2408	2.34
High	0.8391	0.3004	2.47	0.9239	0.3675	2.93	0.9712	0.4169	3.63
H-L	0.4529	0.5195	2.68	0.6664	0.7360	2.83	0.6077	0.7514	3.61

Panel E. Value-weighted portfolios formed from the largest 500 stocks. The implied price difference is calculated using option bid-ask midpoint prices.

Port.	R_{t+1}	α_{t+1}	T Stat.	R_{t+2}	α_{t+2}	T Stat.	R_{t+3}	α_{t+3}	T Stat.
Low	0.3985	-0.2102	-1.53	0.3500	-0.2608	-1.46	0.3885	-0.3028	-1.90
2	0.4274	-0.1097	-1.03	0.4240	-0.1769	-1.25	0.5608	-0.1384	-1.00
3	0.4710	-0.0841	-1.03	0.5096	-0.0789	-0.74	0.5575	-0.0205	-0.19
4	0.8924	0.3876	4.16	0.8539	0.3284	3.02	0.7581	0.1922	1.64
High	0.7263	0.1792	1.36	0.7979	0.2288	1.95	1.0525	0.5061	4.25
H-L	0.3278	0.3894	1.99	0.4478	0.4896	1.93	0.6640	0.8099	4.01

Panel F. Cumulative returns of value-weighted portfolios in percents, months t+1 through t+3

	All Stocks, Trade Prices			All Stocks, Midpoints			500 Large Stocks, Trades		
Port.	$R_{t+1,t+3}$	$\alpha_{t+1,t+3}$	T Stat.	$R_{t+1,t+3}$	$\alpha_{t+1,t+3}$	T Stat.	$R_{t+1,t+3}$	$\alpha_{t+1,t+3}$	T Stat.
Low	1.6539	-1.2834	-3.70	1.6661	-1.2773	-3.80	1.4419	-0.9899	-3.13
2	1.8047	-0.5201	-2.48	1.7861	-0.5138	-1.97	1.6877	-0.3335	-1.31
3	1.7557	-0.3400	-2.58	1.8370	-0.2044	-1.53	1.8812	-0.1475	-0.73
4	2.4998	0.5270	3.44	2.6283	0.6587	4.73	2.5651	0.6870	3.10
High	3.2174	1.2088	5.34	3.0972	0.8926	4.00	3.0227	1.2896	6.29
H-L	1.5635	2.4922	5.83	1.4311	2.1699	5.29	1.5808	2.2796	5.40

Table 4. Correlations of variables. IPD is the percentage difference between the implied and actual stock price, ESS is the volume-weighted average effective spread of the stock, ESO is the volume-weighted average effective spread of the stock, Size is the natural logarithm of the market capitalization, SOI (stock order imbalance) is the stock buy volume minus the stock sell volume divided by total volume, OOI is the options order imbalance calculated as the difference between synthetic positive and negative option volume, σ_{Stock} is the standard deviation of the daily stock return over the month, O/S is the natural logarithm of the option to stock volume ratio, PIN is the probability of informed trading, Turn is the mean daily turnover of the stock over the month, CW is the difference between call and put implied volatilities as in Cremers and Weinbaum (2012), CVOL is the implied volatility of a 30 day call with delta = 0.5, PVOL is the implied volatility of a 30 day put with delta = -0.5, Skew is the difference between the implied volatility of a put with a delta =-0.2 and the average implied volatility of puts and calls with deltas with values |0.5|, IVOL is the average implied volatility of at the money calls and puts with 30 days to maturity and PP is the ratio of volume from buys to open put positions to buys to open both put and call positions.

Panel A. All stocks.

	IPD	ESS	ESO	Size	SOI	OOI	σ_{Stock}	O/S	PIN	Turn	CW	CVOL	PVOL	Skew	IVOL	PP
IPD	1.000															
ESS	-0.106	1.000														
ESO	-0.031	0.248	1.000													
Size	0.021	-0.107	-0.220	1.000												
SOI	0.024	-0.238	-0.107	0.057	1.000											
OOI	0.045	0.008	0.028	0.006	-0.022	1.000										
σ_{Stock}	-0.105	0.459	0.094	-0.103	-0.105	-0.024	1.000									
O/S	-0.064	0.009	-0.196	0.196	0.014	-0.092	0.008	1.000								
PIN	-0.066	0.588	0.290	-0.202	-0.272	0.006	0.199	-0.012	1.000							
Turn	-0.054	-0.168	-0.157	0.009	0.066	-0.083	0.240	0.129	-0.198	1.000						
CW	0.399	-0.007	0.001	0.024	-0.016	0.100	-0.071	-0.137	-0.039	-0.102	1.000					
CVOL	0.019	0.010	0.010	0.000	-0.032	0.021	0.025	0.003	0.004	-0.016	0.037	1.000				
PVOL	-0.003	0.018	0.009	0.001	-0.020	0.007	0.036	0.004	0.004	-0.006	0.007	0.548	1.000			
Skew	-0.111	-0.026	0.193	-0.068	0.006	-0.052	0.017	-0.018	0.056	-0.019	-0.226	-0.033	-0.014	1.000		
IVOL	-0.157	0.494	0.147	-0.207	-0.214	-0.024	0.656	0.053	0.323	0.234	-0.178	0.259	0.260	0.009	1.000	
PP	0.020	0.007	0.274	-0.169	-0.021	-0.055	-0.100	-0.205	0.072	-0.239	0.027	0.018	0.019	0.167	-0.091	1.000

Panel B. Largest 500 stocks

	IPD	ESS	ESO	Size	SOI	OOI	σ_{Stock}	O/S	PIN	Turn	CW	CVOL	PVOL	Skew	IVOL	PP
IPD	1.000															
ESS	-0.049	1.000														
ESO	-0.025	0.132	1.000													
Size	-0.002	-0.098	-0.234	1.000												
SOI	-0.015	-0.278	0.056	-0.010	1.000											
OOI	0.029	-0.014	0.022	0.026	-0.028	1.000										
σ_{Stock}	-0.067	0.258	-0.041	-0.139	-0.117	-0.066	1.000									
O/S	-0.039	-0.053	-0.322	0.303	0.024	-0.010	0.023	1.000								
PIN	-0.017	0.558	0.187	-0.185	-0.207	-0.011	0.187	-0.051	1.000							
Turn	-0.046	-0.074	-0.149	-0.102	-0.019	-0.042	0.469	0.180	-0.011	1.000						
CW	0.262	-0.070	-0.053	0.040	-0.023	0.098	-0.107	-0.100	-0.058	-0.088	1.000					
CVOL	0.010	0.033	0.032	-0.002	-0.030	0.052	0.121	-0.002	0.037	0.037	0.049	1.000				
PVOL	-0.012	0.028	0.026	-0.002	-0.024	0.040	0.128	-0.008	0.030	0.037	0.034	0.835	1.000			
Skew	-0.072	0.115	0.223	-0.111	-0.076	-0.099	0.404	-0.078	0.130	0.174	-0.236	0.009	0.008	1.000		
IVOL	-0.071	0.352	-0.072	-0.183	-0.109	-0.053	0.839	0.047	0.169	0.557	-0.151	0.259	0.258	0.393	1.000	
PP	0.006	-0.059	0.355	-0.228	-0.006	-0.116	-0.099	-0.232	0.101	-0.178	-0.008	0.020	0.017	0.201	-0.115	1.000

Table 5. Regressions of stock returns for month t+1 on explanatory variables from month t. Each month over 2004 – 2013, cross sectional regressions of stock returns are run on explanatory variables from the previous month. As in Fama and MacBeth (1973), coefficients are averaged across months and t-statistics are based on the Newey-West three lag adjusted standard errors of coefficient estimates across months. When PP or OOI is included, the sample period is restricted to 2005-2013.

