The Intertwined Price Discovery Processes in Equity and Option Markets

Yunjiang Dong
Queen's University

July 30, 2025

Abstract

I analyze the linkage between the price discovery processes of the underlying equity and the option markets after option transactions. The price impacts in the two markets are heterogeneous and are connected through the implied volatility. This suggests that option trades contain two dimensions of information: price and volatility. Option trades associated with single-leg execution contain price information but not volatility information, while trades through the limit order book contain both price and volatility information. Furthermore, change in implied volatility is negatively related to underlying price impact, suggesting that part of the option trade information is hidden from the underlying market prices.

Keywords: Option Price Discovery, Market Information Share, Implied Volatility Impact, Volatility Information

JEL Codes: G11, G12, G14

1 Introduction

Compared to equities, options contain several unique characteristics that may cater to some investors' specific needs. First, options have embedded leverage, which allows investors to gain market exposure with minimal capital commitment. Second, options present investors with a cheap and easy-to-manage method due to its unique payout structure. In the absence of options, such a payout structure can only be replicated by a dynamic portfolio of equities and risk-free assets, which is costly in both transaction and management costs. The relationship between equity and option prices has been well-studied in financial literature. In both binomial and Black-Scholes option pricing models, option and equity prices are connected through the volatility of the underlying equity price. It is, however, unclear how such linkage manifests in the real world in a market microstructural sense.

In this paper, I study the linkage between the options and equities markets by comparing the price reactions of these markets after informed option trades. I first identify informed option trades and compare their role in the equity market price discovery process with stock trades by using the vectorautoregression (VAR) model. Overall, the variance decomposition results show that option trades contribute much less to the underlying price discovery than stock trades. Option trades contribute to 0.4% of the total return variance, while stock trades contribute to 44%. However, the impulse response function of the same model shows a different story. There is significant heterogeneity in the underlying price reaction upon option trades. Option trades priced through the limit order book lead to impulse responses that are a quarter of that caused by a trade in the underlying, whereas option trades priced through auctions lead to almost zero impulse responses. The relative impulse response is significant since part of the option trade information is dampened as the information flows across the markets through the trading activities of market makers and arbitrageurs. In addition, part of the option trade information is absorbed into the change in implied volatility, which allows the underlying price to react less to option trades. In the empirical analysis section of this paper, I show that the change in implied volatility and the underlying price impact have a negative relationship.

I then focus my analysis on the various effects of informed option trades. Informed option trades should, in theory, affect both the option price and the underlying price. The former reacts to option trading activity simply by market forces, where buyers tend to increase the price and sellers tend to decrease the price. The latter, however, reacts to option trading activity in a more subtle way. By construction, the payout of option contracts is directly linked to its underlying price. However, due to the non-linear relationship between option and the underlying prices, the co-movement of these prices depends on another variable, the implied volatility. This parameter is usually assumed to be constant, at least in the short term. For instance, when computing the synthetic underlying price based on option prices, Chakravarty, Gulen and Mayhew (2004) used the 30-minute lagged implied volatility, Muravyev, Pearson and Broussard (2013) used the rolling average implied volatility of the past 30 minutes. When studying the change in implied volatility upon option trades, I found that although the median change is close to 0, the distribution is fairly wide, with large magnitudes, especially when the option trade is informative. Both the average and median implied volatility changes are also non-zero when in subsample analysis, separating the option trading observation into buy calls, buy puts, sell calls, and sell puts. Buying/selling options, whether a call or put, tend to increase/decrease the implied volatility of the option contract.

Through this study, I look into the effect of option trading on both option and equity markets simultaneously. I examine the impacts of option trading on 1) the underlying price, 2) the option price, and 3) the option implied volatility. To author's knowledge, this paper is the first to measure and interpret the impact of option trades on the implied volatility impact. In addition, I find that the change of implied volatility upon option trades plays a crucial role in linking the impacts of the two markets. I use several different approaches (VAR, linear regression, panel regression, and instrumental variables) to analyze this intertwined relationship. Most of the time, different approaches have consistent results and interpretation: Option trades that are based on single-leg strategy, not involved in auction price improvements, and close to the money, are seemingly more informative and lead to higher price impacts in both the underlying and the option markets in all empirical approaches. However, when I include the change of implied volatility as a variable, whether on the left-hand side or the right-hand side, I observe results that are more intriguing. Despite having significant positive price impacts on both the option and underlying market, single-leg option trades have no impact on implied volatility. On the other hand, auction option trades retained their significant negative impact on implied volatility. This suggests that option trades can be informative in two ways: price and volatility. Single-leg option trades are informative in price but not volatility, while trades through the limit order book are informative in both price and volatility. The instrumental variable regressions reveal that underlying price impacts are negatively related to the change of implied volatility. This suggests that part of the information included in informed option trades acts upon the implied volatility and does not transmit into the underlying market in terms of price information.

The remaining part of the paper is structured as follows: Section 2 reviews the related literature and highlights this paper's contribution on top of the existing literature. Section 3 details the characteristics of the OPRA option trade data, and explains variable and sample construction processes. Section 4 presents the empirical methods and analyses of this paper. Section 5 concludes.

2 Literature Review

This paper contributes to the literature on the role and impact of option trading on the price discovery of the underlying equity. Previous work such as Chakravarty et al. (2004) and Holowczak, Simaan and Wu (2006) has identified the relative contribution to price discovery by option trades, but with the assumption that implied volatility stays constant after option trades. This paper also contributes to the discussion of the lead-lag relationship between option and equity markets (Muravyev et al. (2013), Hu (2014)). In this paper, I contribute to the discussion by being the first to measure the information share of option trades through the usage of high-frequency option trading data, which does not rely on the assumption of constant implied volatility. I also expand the sample size to include all option trades associated with S&P 500 firms as the underlying. In addition, I calculate and dissect the implied volatility impacts of option trades, showing that option trades contain volatility level information, which is omitted in prior studies, as price information has been the main focus in this part of the literature.[1]

This paper also contributes to the literature on the study of option-induced volatility change. Recent papers by Ni, Pearson, Poteshman and White (2021) and Lipson, Tomio and Zhang (2023) have documented the involvement of retail option traders and their impacts on the underlying volatility. Whereas from a theoretical standpoint, Back (1993) explained that the addition of option market would cause the underlying stochastic volatility to become stochastic. I contribute to this literature through the empirical analysis on the change of implied volatility upon option trades.

This paper is also related to the literature on the economic interpretation of implied volatility. Christensen and Prabhala (1998), Busch, Christensen and Nielsen (2011) have shown that implied volatility can be used to predict realized volatility, which allows us to connect the change of implied volatility upon option trading to identify option trades that contain volatility information.

Finally, this paper contributes to the recent option studies that are related to retail option trading, payment for order flow, and option auctions as the mechanism of price improvements(Lipson et al. (2023), Bryzgalova, Pavlova and Sikorskaya (2023), Hendershott, Khan and Riordan (2022)). I supplement their findings by showing that option trades that are priced through auctions contain significantly less information than order book trades,

[1] Volatility information of option trading has been studied in Ni, Pan and Poteshman (2008), which shows that the total option orderflow contains future volatility information. This paper supplements their findings by studying the volatility information content at the individual trade level, as well as connecting it to the equity market price discovery process.

both in terms of price and volatility information.

Trading activities on the exchange markets convey information. Microstructure theories such as Glosten and Milgrom (1985) and Kyle (1985) suggest that informed investors trade to profit off their private information. Consequently, such informed trades create order imbalances, resulting in market makers adjusting the bid-ask quotes in response to informed trading and adverse selection. This leads to price impact upon informed trading. Through changes in the prices and quotes, private information is converted to public information. The magnitude of the price impact is correlated with the amount of private information conveyed, the market depth, and the informativeness of the market. Therefore, these parameters are crucial for empirical research. The empirical foundation of market informativeness research originates from the vector autoregression (VAR) model introduced by Hasbrouck (1991a, 1991b, 1995). Two estimates from the VAR model that are proven to be exceptionally useful are impulse response and variance decomposition. The former measures the price change induced by trading activities, which correlates to price discovery and information intensity. The latter measures the information contribution of explanatory variables, in which the list of explanatory variables depends on the setting of the study. For instance, the explanatory variables could be dummy variables for exchange, type of investor, type of orders, type of market, etc. The vector autoregression (VAR) model was employed by Hasbrouck (1991a, 1991b) to show that larger trades and trades associated with smaller firms have higher price impact based on impulse response measures, and hence contain more information. The vector error correction model (VECM), an extension of the VAR model, was employed by Hasbrouck (1995) to show that based on the variance decomposition measure, the New York Stock Exchange (NYSE) contributed to more than 90 percent of the total price discovery process at the time of the study. Apart from studying the relative information contribution of different exchanges, the VAR and VECM frameworks are proven to be effective empirical models for similar studies under various settings. Using VECM, Chakravarty et al. (2004) found that the option market contributes 17% in the price discovery process. Hendershott, Jones and Menkveld (2011) shows that the average impulse response on prices decreases over time as algorithm trading activities increase, particularly for large firms. Brogaard, Hendershott and Riordan (2019) studied the information content of limit orders, and showed that limit orders, particularly those from high-frequency traders (HFTs), contain a significant amount of information comparable to market orders and marketable limit orders. In this paper, I heavily utilize the VAR framework to measure the information contribution of different types of option trades.

Theories associated with strategic informed trading suggest that informed investors split their orders to hide their information within uninformed order flow. Kyle (1985) shows that in a single market scenario, informed orders are split across time, corresponding to the magnitude of the uninformed order flow. In a multi-market scenario, informed orders would split trading venues. Easley, O'hara and Srinivas (1998) show that if informed investors have access to both stock and option markets, there exists a "pooled equilibrium" where informed investors submit orders in both markets to maximize their profit, when embedded leverage and option liquidity are sufficiently high. Therefore, it is reasonable to believe that informed option trades exist and convey information. The question, however, is whether the option market contains incremental information that is absent in the equity market.

Regarding the information contribution of stock versus option markets, there is an ongoing debate about the lead-lag relationship between the two markets. Muravyev et al. (2013) show that during disagreement events between the stock price and its option-implied stock price, stock prices are persistent while option quotes re-adjust to eliminate the disagreement. This suggests that stock prices dominate option prices informationally. On the other hand, Hu (2014) shows that option-induced order imbalance in the underlying market predicts future stock returns, while order imbalances unrelated to option trading do not. This suggests that the direction of information flow is from option market to stock market.

The option market provides valuable data for volatility related studies. First, implied volatility can be computed based on option and underlying prices using both the binomial tree model (Cox, Ross and Rubinstein, 1979) and the Black-Scholes option pricing model (Black and Scholes, 1973). Second, implied volatility contains information about future realized volatility (Christensen and Prabhala (1998), Busch et al. (2011)). As a result, implied volatility is a suitable candidate for studying the stochastic behavior of volatility in a high-frequency setting. In this study, I find that implied volatility changes significantly upon an average option trade, especially for out-of-the-money options.

The recent rise of retail investors in the option and derivative markets has led to a significant change in the market landscape. Using OPRA data, Bryzgalova et al. (2023) document the trading behavior and tendencies of retail option investors, as well as the option wholesalers who offer commission-free services to investors and internalize retail investors' orders. Hendershott et al. (2022) demonstrate the pricing mechanism between option market makers and payment for order flow (PFOF) wholesalers. In addition, retail option trading can also impact the market quality of the underlying market. Lipson et al. (2023) document the increase in underlying volatility when the retail option trading activity is high. Ni et al. (2021) has provided an explanation for this phenomenon. The change in underlying return volatility and price movements are caused by option market maker rebalancing their portfolio based on their option positions. In this paper, I analyze the heterogeneous impacts of different types of option trades on equity prices, option prices and option implied volatility. I find that option trades through auctions, which are likely classified as retail trading activity in previous studies, have minuscule impact on all three measures.

3 Data and Sample Construction

3.1 Option Price Reporting Authority (OPRA)

I obtain transactional level option data from the Option Price Reporting Authority (OPRA)[2]. OPRA is a centralized system in the U.S. that consolidates and disseminates real-time price

[2] A detailed descriptive overview of the OPRA data can be found in Andersen et al. (2021)

and trade data for options contracts traded on various options exchanges. OPRA serves investors, traders, and financial institutions by providing transparency and efficiency in the options market. It collects and distributes data on options prices, trades, and related information from multiple exchanges, helping market participants make informed decisions and maintain market integrity. OPRA collects all option transactions from its 16 participant exchanges[3], which are the National Market System (NMS) exchanges that are approved by the Securities and Exchange Commission (SEC) to trade listed derivatives. Since all listed options must be traded on one of the approved exchanges, the OPRA data contains all the U.S. listed option transactions[4]. In 2020, the Option Clearing Corporation (OCC) reported an average daily volume of 30 million contracts for listed options[5]. For every option transaction, the data include underlying symbol, date and timestamp to the millisecond, sequence number, expiration date, strike price, option type (call or put), trade size, trade price, trade condition ID, national best bid and offer (NBBO) of the option, and NBBO of the underlying, at the time of the transaction. The trade condition ID provides several important pieces of information about the option transactions[6]. Based on the trade condition ID, I identify the pricing mechanism (through auction or limit order book), and investors' trading strategy (single-leg or multi-leg).

3.2 Variable Construction

For each option trade observation, I compute the implied volatility and delta using a binomial option pricing model Cox et al. (1979) to account for dividends and the nature of American options. I then measure the price impact of option trade on its underlying price with the

[3]The 16 exchanges are: NYSE (AMEX and ARCA), Boston Options Exchange (BOX), Chicago Board Options Exchange (CBOE, C2, and BATS), Miami International Holdings (EMERALD and MIAX), Nasdaq Inc. (GEMX, ISE, MRX, NASD, BX, and PHLX).