Panel A. All stocks.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
IPD	0.1500***	0.0607**	0.1435***	0.1335***	0.1348***	0.0392	0.1277***
	(5.58)	(2.14)	(4.74)	(4.80)	(4.48)	(0.88)	(2.70)
ESS	-1.5270**	-1.1869*	-1.6682**	-1.7470**	-1.2076*	-0.8692	-1.2971
	(-2.18)	(-1.68)	(-2.37)	(-2.51)	(-1.68)	(-0.95)	(-1.37)
Size	-0.0012*	-0.0013*	-0.0013*	-0.0013**	-0.0011	-0.0017**	-0.0018**
	(-1.89)	(-1.95)	(-1.95)	(-2.02)	(-1.65)	(-2.21)	(-2.42)
SOI	0.0118*	0.0127**	0.0123*	0.0133**	0.0113	0.0017	0.0024
	(1.80)	(1.99)	(1.84)	(2.00)	(1.64)	(0.19)	(0.28)
R_t	-0.0026	-0.0015	-0.0040	-0.0024	0.0009	0.0028	0.0021
	(-0.31)	(-0.18)	(-0.48)	(-0.29)	(0.10)	(0.33)	(0.25)
R_{t-1}	-0.0034	-0.0023	-0.0038	-0.0024	-0.0000	-0.0012	-0.0013
	(-0.44)	(-0.30)	(-0.49)	(-0.32)	(-0.00)	(-0.14)	(-0.16)
MOM	-0.0026	-0.0026	-0.0026	-0.0026	-0.0029	-0.0011	-0.0010
	-0.60)	(-0.61)	(-0.60)	(-0.59)	(-0.61)	(-0.27)	(-0.23)
σ_{Stock}	-0.0743	-0.0844	-0.0538	-0.0690	-0.0553	-0.0747	-0.0624
	(-1.19)	(-1.35)	(-0.91)	(-1.12)	(-0.88)	(-1.27)	(-1.08)
ESO	-0.0040	-0.0032	-0.0038	0.0008	-0.0059	-0.0052	-0.0051
	(-0.49)	(-0.39)	(-0.46)	(0.10)	(-0.68)	(-0.42)	(-0.41)
OS	-0.0036	-0.0003	-0.0032	-0.0019	-0.0064*	-0.0011	-0.0027
	(-0.83)	(-0.06)	(-0.75)	(-0.44)	(-1.67)	(-0.26)	(-0.65)
B/M	0.0017	0.0016	0.0016	0.0018	0.0011	0.0009	0.0011
	(1.16)	(1.07)	(1.09)	(1.24)	(0.70)	(0.60)	(0.72)
Turnover	-0.0005*	-0.0003	-0.0005*	-0.0004	-0.0005	-0.0083	-0.0108
	(-1.68)	(-1.13)	(-1.81)	(-1.58)	(-1.64)	(-0.69)	(-0.90)
PIN						-0.0005	-0.0005*
						(-1.51)	(-1.75)
CW		0.0727***				0.0582***	
		(5.14)				(3.39)	
CVOL			0.0093			0.0072	0.0085
			(1.44)			(0.77)	(0.90)
PVOL			-0.0196***			-0.0123	-0.0127
			(-3.24)			(-1.58)	(-1.61)
Skewness				-0.0410***		-0.0332**	-0.0495***
				(-3.55)		(-2.15)	(-3.27)
IVOL						-0.0052	-0.0073
						(-0.40)	(-0.56)
PP						-0.0030	-0.0028
						(-1.29)	(-1.19)
OOI					2.1602***	1.5883***	1.7714***
					(3.72)	(2.71)	(2.98)
Adj R^2	0.0591	0.0597	0.0617	0.0605	0.0581	0.0820	0.0797
N Stocks	1,563	1,554	1,561	1,563	1,583	1,270	1,270

Table 5 Panel B. 500 Largest stocks, based on capitalization at the end of the previous month. ***denotes significance at the 1% level (two-tailed test), **denotes significance at the 5% level and *is used to denote significance at the 10% level.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
IPD	0.4034***	0.4897***	0.3823***	0.4007***	0.3621***	0.5489***	0.4191***
	(3.31)	(3.91)	(3.00)	(3.74)	(3.16)	(4.09)	(3.24)
ESS	5.1652**	5.0601**	5.5121**	5.1857**	5.7226**	4.5066**	5.0223**
	(2.22)	(2.18)	(2.44)	(2.30)	(2.43)	(2.32)	(2.48)
Size	-0.0012	-0.0013	-0.0012	-0.0012	-0.0014*	-0.0012	-0.0011
	(-1.44)	(-1.51)	(-1.50)	(-1.42)	(-1.79)	(-1.57)	(-1.60)
SOI	0.0317*	0.0310*	0.0307*	0.0305*	0.0305*	0.0264	0.0284
	(1.78)	(1.74)	(1.82)	(1.73)	(1.75)	(1.44)	(1.52)
R_t	-0.0074	-0.0072	-0.0075	-0.0077	-0.0077	-0.0024	-0.0018
	(-0.57)	(-0.55)	(-0.57)	(-0.60)	(-0.62)	(-0.21)	(-0.16)
R_{t-1}	-0.0111	-0.0104	-0.0104	-0.0115	-0.0110	-0.0148	-0.0145
	(-0.96)	(-0.89)	(-0.90)	(1.01)	(-0.97)	(-1.23)	(-1.20)
MOM	-0.0039	-0.0037	-0.0037	-0.0041	-0.0036	-0.0024	-0.0027
	(-0.76)	(-0.71)	(-0.72)	(-0.80)	(-0.73)	(-0.45)	(-0.51)
σ_{Stock}	-0.1776	-0.1776	-0.1245	-0.1647	-0.1571	-0.0988	-0.1203
	(-1.44)	(-1.43)	(-1.03)	(-1.41)	(-1.58)	(-1.06)	(-1.29)
ESO	-0.0115	-0.0105	-0.0098	-0.0101	-0.0135	-0.0200	-0.0179
	(-0.72)	(-0.66)	(-0.62)	(-0.64)	(-0.96)	(-1.24)	(-1.12)
OS	0.0095	0.0094	0.0096	0.0110	0.0116	0.0001	0.0023
	(1.18)	(1.17)	(1.21)	(1.42)	(1.47)	(0.01)	(0.42)
B/M	0.0008	0.0006	0.0004	0.0011	0.0008	-0.0010	-0.0010
	(0.40)	(0.30)	(0.23)	(0.55)	(0.43)	(-0.45)	(-0.43)
Turnover	-0.0001	-0.0001	-0.0003	-0.0001	-0.0001	0.0112	0.0117
	(-0.17)	(-0.19)	(-0.44)	(-0.24)	(-0.19)	(0.51)	(0.54)
PIN						-0.0005	-0.0004
						(-0.81)	(-0.73)
CW		-0.0072				-0.0666	
		(-0.18)				(-1.64)	
CVOL			0.0255			0.0203	0.0197
			(1.43)			(0.99)	(1.01)
PVOL			-0.0384*			-0.0299	-0.0279
			(-1.89)			(-1.44)	(-1.49)
Skewness				-0.0028		-0.0055	0.0052
				(-0.11)		(-0.24)	(0.20)
IVOL					-0.0088	0.0020	-0.0027
					(-0.61)	(0.11)	(0.15)
PP						0.0004	0.0001
						(0.12)	(0.04)
OOI						-0.2400	-0.3857
						(-0.16)	(-0.27)
Adj R^2	0.1191	0.1224	0.1226	0.1218	0.1274	0.1405	0.1381
N Stocks	440	438	440	439	439	449	449