[4]Specifically, "standardized listed options" are only traded on national security exchanges. Moreover, the Options Clearing Corporation (OCC) only accepts clearing for standardized listed options that are traded on the exchanges. See staff report equity options market structure conditions by SEC.

[5]See historical volume statistics by OCC.

[6]See Appendix Table A1 for a complete description of trade condition IDs assigned by ORPA.

following formula:

$$PI_u = \frac{mid_{u,t+5} - mid_{u,t}}{mid_{u,t}} \times B \tag{1}$$

where PI_u is the underlying price impact, $mid_{u,t}$ is the underlying midprice at the time of the option trade, and $mid_{u,t+5}$ is the underlying midprice 5 minutes after the trade. The variable B indicates the trading sentiment direction, where +1 signifies bullish trades and -1 signifies bearish trades. To determine the trading sentiment direction, I first use the Lee and Ready (1991) tick test to assign trade directions. I then classify trades in the following manner: A market buy of call option or a market sell of put option is classified as a bullish trade, while a market sell of call option or a market buy of put option is considered bearish.

Apart from impacting underlying prices, option trades also have impacts on their own prices. To calculate the option on option price impact, I do:

$$PI_o = \frac{mid_{o,t+5} - mid_{o,t}}{mid_{o,t}} \times D \tag{2}$$

where PI_o is the option price impact, D is the direction of the trade (+1 for market buy, and -1 for market sell). This formula is similar to that of the underlying price impact, with the subscript o indicating the price impact and midprice of the option.

The measure of implied volatility impact can be constructed similarly:

$$IVI = \frac{IV_{t+5} - IV_t}{IV_t} \times D \tag{3}$$

where IVI is the implied volatility impact.

Both $mid_{o,t+5}$ and IV_{t+5} are collected based on subsequent option trades of the same contract. However, due to variability in strike prices and expiration dates, not all contracts receive a price update through trading within 5 minutes. In my sample, only 50% of option trades are followed by a trade of the same contract within a 4 to 6 minute window. For these trades, I take the observed future option midprice as $mid_{o,t+5}$ to compute the observed price impact through equation 2.

3.3 Sample Construction

I collect option transaction level data from OPRA, including option contract information (Underlying ticker, strike price, expiration date, option type as in call or put), trade price, bid-ask quotes, trade size, and the trade conditional ID. The data from OPRA covers the period from January 2nd, 2020 through October 21st, 2020. I focus on equity options with S&P 500 stocks as the underlying. I remove canceled trades, floor trades[7], trades and quotes with nonpositive size or price, and quotes with a negative spread. Trades with the same option contract, trade condition ID, and timestamp are aggregated into one consolidated trade. I also remove trades reported during the first 15 minutes and the last 5 minutes of the trading hour.

I collect relevant information from several other datasets and incorporate it into the sample. I obtain daily closing prices, daily returns, market capitalization, dividend amount, and ex-dividend dates from the Center for Research in Security Prices (CRSP). I obtain intraday trading cost measures (quoted spread, effective spread, realized spread, and price impacts) of equities from the Trade and Quotation (TAQ) database.

For each option trade observation, I compute the delta and implied volatility of each option transaction using a binomial model, in order to account for dividends and the nature of American options. I also compute the 1-minute option and underlying price impact after an option trade. I assign trade directions using the Lee and Ready (1991) tick test. A market buy of call option or a market sell of put option is classified as a bullish trade, while a market sell of call option or a market buy of put option is considered a bearish trade. The moneyness of the option is defined as the log difference of underlying mid price and the strike price for call options, and multiplied by -1 for put options. An option trade is classified as in-the-money (ITM) if its moneyness is greater than 0.03, out-of-the-money (OTM) if the

[7] Prior to removal, floor trades only constitute to 0.12% of all transactions.

moneyness is less than -0.03, and at-the-money (ATM) if in between. The trade condition ID allows for classification of option trades based on their pricing mechanism and execution type. An option trade is priced either through auction (with price improvements from NBBO), or through the limit order book (traded at the bid or ask quote). An option trade could either be involved in a single-leg or multi-leg trading strategy.

3.4 Option Trade Characteristics

Table 1 reports the option trade characteristics. Though the list of options available on the market is symmetric (i.e. For every call option at a specific strike price, there is a put option available at the same strike price), call options (64% of total number of trades) are traded more often than put options (36%). The frequency of bullish and bearish trades are both very close to 50% of the trades.

Option trades of ATM (42%) and OTM (50%) options happen more frequently than that of ITM (8%) options. As the moneyness of option increases, the delta of option contract approaches 1. Such ITM options will have the same payout as underlying equity if the underlying price stays within the money, and less downside risk if the underlying price moves out of the money. In order for an investor to consider ITM option over equity investment, the extra cost of trading option over equity must be justified by the downside risk of equity and the cost of borrowing when trading equity. The low percentage of ITM option transactions indicates that this is likely not the case.

In terms of the pricing mechanism, 23% of the trades are priced by auctions while the remaining 77% are priced by the limit order book. Auction trades are largely made up with orderflow from wholesalers which originates from retail investors Hendershott et al. (2022).

In terms of the trading strategy, 77% of the trades are single-leg trades and 23% of the trades are multi-leg trades. Single-leg options are simpler and less costly to trade, while multi-leg options provide flexibility to investors for their specific hedging purposes. Therefore, single-leg options are more likely used by speculators and informed traders, while multi-leg options are more likely used by hedgers and volatility traders.

Insert Table 1 Here

3.5 Option Trade Summary

Table 2 reports the descriptive statistics of option trades in the sample at the transactional level. The option trade price is heavily skewed right with a small number of trades at very high prices, with a mean of \$9.37 and median of \$2.14. This is consistent with the mean and median of days to expiration (38.72 and 8.46 days respectively), as trading of options happens more often when the expiration date is close, while options with long expiration tend to increase in value due to the time value of options. The average trade size is 5.51, while the median trade size of 1 indicates that the option market is heavily favored by individual and retail investors who often tend to trade at low volume (Bryzgalova et al., 2023). The average moneyness is -0.07, and the median is -0.03, indicating that most option trades revolve around options that are OTM or ATM. This is because ITM options have low embedded leverage, investors are better off trading ATM or OTM options which have higher embedded leverage, or trading the underlying stock directly which have higher liquidity (Easley et al., 1998). The average quoted spread average is 8.81 bps and the average effective spread average is 6.69 bps. The difference mostly arises from price improvement from NBBO offered by market makers to wholesalers through auctions (Hendershott et al., 2022). The price impact of the option (with a mean of 0.59 bps) is much lower than the effective spread. This indicates that option trades are unlikely to contain information which generates a permanent price impact.

Insert Table 2 Here

4 Empirical Analyses

4.1 Price Discovery in the Underlying Market upon Option Trade

To study the effect of option trading on the price discovery process of the underlying market, I use the vectorautoregressive (VAR) model. Originating from Hasbrouck (1991a), the VAR model was used to measure the effects of equity trading on the price discovery process in the stock market. Brogaard et al. (2019) extended the empirical methodology to study the effect of orders and messages, including limit orders that may or may not be executed in the future. In this paper, I use the VAR model to estimate the effect of not only stock trades, but also option trading activities, on the price discovery process of the underlying stock. The VAR model incorporates a set of equations where each variable is written as a linear combination of all other variables and their lags, and its own lags. I estimate the following model for every stock-day:

$$X_t^i = \sum_{k=1}^{5} \beta_k^i X_{t-k}^i + \sum_{j \neq i, j \in N} \sum_{k=0}^{5} \beta_k^j X_{t-k}^j \quad \forall i \in N \tag{4}$$

where N denotes the set of variables involved in the VAR model, including event-time underlying returns, stock trading characteristics, and option trade characteristics. Superscript i and j indicate ith and jth variable within set N. Subscript k indicates the number of lags (in event time, with $k = t$ indicates that the variable is taken at the time of the trading event). For instance, $X_t^i - 2$ denotes the value of the ith variable two trades prior to the current trade. β_k^j is the coefficient to be estimated for the kth lag of jth variable. Note that for $j = i$, I do not include the contemporaneous value of X_t^i on the right hand side, as this would trivialize the coefficient estimates. For $j \neq i$, I allow for contemporaneous effects from other variables.

The VAR is estimated in event time, which advances by 1 whenever an option or stock trade takes place. The X variables included are: *Event Time Underlying Return*, *Stock Trade - Price Change*, *Underlying Trade - Same Price*, *Option Trade - ITM Auction*, *Option*

Trade -ITM order book, *Option Trade - OTM Auction*, *Option Trade - OTM order book*, *Option Trade - ATM Auction*, *Option Trade - ATM order book*. The variables are defined as follows: *Underlying Trade - Price Change* captures all the stock trades with a trade size greater than the NBBO depth and hence moves the NBBO. *Underlying Trade - Same Price* captures all the stock trades with a smaller trade size than the NBBO depth and hence do not move the NBBO. *Option Trade - ITM Auction, ITM order book, OTM Auction, OTM order book, ATM Auction, ATM order book* each captures all the options trades within its specific category based on moneyness (ITM/OTM/ATM) and pricing mechanism (auction/order book).

Apart from *Event Time Underlying Return*, all variables indicate the characteristics of the trading event, are mutually exclusive, and sum to 100% of all the observations used in the VAR model estimation. Table 3 reports the frequencies and variance contributions of each trading characteristic on the underlying return. Variance contribution is computed as in Hasbrouck (1995), which indicates the relative contribution to the price changes of each trade type. The variance contribution of option trades sum to 0.4%, while the trading frequency of options sum to 8%. This shows that options contribute less to the development of the underlying prices. This agrees with Muravyev et al. (2013), which shows that option trades rarely lead to permanent price changes in the underlying. However, the relative differences across different types of option trades suggest that informed investors may have preferences for specific types of option trades.

Insert Table 3 Here

With 9 variables and 5 lags included in the VAR model, there are $((6 \times 8) + 5) \times 9 = 477$[8] coefficients estimated from each stock-day. The interpretation from coefficient estimates is hence not intuitive. An easier method to interpret the model results is to construct the impulse response function (IRF). The impulse response function is constructed by giving

[8]For each equation, the other 8 variables each contribute 5 lagged coefficients and 1 contemporaneous coefficient, while the variable from LHS contributes 5 lagged coefficients. Therefore, a VAR system with 9 variables and 5 lags leads to 477 coefficient estimates.

the fitted model an impulse of 1 for the "explanatory variable", and recording the change in value (response) for the "dependent variable". The VAR model does not constitute a dependent variable in the traditional sense, as the set of equations puts each variable on the left-hand side once and regresses on all other variables, their lags, and the variable's own lags. In this study, the variable of interest is the *Event Time Underlying return*, which reacts to trading events.

Table 4 reports the average stock-day impulse response function of underlying return upon stock or option trades up to 20 events forward (Figure 1 provides a visualized version). At $t = 20$, *Stock Trade - Price Change* and *Stock Trade - Same Price* generate a return response of 0.98 bps and 0.91 bps, respectively. On the other hand, option trades generate lower return responses. Trades priced through the electronic limit order generate a return response of 0.23 (ITM), 0.16 (OTM), and 0.24 (ATM) bps. Trades priced through auction generate a return response of 0.01 (ITM and OTM), 0.02 (ATM) bps. This is consistent with Hendershott et al. (2022) that market makers expect auction trades to be from uninformed order flow, and choose to not immediately rebalance their underlying equity holdings after a change in their inventory. OTM trades through the limit order book generate a lower return response than ITM and ATM through the limit order book. This is because ITM and ATM options have higher delta than OTM options, requiring market makers to make larger trades in the underlying in order to delta hedge the change in their option positions. The return responses of ITM order book trades and ATM order book trades are similar despite ITM options having a larger delta. This is because most ITM trades are still near-the-money (Table 2 shows that the 95th percentile of moneyness is at 0.06). On the other hand, informed option trading tends to happen at-the-money, where implied volatility is low, liquidity is high, leading to lower trading costs for informed investors.

Insert Table 4 Here

The VAR setting that includes both stock and option trading observations allows us to compare the relative informativeness of stock and option trading. To specifically further understand the role of option trading on the underlying price discovery, I use a linear regression approach where the dependent variable is the 1-minute underlying price impact, and the explanatory variables are option trading characteristics. Figure 2 shows that the price impact on the underlying peaks out at around 1 minute after an average option trade. This approach has several benefits. First, this approach is in real-time instead of event time, and its results can supplement results and discussions originated from table 3 and 4, and showcase the robustness of results regarding the timing method. Second, this approach allows us to incorporate control variables, such as market volatility, trade size, etc. Finally, the VAR setup forces us to categorize trades in a way that characteristics are mutually exclusive and sum up to 100% of the sample. With linear regression, I can use trade characteristic dummies without the above restrictions, while also allowing for interaction terms among trading characteristics[9]. The regression model is:

$$PriceImpact = \alpha + \beta X + \gamma Controls + \epsilon \tag{5}$$

where X includes option trading characteristics and their interaction terms. Table 5 reports the average stock-day estimates of the linear regression model. Model (1) includes only option trading characteristics (ITM, OTM, Single-leg, Auction). Model (2) includes option trading characteristics and their interaction terms. Model (3) includes option trading characteristics and control variables (*Trade Size, Underlying Volatility, SPY Volatility, Realized Spread, Lagged SPY Return, Limit Order Book Imbalance, and Implied Volatility*). Model (4) includes option trading characteristics, interaction terms, and control variables.