Table 6. Returns for the three months following portfolio formation for IPD (Implied Price Difference from Trades) based portfolios when IPD is calculated using options with difference times to expiration. Returns and α's are in percentage terms. The Fama-French-Carhart four factor model is used to calculate α's. Newey-West three lag adjusted standard errors are used to calculate t-statistics.

Port.	R_{t+1}	α_{t+1}	T Stat.	R_{t+2}	α_{t+2}	T Stat.	R_{t+3}	α_{t+3}	T Stat.
Panel A. 10 – 29 days to expiration									
Low	0.7275	0.0058	0.04	0.6232	-0.1492	-1.33	0.7590	-0.0095	-0.08
2	0.7464	0.1082	1.14	0.6101	-0.0666	-0.61	0.6890	-0.0021	-0.02
3	0.6343	0.1052	1.78	0.5980	0.0063	0.08	0.4859	-0.1288	-2.21
4	0.3688	-0.2031	-3.04	0.5323	-0.0559	-0.76	0.6355	0.0266	0.40
High	0.5464	-0.1282	-0.76	0.6953	0.0308	0.20	0.9521	0.2585	1.65
H-L	-0.1811	-0.1340	-0.66	0.0721	0.1800	0.85	0.1931	0.2680	1.42
Panel B. 30 – 59 days to expiration									
Low	0.6346	-0.0754	-0.67	0.6641	-0.1008	-0.65	0.6521	-0.1189	-0.84
2	0.6784	0.0730	0.82	0.6833	0.0696	0.81	0.5465	-0.1320	-1.51
3	0.5194	-0.0336	-0.58	0.5879	0.0052	0.09	0.5281	-0.0913	-1.44
4	0.5352	-0.0061	-0.10	0.4906	-0.1117	-1.96	0.6729	0.0797	1.06
High	0.6347	-0.0296	-0.23	0.7875	0.1381	1.08	0.8787	0.1794	1.35
H-L	0.0001	0.0458	0.28	0.1235	0.2389	1.09	0.2265	0.2983	1.57
Panel C. 60 – 89 days to expiration									
Low	0.3570	-0.4060	-1.94	0.4212	-0.3391	-2.12	1.0601	0.2525	1.48
2	0.5141	-0.0794	-0.65	0.5018	-0.1370	-1.31	0.4287	-0.2907	-2.00
3	0.4826	-0.0805	-0.87	0.5177	-0.0709	-0.75	0.6409	0.0396	0.48
4	0.5785	0.0423	0.47	0.5943	0.0261	0.35	0.5104	-0.0758	-0.72
High	0.6064	-0.0032	-0.02	0.7984	0.1661	1.03	0.9140	0.2724	1.66
H-L	0.2494	0.4028	1.78	0.3772	0.5052	2.25	-0.1461	0.0199	0.08
Panel D. 90 – 179 days to expiration									
Low	0.2238	-0.5274	-2.88	0.4439	-0.3478	-1.86	0.5134	-0.2626	-1.36
2	0.5236	-0.1299	-1.18	0.5067	-0.1500	-1.32	0.4075	-0.3199	-2.46
3	0.5005	-0.0537	-0.83	0.5410	-0.0427	-0.61	0.5647	-0.0409	-0.68
4	0.5564	0.0313	0.53	0.5458	-0.0341	-0.58	0.6332	0.0468	0.59
High	0.7350	0.1675	1.49	0.9601	0.3651	2.98	0.8683	0.2482	2.15
H-L	0.5112	0.6949	3.11	0.5162	0.7129	3.20	0.3549	0.5108	2.38
Panel E. ≥ 180 days to expiration									
Low	0.2084	-0.4502	-2.03	-0.1268	-0.8938	-3.14	0.1766	-0.6680	-2.62
2	0.3094	-0.3176	-3.88	0.3820	-0.2734	-3.06	0.4528	-0.2427	-2.81
3	0.5990	0.0546	0.62	0.5761	0.0059	0.06	0.6541	0.0595	0.80
4	0.6015	0.0866	1.02	0.7009	0.1352	1.35	0.4954	-0.0644	-0.72
High	0.9576	0.3716	2.74	0.9952	0.3967	2.77	0.9462	0.3373	2.29
H-L	0.7492	0.8218	2.88	1.1219	1.2904	4.05	0.7696	1.0053	3.45

Table 7. IPD estimated from options with at least 90 days to expiration and future stock returns. Regressions of stock returns for month t+1 on explanatory variables from month t. Each month over 2004 – 2013, cross sectional regressions of stock returns are run on explanatory variables from the previous month. As in Fama and MacBeth (1973), coefficients are averaged across months and t-statistics are based on the three-lag adjusted Newey-West standard errors of coefficient estimates across months. When PP or OOI is included, the sample period is restricted to 2005-2013.

Panel A. All stocks.

	R_{t+1}	R_{t+1}	R_{t+1}	R_{t+1}	R_{t+1}	R_{t+1}
IPD	0.1321***	0.1284***	0.0595***	0.1305***	0.0638**	0.1092***
	(5.01)	(4.84)	(2.64)	(4.87)	(2.52)	(3.84)
ESS	-1.0700	-1.0253	-0.7021	-1.2016*	-0.3232	-0.7764
	(-1.49)	(-1.43)	(-0.97)	(-1.66)	(-0.34)	(-0.81)
Size	-0.0010	-0.0011	-0.0011	-0.0011	-0.0017**	-0.0018**
	(-1.53)	(-1.63)	(-1.63)	(-1.59)	(-2.09)	(-2.27)
SOI	0.0092	0.0071	0.0097	0.0099	-0.0027	-0.0023
	(1.22)	(0.93)	(1.30)	(1.29)	(-0.25)	(-0.22)
R_t	-0.0015	-0.0008	-0.0006	-0.0029	0.0040	0.0036
	(-0.17)	(-0.09)	(-0.07)	(-0.35)	(0.48)	(0.43)
R_{t-1}	-0.0036	-0.0036	-0.0030	-0.0043	-0.0020	-0.0018
	(-0.45)	(-0.46)	(-0.38)	(-0.54)	(-0.24)	(-0.21)
MOM	-0.0026	-0.0023	-0.0026	-0.0026	-0.0011	-0.0010
	(-0.57)	(-0.52)	(-0.58)	(-0.57)	(-0.24)	(-0.21)
σ_{Stock}	-0.0795	-0.0739	-0.0884	-0.0555	-0.0687	-0.0569
	(-1.25)	(-1.12)	(-1.35)	(-0.90)	(-1.08)	(-0.91)
ESO	-0.0046	-0.0053	-0.0040	-0.0063	-0.0123	-0.0111
	(-0.46)	(-0.52)	(-0.40)	(-0.61)	(-0.79)	(-0.72)
OS	-0.0021	-0.0015	0.0012	-0.0026	0.0005	-0.0006
	(-0.46)	(-0.31)	(0.25)	(-0.57)	(0.10)	(-0.13)
B/M	0.0019	0.0022	0.0016	0.0018	0.0010	0.0013
	(1.29)	(1.47)	(1.08)	(1.23)	(0.63)	(0.81)
Turnover	-0.0005	-0.0005*	-0.0003	-0.0005*	-0.0065	-0.0090
	(-1.53)	(-1.82)	(-1.06)	(-1.68)	(-0.47)	(-0.67)
PIN		-0.0097			-0.0005	-0.0005*
		(-0.76)			(-1.56)	(-1.73)
CW			0.0740***		0.0504***	
			(5.01)		(2.75)	
CVOL				0.0110	0.0114	0.0110
				(1.50)	(1.04)	(1.00)
PVOL				-0.0223	-0.0172*	-0.0153
				(-3.28)	(-1.68)	(-1.49)
Skewness					-0.0272	-0.0445**
					(-1.47)	(-2.54)
IVOL					-0.0074	-0.0090
					(-0.56)	(-0.69)
PP					-0.0033	-0.0030
					(-1.31)	(-1.20)
OOI					1.7740***	1.9209***
					(2.97)	(3.25)
Adj R^2	0.0632	0.0638	0.0660	0.0665	0.0871	0.0847
Mean Stocks	1,313.4	1,306.3	1,313.0	1,312.7	1,133.6	1,133.7

Table 7 (continued). Predictive regressions using maturities greater than 90 days.
Panel B. 500 Largest stocks based on market capitalization at the end of the previous month.

	R_{t+1}	R_{t+1}	R_{t+1}	R_{t+1}	R_{t+1}	R_{t+1}
IPD	0.2542***	0.2553***	0.3146***	0.2451***	0.2912***	0.2328***
	(4.10)	(4.14)	(4.46)	(3.90)	(4.52)	(4.11)
ESS	4.6804**	4.3162*	4.4217*	5.1650**	4.5407**	5.0442***
	(2.06)	(1.87)	(1.88)	(2.31)	(2.37)	(2.60)
Size	-0.0015*	-0.0015	-0.0016*	-0.0015*	-0.0013*	-0.0013*
	(-1.75)	(-1.64)	(-1.83)	(-1.81)	(-1.71)	(-1.81)
SOI	0.0277	0.0287	0.0282	0.0270	0.0262	0.0278
	(1.56)	(1.62)	(1.57)	(1.64)	(1.42)	(1.49)
R_t	-0.0060	-0.0060	-0.0053	-0.0063	-0.0012	-0.0009
	(-0.48)	(-0.47)	(-0.42)	(-0.49)	(-0.11)	(-0.08)
R_{t-1}	-0.0125	-0.0140	-0.0115	-0.0113	-0.0148	-0.0146
	(-1.06)	(-1.19)	(-0.95)	(-0.95)	(-1.21)	(-1.19)
MOM	-0.0035	-0.0036	-0.0034	-0.0033	-0.0023	-0.0025
	(-0.68)	(-0.69)	(-0.66)	(-0.64)	(-0.43)	(-0.48)
σ_{Stock}	-0.1755	-0.1572	-0.1750	-0.1319	-0.0881	-0.1065
	(-1.44)	(-1.31)	(-1.41)	(-1.09)	(-0.91)	(-1.10)
ESO	-0.0145	-0.0138	-0.0143	-0.0142	-0.0224	-0.0205
	(-0.77)	(-0.74)	(-0.75)	(-0.76)	(-1.26)	(-1.17)
OS	0.0104	0.0101	0.0105	0.0104	0.0013	0.0035
	(1.30)	(1.23)	(1.30)	(1.33)	(0.23)	(0.65)
B/M	0.0011	0.0011	0.0010	0.0007	-0.0008	-0.0008
	(0.58)	(0.56)	(0.47)	(0.34)	(-0.35)	(-0.33)
Turnover	-0.0002	-0.0002	-0.0002	-0.0003	0.0117	0.0114
	(-0.25)	(-0.27)	(-0.31)	(-0.48)	(0.51)	(0.50)
PIN		0.0123			-0.0005	-0.0005
		(0.57)			(-0.91)	(-0.84)
CW			-0.0103		-0.0642	
			(-0.23)		(-1.57)	
CVOL				0.0284	0.0210	0.0213
				(1.63)	(0.97)	(1.04)
PVOL				-0.0385*	-0.0300	-0.0286
				(-1.95)	(-1.38)	(-1.45)
Skewness					-0.0003	0.0106
					(-0.01)	(0.40)
IVOL					0.0002	0.0004
					(0.01)	(0.02)
PP					0.0002	-0.0001
					(0.05)	(-0.04)
OOI					-0.0294	-0.4868
					(-0.20)	(-0.33)
Adj R^2	0.1214	0.1237	0.1255	0.1245	0.1425	0.1400
Mean Stocks	422.6	420.5	422.6	424.9	438.8	438.8

Table 8. Coefficients on IPD for regressions with stock returns from one to ten months later.
Each month over 2004 – 2013, cross sectional regressions of stock returns are run on explanatory variables from the previous month. As in Fama and MacBeth (1973), coefficients are averaged across months and t-statistics are based on the standard errors of coefficient estimates across months. When PP and OOI are included, the sample period is restricted to 2005-2013. In all regressions, IPD is the percentage difference between the implied and actual stock price, ESS is the volume-weighted average effective spread of the stock, ESO is the volume-weighted average effective spread of the stock, Size is the natural logarithm of the market capitalization, SOI (stock order imbalance) is the stock buy volume minus the stock sell volume divided by total volume, σ_{Stock} is the standard deviation of the daily stock return over the month, PIN is the probability of informed trading, and Turn is the mean daily turnover of the stock over the month. In regressions that also includeoption-based variables, O/S is the natural logarithm of the option to stock volume ratio, OOI is the options order imbalance calculated as the relative difference between synthetic positive and negative option volume, CW is the difference between call and put implied volatilities as in Cremers and Weinbaum (2012), CVOL is the implied volatility of a 30 day call with delta = 0.5, PVOL is the implied volatility of a 30 day put with delta = -0.5, Skew is the difference between the implied volatility of a put with a delta =-0.2 and the average implied volatility of puts and calls with deltas with values |0.5|, IVOL is the average implied volatility of at the money calls and puts with 30 days to maturity and PP is the ratio of volume from buys to open put positions to buys to open both put and call positions. T statistics are based on Newey West standard errors adjusted for predictive horizon (months) plus one lags, and with 3 lags for t+1, and t+2 return regressions.