Insert Table 5 Here

From table 5, the coefficients for ITM and OTM are significantly negative throughout most model specifications. This indicates that ATM option trades have the highest underly-

[9]If I include single-leg vs multi-leg in the VAR setup, it would increase the number of variables to 25, as I must assign every possible combination of option trade characteristics a variable. This also increases the computation time of the VAR model tremendously. In the linear regression model, I am able to include trading strategy characteristics as an explanatory variable.

ing price impact. This suggests that informed investors prefer to trade ATM options, which typically have lower implied volatility and higher liquidity. This is consistent with Easley et al. (1998), which states that informed investors would only choose to invest in options when the benefits of the embedded leverage outweigh the trading cost.

The coefficients for single-leg option trades are significantly positive for all model specifications. This suggests that single-leg trades are likely to be more informative on the underlying prices than multi-leg trades. As multi-leg trades are mostly used for hedging purposes, such trades are less likely to be informed. However, the interaction term *OTM * Single-leg* is significantly negative, while *ITM * Single-leg* is statistically insignificant. This suggests that OTM single-leg option trades are likely to be from speculators, instead of informed investors, who minimize their trading costs by trading ATM or slightly ITM options.

The coefficients for auction trades are significantly negative for all model specifications. This suggests that option trades that are priced through auctions and receive price improvements are unlikely to be informed. This result corroborates the IRF results reported in table 4. Auction trades largely come from PFOF wholesalers who internalized orders from retail investors (Bryzgalova et al., 2023), and are deemed uninformed by market makers, who provide price improvements as incentives to facilitate the trade (Hendershott et al., 2022). However, both *ITM * Auction* and *OTM * Auction* have statistically significant positive coefficients. A possible explanation for this result is that ATM auction trades contain less sentiment information from retail traders, while ITM and OTM auction trades may come from retail herding (Hsieh, Chan and Wang, 2020), and puts more price pressure on the underlying.

The coefficients for control variables are mostly expected. Underlying price impacts increase with option trade size and limit order book imbalance. Higher option trade size leads to higher hedging trades in the underlying market by option market makers. Higher limit order book imbalance is associated with higher probability and magnitude of price movements that can be stimulated by option trades. Both underlying volatility and implied volatility are negatively correlated with underlying price impacts. This is consistent with Easley et al. (1998) that informed investors move away from the option market when trading cost increase. By design of the option contract, option prices and trading costs increase with underlying volatility and implied volatility.

When an option trade takes place, both the option price and the underlying price are impacted. Therefore, it is interesting to compare the underlying price impact and the option price impact after option trades. The empirical challenge for measuring option price impact in real time is that option prices are not comparable across option contracts with different strike prices, expiration dates, and option types. To tackle this challenge, I group option trades by their option contracts and compute the event time option price impact after each option trade. In addition, I also compute the underlying price impact under the same event time scheme.

The VAR model and the linear regression produce consistent results about the price impact of option trades on the underlying stock on the stock-day level. To generalize the result to the full sample, I perform a panel regression on the event time underlying price impacts. Table 6 reports the regression results. Model (1) is the pooled regression with no fixed effects. Model (2) reports the estimates with firm fixed effects. Model (3) reports the estimates with date fixed effects. Model (4) reports the estimates with both firm and date fixed effects. Consistent with table 4 and 5, the coefficients for single-leg option trades are positive and statistically significant under model (2) and (4). The coefficients for auction trades are negative and statistically significant in all 4 model specifications. Surprisingly, ITM and OTM option trades have positive statistically significant price impact on the underlying in the event time panel setting, but not in VAR and stock-day models. One possible explanation is that the event time underlying price impact captures the immediate price reaction after option trades, which may be different from permanent price impacts captured by the VAR model and the 1-minute price impacts captured by the stock-day regression model.

Insert Table 6 Here

With the same panel model, I regress option price impacts on option trade characteristics and fixed effects and reported the results in table 7, with the same respective model specifications. The signs for option trade characteristics are consistent with table 6, but with a larger magnitude. This suggests option trades have immediate price impacts on options, but not on the underlying stock.

Insert Table 7 Here

4.2 Change in Implied Volatility after Option Trade

An option trade has impacts on prices of both the option itself and the underlying. Such impacts are assumed to be connected via a constant implied volatility assumption. This assumption ignores the effect of an option trade on the option itself versus the underlying. Price impact in the option market is largely contributed by the low liquidity and small number of attendants in the market, such price impacts are more likely to be transient, and less likely to carry private information. On the other hand, price impacts exerted from the option market to the underlying are more likely to be permanent and carry private information. I measure the implied volatility impact of each option trade by comparing the implied volatilities of two consecutive option trades in the same option contract (i.e., option contract with the same underlying, option type, strike price, and expiration) in event time. An option trade with positive implied volatility impact increases the implied volatility, which elevates option prices relative to the underlying. Conversely, negative implied volatility impact means underlying prices are more elevated relative to the option price. An implied volatility impact of 0 indicates that option prices and underlying prices impacts are synchronized. Table 8 shows the theoretical relationship between option trade, underlying price impact, and implied volatility impact.

Insert Table 8 Here

Table 9 reports the descriptive statistics of the implied volatility impacts of different types of option trades. The implied volatility impact has a positive mean and median, albeit very close to 0. A positive average implied volatility indicates that option trades, on average, do not contain new information. This is because option trades based on private information should affect underlying prices permanently. When the impact of an option trade is absorbed into option price and implied volatility, but not underlying price, such impacts are less likely to contain new information. The variance of the implied volatility impact is large. On average, bullish trades have negative implied volatility impact (-8.57bp) while bearish trades have positive implied volatility impact (12.79bp). In terms of moneyness, ITM option trades have the highest average implied volatility impact (4.65bp), as well as the highest implied volatility impact variance. This suggests that in comparison to OTM and NTM trades, ITM options trades are rarely based on new private information. Rather, those trades are based on liquidity needs or hedging purposes.

Insert Table 9 Here

With the implied volatility impact in mind, I perform a panel regression similar to that of table 5 and 6, with the same explanatory variables and model specifications. Table 10 reports the regression estimates. As the dependent variable is the implied volatility impact, a significant coefficient would suggest an association with volatility information. The coefficients for implied volatility impact are statistically insignificant for single-leg and OTM option trades when either firm or date fixed effects are introduced. This indicates that while these option trades may contain price information of the underlying stock, they do not contain volatility information. The coefficients for auction trades are significantly negative, indicating that trades through order book contain both price and volatility information for the underlying stock. Finally, ITM options contain both price and volatility information.

4.3 Price and Volatility Information of Option Trades

A stylized fact arises by comparing the results of table 6 and 10. While some option trade characteristics (limit order book and ITM trades) impact both underlying price and implied volatility in the same direction, some characteristics (Single-leg and OTM trades) impact in opposite directions. A price-informed option trade (positive underlying price impact) may or may not be volatility-informed (implied volatility impact may be positive, negative, or near zero). For instance, table 6 indicates that single-leg option trades have a higher underlying price impact than multi-leg option trades, indicating a higher probability of price-informed option trading. However, their implied volatility impacts are similar, indicating that they both hold similar levels of volatility information. The same phenomenon applies to OTM versus ATM trades. On the other hand, limit order book and ITM trades contain both more price information and volatility information than auction and ATM trades respectively.