	Options with more than 90 days to expiration		Options with 11-59 days to expiration	
	All Stocks	Largest 500	All Stocks	Largest 500
R_{t+1}	0.0638**	0.2912***	-0.0250	0.1271
	(2.52)	(4.52)	(-0.41)	(0.78)
R_{t+2}	0.0631**	0.2007***	0.1000	0.0075
	(2.10)	(3.51)	(1.20)	(0.05)
R_{t+3}	0.0704**	0.1751***	-0.0680	0.0081
	(2.55)	(3.03)	(-1.05)	(0.05)
R_{t+4}	0.0369	0.1072*	-0.0440	0.2114
	(1.27)	(1.82)	(-0.72)	(1.10)
R_{t+5}	0.0743***	0.1185**	0.0578	0.0375
	(2.60)	(2.06)	(0.85)	(0.22)
R_{t+6}	0.0313	0.1669***	0.1426*	0.1448
	(0.94)	(2.89)	(1.83)	(0.85)
R_{t+7}	0.0896***	0.1068**	0.0570	0.0575
	(3.09)	(2.03)	(0.89)	(0.34)
R_{t+8}	0.0953***	0.0745	0.1479**	0.1497
	(3.14)	(0.83)	(2.00)	(0.87)
R_{t+9}	0.0599**	0.1187*	0.0417	-0.1880
	(1.91)	(1.93)	(0.60)	(-1.28)
R_{t+10}	0.0591**	0.1527**	0.0401	0.3134
	(2.45)	(2.13)	(0.63)	(1.22)
R_{t+11}	-0.0002	0.1706***	0.0277	0.0322
	(-0.06)	(2.66)	(0.45)	(0.19)
R_{t+12}	-0.0240	0.1650*	0.0307	-0.2460
	(-0.75)	(1.81)	(0.34)	(-1.43)
R_{t+13}	0.0453	0.0569	0.0799	-0.3200*
	(1.43)	(0.79)	(1.21)	(-1.75)
R_{t+14}	0.0003	0.0587	-0.0250	-0.2240
	(0.11)	(0.79)	(-0.28)	(-0.77)

Table 9. Month ahead four-factor alphas by long-term holding tercile and IPD quintiles.
Turnover is calculated for each institution each month as the average turnover of their portfolio over the previous four quarters. Long-term investors are defined as those in the tercile with the slowest turnover. Each month, stocks are sorted into terciles based on the proportion of shares held by long-term investors. Each tercile is then sorted into quintiles based on IPD, the percentage difference between implied and actual stocks prices. Alphas are reported in percentages.

Panel A. All stocks

	Low IPD	2	3	4	High IPD	High - Low	H-L t-stat
Low LT	-1.2407	-0.4705	-0.0681	-0.1224	0.0677	1.3084	5.22
2	-0.0204	0.0900	0.0698	-0.0357	0.2565	0.2770	1.36
High LT	0.1672	0.0366	0.0517	0.0649	0.2880	0.1208	0.57
High - Low	1.4079	0.5071	0.1198	0.1873	0.2203		
H-L t-stat	5.64	2.61	0.67	1.21	0.94		

Panel B. Largest 500 stocks.

	Low IPD	2	3	4	High IPD	High - Low	H-L t-stat
Low LT	-0.6444	-0.0610	0.5119	0.0326	0.5751	1.2195	3.66
2	-0.4100	0.0932	0.1784	0.2066	0.3570	0.7671	2.71
High LT	0.0233	0.0551	-0.0590	0.2946	0.3022	0.2788	1.16
High - Low	0.6678	0.1161	-0.5709	0.2620	-0.2729		
H-L t-stat	1.84	0.49	-2.41	1.57	-1.05		

Table 10.

Regressions of stock returns on interactions between IPD and dummy variables for above and below average holdings by long-term institutional investors.

Long-term investors are those with portfolio turnover in the previous four quarters that is in the bottom tercile. LT < Avg. is a dummy variable that equals one if the proportion of a stocks shares that is held by long-term investors is less than average. LT > Avg. is a dummy variable that equals one if the proportion of a stocks shares that is held by long-term investors is greater than average. IPD is the percentage difference between the implied and actual stock price, ESS is the volume-weighted average effective spread of the stock, ESO is the volume-weighted average effective spread of the stock, Size is the natural logarithm of the market capitalization, SOI (stock order imbalance) is the stock buy volume minus the stock sell volume divided by total volume, OOI is the options order imbalance calculated as the difference between synthetic positive and negative option volume, σ_{Stock} is the standard deviation of the daily stock return over the month, O/S is the natural logarithm of the option to stock volume ratio, PIN is the probability of informed trading, Turn is the mean daily turnover of the stock over the month, CW is the difference between call and put implied volatilities as in Cremers and Weinbaum (2012), CVOL is the implied volatility of a 30 day call with delta = 0.5, PVOL is the implied volatility of a 30 day put with delta = -0.5, Skew is the difference between the implied volatility of a put with a delta =-0.2 and the average implied volatility of puts and calls with deltas with values |0.5|, IVOL is the average implied volatility of at the money calls and puts with 30 days to maturity and PP is the ratio of volume from buys to open put positions to buys to open both put and call positions.

	All Stocks		500 Largest Stocks	
	R_{t+1}	R_{t+2}	R_{t+1}	R_{t+2}
LT < Avg.	0.0358**	0.0321**	0.0319**	0.0222
	(2.08)	(1.98)	(2.20)	(1.22)
LT > Avg.	0.0396**	0.0350**	0.0315**	0.0213
	(2.32)	(2.17)	(2.18)	(1.18)
LT < Avg. x IPD	0.1580***	0.0663	0.4262***	0.4251***
	(3.96)	(1.41)	(4.39)	(4.44)
LT > Avg. x IPD	0.0299	0.0761	0.2409**	0.1398
	(0.64)	(1.39)	(2.29)	(1.28)
Lt. Inst Holding	-0.0001	-0.0000	0.0001	0.0001
	(-1.44)	(-0.36)	(0.19)	(0.15)
ESS	-0.4986	-1.0138	6.3037***	4.1409*
	(-0.51)	(-1.02)	(2.93)	(1.67)
Size	-0.0016*	-0.0015*	-0.0015	-0.0010
	(-1.91)	(-1.91)	(-2.06)	(-1.13)
SOI	-0.0029	-0.0121	0.0280	0.0106
	(-0.28)	(-0.84)	(1.45)	(0.81)
R_t	0.0054	0.0032	-0.0002	-0.0186
	(0.65)	(0.33)	(-0.01)	(-1.40)
R_{t-1}	0.0001	0.0052	-0.0162	-0.0109
	(0.01)	(0.59)	(-1.30)	(-1.04)
MOM	-0.0011	-0.0023	-0.0021	-0.0022
	(-0.24)	(-0.56)	(-0.41)	(-0.45)
σ_{Stock}	-0.0931	-0.1603**	-0.0855	-0.0299
	(-1.44)	(-2.51)	(-0.87)	(-0.27)
ESO	-0.0091	-0.0035	-0.0294*	0.0034
	(-0.59)	(-0.25)	(-1.70)	(0.17)
OS	0.0019	-0.0004	0.0017	0.0043
	(0.40)	(-0.08)	(0.31)	(0.58)
B/M	0.0014	-0.0003	-0.0004	-0.0011
	(0.87)	(-0.16)	(-0.18)	(-0.49)
Turnover	-0.0040	-0.0147	0.0084	0.0098
	(-0.28)	(-1.33)	(0.37)	(0.49)
PIN	-0.0006*	-0.0001	-0.0007	-0.0006
	(-1.83)	(-0.41)	(-1.09)	(-1.21)
CW	0.0339*	0.0404**	-0.0539	-0.0080
	(1.73)	(2.36)	(-1.20)	(-0.21)
CVOL	0.0104	-0.0071	0.0291	-0.0361*
	(0.88)	(-0.75)	(1.34)	(-1.85)
PVOL	-0.0137	-0.0064	-0.0347	0.0200
	(-1.22)	(-0.68)	(-1.56)	(0.96)
Skewness	-0.0239	-0.0246	0.0019	-0.0374*
	(-1.30)	(-1.31)	(0.08)	(-1.78)
IVOL	0.0009	0.0046	-0.0028	-0.0035
	(0.06)	(0.42)	(-0.15)	(-0.22)
PP	-0.0035	-0.0015	-0.0002	0.0001
	(-1.42)	(-0.55)	(-0.06)	(0.03)
OOI	2.0259***	-1.7699**	-0.2927	-2.6712**
	(3.16)	(-2.47)	(-0.19)	(-2.14)
Adj R^2	0.2171	0.2170	0.3038	0.3014
	1,087.6	1,087.6	427.2	427.2

Table 11.
Month ahead four-factor (Fama-French and momentum) alphas for portfolios sorted first on implied volatility and then IPD. IPD is the percentage difference between the implied and actual stock price. Each month stocks are sorted into five quintiles based on IPD, and each quintile is sorted into three terciles by implied volatility. IPD and implied volatility are estimated jointly.

Panel A. All Stocks

	Low IPD	2	3	4	High IPD	High-Low
Low σ	-0.0694	0.0513	0.2608	0.0315	0.2752	0.3446
						(1.57)
Middle σ	-0.4513	-0.1621	0.0361	0.1120	0.0636	0.5149***
						(3.13)
High σ	-0.9779	-0.4875	-0.1510	-0.1285	0.0510	1.0289***
						(3.87)
High-Low	-0.9085***	-0.5387**	-0.4118**	-0.1599	-0.2241	
	(-2.89)	(-2.50)	(-2.06)	(-0.85)	(-1.10)	

Panel B. Largest 500 stocks.

	Low IPD	2	3	4	High IPD	High-Low
Low σ	-0.0801	0.1829	0.1288	0.3126	0.4138	0.4939**
						(2.11)
Middle σ	-0.4251	0.0742	0.2392	0.0364	0.3406	0.7657**
						(2.38)
High σ	-0.5247	-0.1210	-0.1377	0.5455	0.3792	0.9039***
						(2.78)
High-Low	-0.4446	-0.3040	-0.2665	0.2329	-0.0346	
	(-1.33)	(-1.63)	(-0.96)	(0.92)	(-0.13)	

Table 12. Each month over 2005 – 2013, cross sectional regressions of stock returns are run on explanatory variables from the previous month. Coefficients are averaged across months and t-statistics are based on the standard errors of coefficient estimates across months. IPD is the percentage difference between the implied and actual stock price calculated from options with at least 90 days to expiration. D_{LowVol}, D_{MedVol}, and $D_{HighVol}$ are dummy variables that are one if the stock is in the bottom, middle, or highest tercile of implied volatility on options with more than 90 days to expiration. IPD and implied volatilities are estimated jointly using option prices over thirty minute intervals. Variables that are included in the regression but not reported here are ESS, the volume-weighted average effective spread of the stock, ESO, the volume-weighted average effective spread of the stock, Size, defined as the natural log of market capitalization, SOI (stock order imbalance) is the stock buy volume minus the stock sell volume divided by total volume, σ_{Stock} is the standard deviation of the daily stock return over the month, PIN is the probability of informed trading, and Turn is the mean daily turnover of the stock over the month, O/S is the natural logarithm of the option to stock volume ratio, OOI is the options order imbalance calculated as the relative difference between synthetic positive and negative option volume, CW is the difference between call and put implied volatilities as in Cremers and Weinbaum (2012), CVOL is the implied volatility of a 30 day call with delta = 0.5, PVOL is the implied volatility of a 30 day put with delta = -0.5, Skew is the difference between the implied volatility of a put with a delta =-0.2 and the average implied volatility of puts and calls with deltas with values |0.5|, IVOL is the average implied volatility of at the money calls and puts with 30 days to maturity and PP is the ratio of volume from buys to open put positions to buys to open call positions.

	All Stocks				Largest 500 Stocks			
	R_{t+1}	t-stat	R_{t+2}	t-stat	R_{t+1}	t-stat	R_{t+2}	t-stat
D_{LowVol}	0.0230*	1.90	0.0249**	1.97	0.0130*	1.87	-0.0001	-0.01
D_{MedVol}	0.0243*	1.95	0.0259**	2.00	0.0128*	1.83	-0.0014	-0.18
$D_{HighVol}$	0.0245**	1.98	0.0237*	1.85	0.0149**	2.19	0.0007	0.09
IPD · D_{LowVol}	0.0071	0.14	0.0484	0.74	0.1881*	1.84	0.1652	1.54
IPD · D_{MedVol}	0.1629**	2.52	0.0951	1.27	0.3439**	2.12	0.4821**	2.36
IPD · $D_{HighVol}$	0.0610	1.60	0.0755	1.55	0.3859***	3.46	0.3227**	2.23
$\sigma_{Implied}$	-0.0365**	-2.11	0.0052	0.33	-0.0045	-0.12	0.0443*	1.66
ESS	-0.1198	-0.13	-0.9020	-0.98	4.4478**	2.36	4.0999*	1.75
Size	-0.0009	-1.52	-0.0010*	-1.65	-0.0006	-1.55	0.0001	0.34
SOI	-0.0034	-0.33	-0.0125	-0.84	0.0279	1.45	0.0082	0.58
R_t	0.0041	0.50	0.0017	0.18	-0.0020	-0.17	-0.0191	-1.46
R_{t-1}	-0.0013	-0.15	0.0059	0.68	-0.0158	-1.31	-0.0118	-1.12
MOM	-0.0012	-0.27	-0.0027	-0.66	-0.0027	-0.50	-0.0025	-0.53
σ_{Stock}	-0.0665	-1.05	-0.1726**	-2.54	-0.0723	-0.75	-0.0070	-0.06
ESO	-0.0045	-0.32	-0.0038	-0.30	-0.0168	-1.02	0.0114	0.62
OS	0.0001	0.02	-0.0030	-0.55	0.0012	0.21	0.0035	0.49
B/M	0.0012	0.70	-0.0005	-0.29	-0.0006	-0.27	-0.0012	-0.53
Turnover	-0.0007	-0.05	-0.0124	-1.07	0.0142	0.63	0.0166	0.84
PIN	-0.0004	-1.26	-0.0000	0.00	-0.0006	-0.98	-0.0006	-1.17
CW	0.0506***	2.84	0.0420**	2.46	-0.0666*	-1.69	-0.0270	-0.72
CVOL	0.0111	1.00	-0.0062	-0.67	0.0218	1.04	-0.0316	-1.64
PVOL	-0.0179*	-1.68	-0.0106	-1.15	-0.0297	-1.37	0.0167	0.80
Skewness	-0.0245	-1.30	-0.0228	-1.26	-0.0019	-0.08	-0.0355*	-1.67
IVOL	-0.0042	-0.31	0.0045	0.42	0.0039	0.20	0.0001	0.01
PP	-0.0022	-0.96	-0.0004	-0.16	0.0008	0.27	0.0016	-0.67
OOI	1.9240***	3.15	-1.3279*	-1.93	-0.5597	-0.38	-2.8733**	-2.26
Adj R^2	0.2189		0.2165		0.3074		0.3031	
Mean Stocks	1,133.2		1,129.0		438.7		437.6	