Insert Table 10 Here

4.4 Linkage between the Option and Equity Markets: An Instrumental Variable Approach

To investigate the relationship across underlying price impact, option price impact, and implied volatility impact, I would like to perform a regression in the following form:

$$\begin{aligned}\text{Underlying Price Impact} = \alpha &+ \beta_1 \text{Option Price Impact} \\ &+ \beta_2 \text{Implied Volatility Impact} \\ &+ \beta_3 \text{Implied Volatility} + \epsilon\end{aligned}$$

However, there are endogeneity concerns because option and underlying price movements are co-integrated. To tackle this issue, I use option trading characteristics and implied volatility as instrumental variable for *Option Price Impact* and *Implied Volatility Impact*. The 2 stage least square (2SLS) regression is in the following form:

First-stage:

$$\text{Option }\widehat{\text{Price}}\text{ Impact} = \delta + \gamma_1 \text{IV} + \gamma_2 \text{Option trade characteristics}$$
$$\text{IV }\widehat{\text{impact}} = \mu + \theta_1 \text{IV} + \theta_2 \text{Option trade characteristics}$$

Second-stage:

$$\begin{aligned}\text{Underlying Price Impact} = \alpha &+ \beta_1 \text{Option }\widehat{\text{Price}}\text{ Impact} \\ &+ \beta_2 \text{Implied }\widehat{\text{Volatility}}\text{ impact} \\ &+ \beta_3 \text{Implied Volatility} + \epsilon\end{aligned}$$

Table 11 reports the esimtates of the IV regression model. Model (1) is the OLS model without consideration for endogeneity. Model (2) is the IV model. Model (3) is the first stage regression for *Option Price Impact*. Model (4) is the first stage regression for *Implied Volatility Impact*. In both models (1) and (2), the coefficients for option price impact are positive and statistically significant, while the coefficients for implied volatility impact are negative and statistically significant. This suggests that upon option trade, option price impact and underlying price impact are significantly linked, i.e. an option trade that causes changes in option price is also going to cause changes in the same direction for the underlying price. However, the direction of implied volatility impact is opposite to that of the underlying price impact. This result suggests that some informed option trading activities contain both price and volatility information, and when the latter causes a change in the implied volatility, it reduces the price discovery effects in the underlying market. In other words, as option trades contain both price and volatility information, only the price information is reflected in the underlying stock prices, while the volatility information is reflected in the change in implied volatility, which is more subtle and unlikely to be observed by the equity market.

Insert Table 11 Here

5 Conclusion

This paper investigates the linkage between the equity market and the option market by studying the price impacts in both markets upon informed option trades. The empirical analyses of this paper demonstrate that option trades associated with single-leg strategy and traded through the limit order book are more likely to come from informed investors and have permanent price impacts in both the option market and the equity market. However, information contained by option trades have two dimensions: price and volatility. Despite having permanent price impacts and hence containing price information, single-leg option trades do not contain a significant amount of volatility information. The results also show that only price information are reflected through the changes in the underlying price, while volatility information are "hidden" as they affect the implied volatility of options. In fact, the volatility part of the information contained by option trades may have eluded from researchers as they underestimated the information content of option trades. Future work in this direction may lead us to understand the price discovery process from another point of view, i.e. volatility information.

Figures

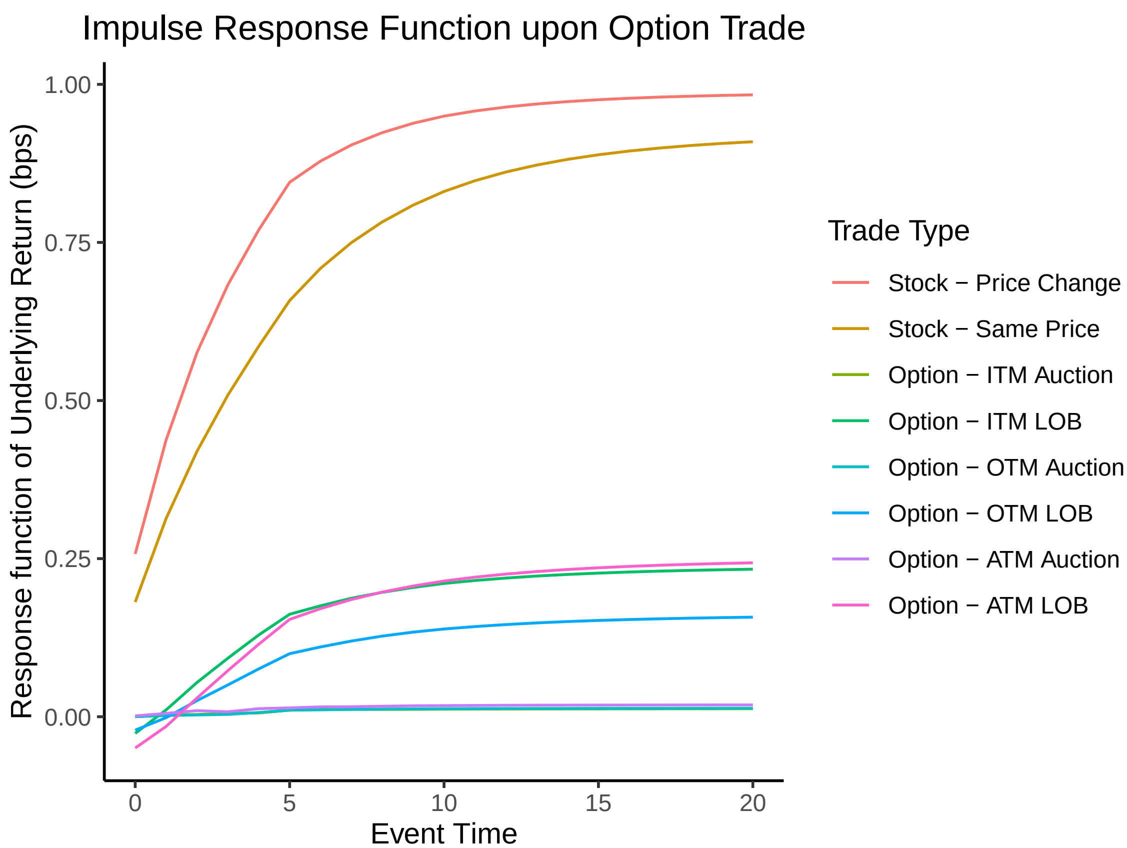

Figure 1: Impulse Response Function (IRF) of Underlying Return. This figure is a visualization of Table 4, presenting the IRF of underlying returns after specific types of trades. The y-axis is the response of underlying return estimated by the vectorautoregression model (Equation 4), measured in basis points. The x-axis is event time, which advances by 1 when any type of trade happen for the given underlying stock.