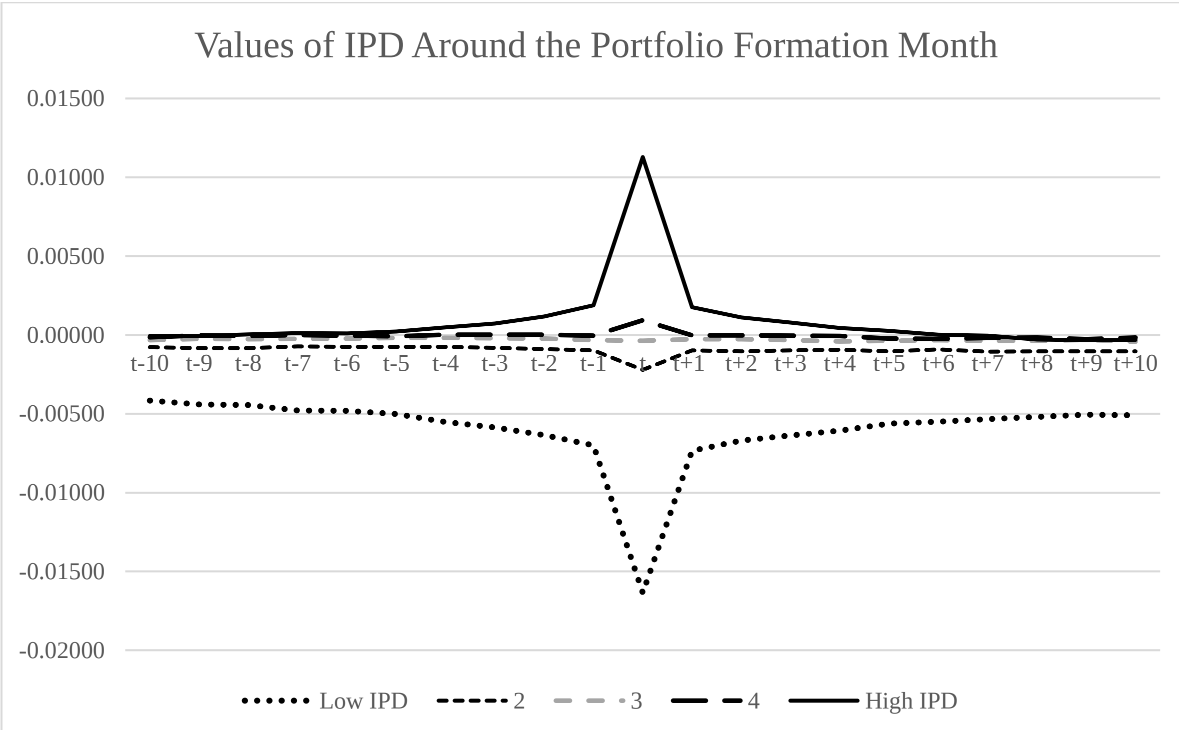


Figure 1.