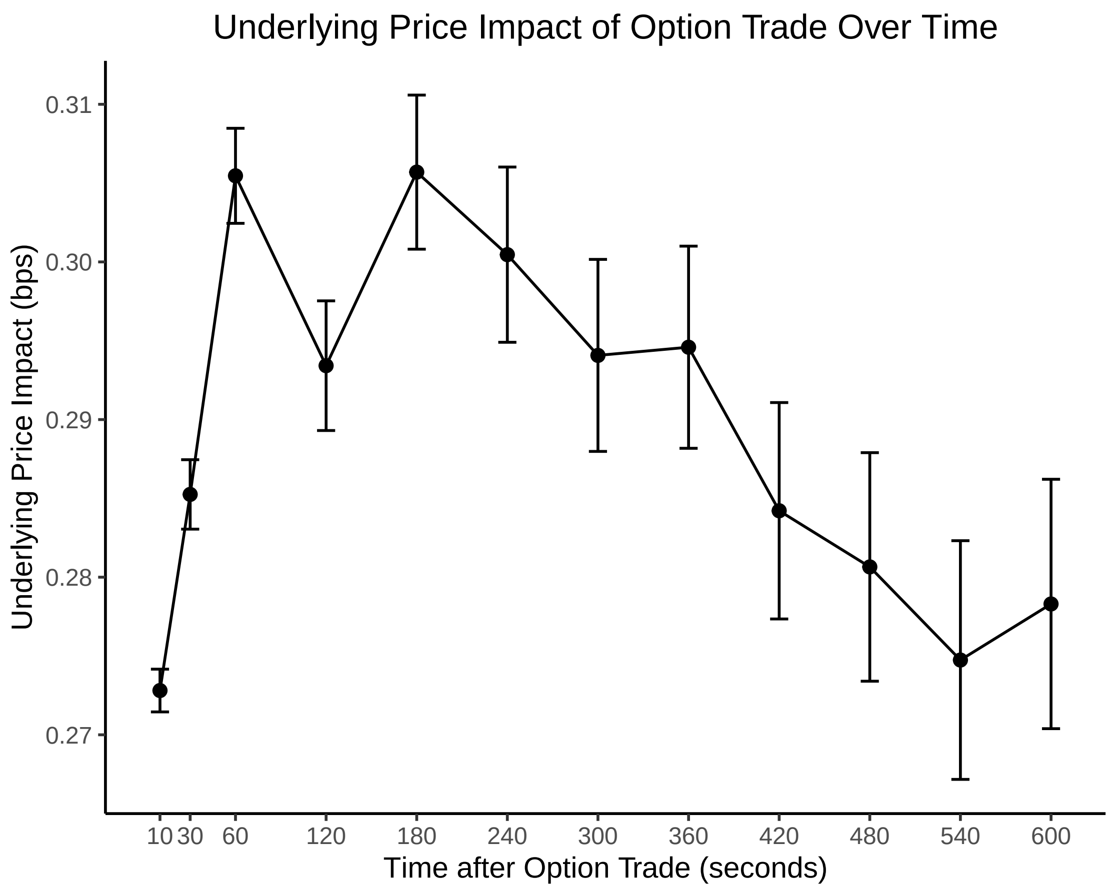

Figure 2: Average Underlying Price Impact after Option Trade, measured at different time intervals ranging from 10 seconds to 10 minutes. *Underlying Price Impact* is defined as the underlying return, in basis points, multiplied by the direction of the option trade sentiment (+1 for bullish, and −1 for bearish, where buying a call/put and selling a put/call are considered bullish/bearish).

Tables

Table 1: Option Trade Characteristics and Frequency

This table reports the frequency of option trade characteristics reported by OPRA. The sample period is from January 2nd to October 21st, 2020. Buying a call option or selling a put option are classified as bullish, while selling a call or buying a put are classified as bearish. The trading direction of option trades are determined by the tick test (Lee and Ready, 1991). ITM, OTM, and ATM are defined as option trades with moneyness greater than 3%, less than -3%, and in between -3% and 3% respectively. Moneyness is the difference between the option mid price and the strike price divided by the strike price, multiplied by -1 for put options. Auction/order book, Single-leg/Multi-leg are determined based on the trade condition ID provided by OPRA as in Appendix Table A2.

Trade Characteristic	Frequency
Call	64.04%
Put	35.96%
Bullish	50.11%
Bearish	49.89%
ITM	8.35%
OTM	49.99%
ATM	41.66%
Auction	22.85%
order book	77.15%
Single-leg	76.62%
Multi-leg	23.38%

Table 2: Descriptive Statistics

This table reports the descriptive statistics of all option trades included in the sample ($N = 138,180,628$). The sample period is from January 2nd to October 21st, 2020. *Price* is the reported option trading price. *Trade Size* is the reported option trading volume (in number of contracts). *Bid*, *Ask*, and *Mid* are reported NBBO quotes. *Moneyness* is the difference between underlying mid price and option strike price divided by option strike strike, and multiplied by -1 for put options. *Implied Volatility* is the implied volatility determined by binomial model for American Options. *Days to Expiration* is the time difference between expiration date at 4 p.m. and the trading datetime. *Underlying Bid*, *Underlying Ask*, and *Underlying Bid* are reported NBBO quotes for the underlying stock. *Quoted Spread* is the difference between option ask price and bid price divided by option mid price. *Effective Spread* is the difference between trade price and option mid price divided by option mid price, multiplied by 2, and multiplied by -1 for market sells. *Price Impact* is the difference between the option mid price 1 minute after trading and the option mid price at the time of trading, divided by the option mid price at the time of trading, multiplied by 2, and multiplied by -1 for market sells.

			Percentile				
	Mean	SD	5%	25%	50%	75%	95%
Price ($)	9.37	31.84	0.10	0.70	2.14	6.58	42.02
Trade Size (Contracts)	5.51	71.38	1.00	1.00	1.00	3.98	18.48
Bid ($)	9.22	31.60	0.08	0.67	2.08	6.44	41.50
Ask ($)	9.52	32.08	0.11	0.73	2.21	6.74	42.55
Mid ($)	9.37	31.84	0.10	0.70	2.15	6.59	42.02
Moneyness	-0.07	0.17	-0.32	-0.09	-0.03	0.00	0.06
Implied Volatility	0.59	0.43	0.22	0.34	0.48	0.71	1.35
Days to Expiration	38.72	91.48	0.00	2.94	8.46	30.53	196.27
Underlying Bid	381.15	642.73	14.82	57.46	156.05	368.73	1,970.12
Underlying Ask	381.37	643.16	14.83	57.48	156.11	368.91	1,971.44
Underlying Mid	381.26	642.95	14.83	57.47	156.08	368.82	1,970.76
Quoted Spread (bps)	8.81	20.04	0.76	1.77	3.44	7.71	34.11
Quoted Spread ($)	0.30	1.01	0.01	0.02	0.08	0.23	1.34
Effective Spread (bps)	6.69	20.42	0.01	0.90	2.05	5.01	26.04
Effective Spread ($)	0.15	0.83	0.00	0.01	0.04	0.12	0.68
Price Impact (bps)	0.59	19.10	-22.01	-3.99	0.01	4.84	24.30
Price Impact ($)	0.04	1.71	-1.09	-0.07	0.00	0.09	1.29