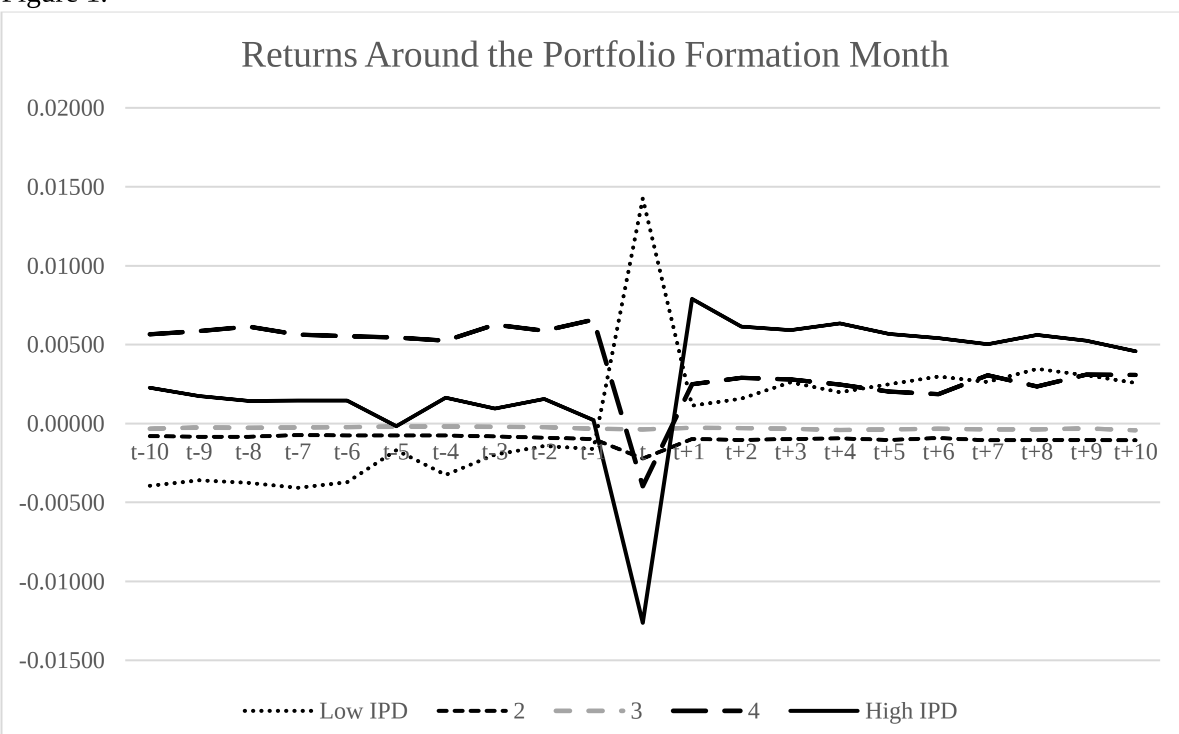
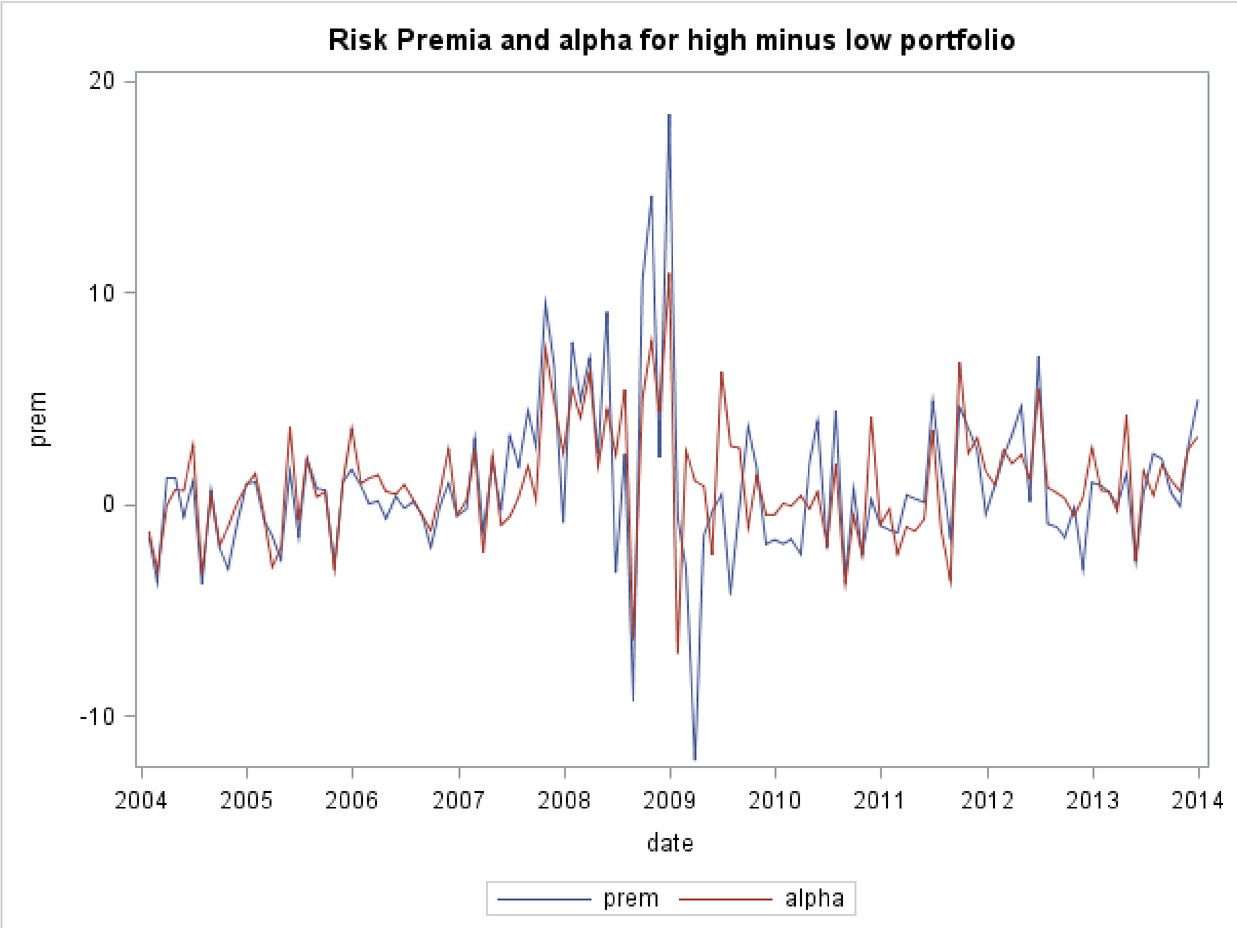


Figure 2.


Risk Premia and alpha for high minus low portfolio
20
10
0
-10
prem
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
date
prem
alpha

References

An, Byeong-Je, Andrew Ang, Turan Bali, and Nusret Cakici, 2014, The joint cross section of stocks and options, Journal of Finance 59 (5), 2279-2337.

Bollen, Nicolas and Robert Whaley, 2004, Does net buying pressure affect the shape of implied volatility functions? Journal of Finance 59 (2), 711-753.

Chakravarty, Sugato, Huseyin Gulen, and Stewart Mayhew, 2004, Informed trading in stock and option markets, Journal of Finance 59, 1235-1257.

Christofferson, Peter, Ruslan Goyenko, Kris Jacobs, and Mehdi Karoui, 2018, Review of Financial Studies 31 (3), 811-851.

Cremers, Martijn, and David Weinbaum, 2010, Deviations from put-call parity and stock return predictability, Journal of Financial and Quantitative Analysis 45(2), 335-367.

Dimson, Elroy, 1979, Risk measurement when shares are subject to infrequent trading, Journal of Financial Economics 7, 197-226.

Easley, David, Nicholas Kiefer, Maureen O'Hara, and Joseph Paperman, 1996, Liquidity, information, and infrequently traded stocks, Journal of Finance 51 (4), 1405-1436.

Easley, David, Maureen O'Hara, and P.S. Srinivas, 1998, Option volume and stock prices: Evidence on where informed traders trade, Journal of Finance 53, 431-465.

Fama, Eugene, and James MacBeth, 1973, Risk, return and equilibrium: Empirical tests, Journal of Political Economy 81 (3), 607-636.

Gaspar, Jose-Miguel, Massimo Massa, and Pedro Matos (2005), Shareholder investment horizons and the market for corporate control. Journal of Financial Economics 76, 135-165.

Ge, Li, Tse-Chun Lin, and Neil Pearson, 2016, Why does the option to stock volume ratio predict stock returns? Journal of Financial Economics 120, 601-622.

Hu, Jianfeng, 2014, Does option trading convey stock price information? Journal of Financial Economics 111, 625-645.

Johnson, Travis and Eric So, 2012, The option to stock volume ratio and future returns, Journal of Financial Economics 106, 262-286.

Lee, Charles, and Mark Ready, 1991, Inferring trade direction from intraday data, Journal of Finance 46, 733-746.

Manaster, Steven and Richard Rendleman, 1982, Option prices as predictors of equilibrium stock prices, Journal of Finance 37 (4), 1043-1057.

Muravyev, Dmitriy, Neil Pearson, and John Broussard, 2013, Is there price discovery in equity options?, Journal of Financial Economics 107, 259-283.

Ni, Sophie, Jun Pan, and Allen Poteshman, 2008, Volatility information trading in the option market, Journal of Finance 63, 1059-1091.

Pan, Jun and Allen Poteshman, 2006, The information in option volume for future stock prices, Review of Financial Studies 19 (3), 871-908.

Roll,Richard, Eduardo Schwartz, and Avanidhar Subrahmanyam, 2010, O/S: The relative trading activity in options and stock, Journal of Financial Economics 96, 1-17.

Xing, Yuhang, Xiaoyan Zhang, and Rui Zhao, 2010, What does the individual option volatility smirk tell us about future equity returns? Journal of Financial and Quantitative Analysis 45 (3), 641-662.