Table 3: Stock-day Average Variance Contribution

This table reports the frequency and the variance decomposition of the VAR model. *Trading Frequency* sum to 100% for all the trading types and shows the relative frequency of each type of trade. *Variance Contribution* indicates the average stock-day variance decomposition value outputted by the VAR model. A VAR model is fitted for each stock-day in the sample including all the stock and option trades of the same ticker and date. Each trade is categorized into exactly one of the following trade groups: *Underlying Trade - Price Change* captures all the stock trades with a trade size greater than the NBBO depth and hence moves the NBBO. *Underlying Trade - Same Price* captures all the stock trades with a smaller trade size than the NBBO depth and hence do not move the NBBO. *Option Trade - ITM Auction, ITM order book, OTM Auction, OTM order book, ATM Auction, ATM order book* capture all the options trades with their specific category in terms of moneyness (ITM/OTM/ATM) and pricing mechanism (auction/order book). Variance contribution of *Return* indicates the percentage of price variation that is not explained by stock or option trading.

	Variance Contribution	Trading Frequency
Underlying Trade		
Price Change	3.11%	4.03%
Same Price	40.90%	87.97%
Option trade		
ITM Auction	0.03%	0.15%
ITM order book	0.06%	0.52%
OTM Auction	0.03%	0.86%
OTM order book	0.09%	3.13%
ATM Auction	0.03%	0.82%
ATM order book	0.16%	2.52%
Return	55.59%	

Table 4: Cumulative Stock-day Average Return Impulse Response

This table reports the stock-day average return impulse responses upon a stock or option trade. Each trading category is a variable involved in the VAR model and can take the value of +1, 0, or -1. *Stock Trade - Price Change* takes the value of +1 if a market buy trade changes the NBBO price, -1 if a market sell trade changes the NBBO price, and 0 otherwise. *Stock Trade - Same Price* takes the value of +1 for a market buy trade that does not change the NBBO price, -1 for a market sell trade that does not change the NBBO price, and 0 otherwise. *Option Trade - ITM/OTM/ATM Auction/order book* takes the value of +1 for a bullish ITM/OTM/ATM option trade priced through auction/order book, -1 for a bearish ITM/OTM/ATM option trade priced through auction/order book, and 0 otherwise. Each trade can only be assigned to one of the trading categories.

	Stock Trade		Option Trade					
t	Price Change	Same Price	ITM Auction	ITM order book	OTM Auction	OTM order book	ATM Auction	ATM order book
0	0.26	0.18	0.00	-0.03	0.00	-0.02	0.00	-0.05
1	0.44	0.31	0.00	0.01	0.00	0.00	0.01	-0.02
2	0.58	0.42	0.00	0.05	0.00	0.03	0.01	0.03
3	0.68	0.51	0.01	0.09	0.00	0.05	0.01	0.07
4	0.77	0.59	0.01	0.13	0.01	0.08	0.01	0.11
5	0.85	0.66	0.01	0.16	0.01	0.10	0.01	0.15
6	0.88	0.71	0.01	0.18	0.01	0.11	0.02	0.17
7	0.90	0.75	0.01	0.19	0.01	0.12	0.02	0.19
8	0.92	0.78	0.01	0.20	0.01	0.13	0.02	0.20
9	0.94	0.81	0.01	0.20	0.01	0.13	0.02	0.21
10	0.95	0.83	0.01	0.21	0.01	0.14	0.02	0.21
11	0.96	0.85	0.01	0.22	0.01	0.14	0.02	0.22
12	0.96	0.86	0.01	0.22	0.01	0.15	0.02	0.23
13	0.97	0.87	0.01	0.22	0.01	0.15	0.02	0.23
14	0.97	0.88	0.01	0.22	0.01	0.15	0.02	0.23
15	0.98	0.89	0.01	0.23	0.01	0.15	0.02	0.24
16	0.98	0.89	0.01	0.23	0.01	0.15	0.02	0.24
17	0.98	0.90	0.01	0.23	0.01	0.15	0.02	0.24
18	0.98	0.90	0.01	0.23	0.01	0.16	0.02	0.24
19	0.98	0.91	0.01	0.23	0.01	0.16	0.02	0.24
20	0.98	0.91	0.01	0.23	0.01	0.16	0.02	0.24

Table 5: Underlying Price Impact upon Option Trade

This table reports the average of OLS regressions on the 1-minute signed price impacts (in basis points) for each stock-day on all option trades associated with the underlying stock. The sample period is from January 2 to October 21, 2020. Only stock-days with at least 50 option trades associated with the underlying were included in the sample. Column (1) includes all option trade characteristics as explanatory variables. Column (2) includes option trade characterstics and their associated interaction terms. Column (3) Includes option trade characteristics and control variables. Column (4) includes option trade characteristics, interaction terms, and control variables. *ITM* takes the value of 1 if the option contract's moneyness is greater than 0.03 at the time of the trade, and 0 otherwise. *OTM* takes the value of 1 if the option contract's moneyness is less than -0.03 at the time of the trade, and 0 otherwise. *Single-leg* takes the value of 1 if the option traded is involved in a single-leg trade, and 0 otherwise. *Auction* takes the value of 1 if the option trade is priced through auction instead of crossing or limit order book, and 0 otherwise. *Stock Volaility* is the absolute value of the past 10-second stock mid price return (in percent). *SPY volatility* is the absolute value of the past 10-second return of the S&P 500 exchange-traded fund, SPY (in percent). *Realized Spread* is the option's half quoted bid-ask spread relative to the midpoint price at the time of the trade (in percent). *Lagged Underlying Return* is the signed 10-second lagged return of the stock mid price (in percent). *Lagged SPY Return* is the signed 10-second SPY return (in percent). *Limit Order Book Imbalance* is defined as the option's (depth at best bid price - depth at best ask price)/(depth at best bid price + death at best ask price) * (1 if buy order, -1 if sell oder). *Implied volatility* is the volatility computed by the Binomial Option Pricing Model, given the mid prices of the option contract and the underlying at the time of the trade. Standard errors clustered to stock and date are reported. *, **, *** indicates statistically significance at 10%, 5%, and 1% respectively.

	(1)		(2)		(3)		(4)	
	mean	se	mean	se	mean	se	mean	se
ITM	-0.04**	0.02	-0.04	0.03	-0.04**	0.02	-0.05*	0.03
OTM	-0.28***	0.02	-0.13***	0.02	-0.22***	0.01	-0.08***	0.02
Single	0.44***	0.03	0.71***	0.05	0.39***	0.02	0.63***	0.04
Auction	-0.45***	0.03	-0.23***	0.02	-0.42***	0.02	-0.26***	0.03
ITM * Single			-0.05	0.04			-0.02	0.04
OTM * Single			-0.28***	0.03			-0.26***	0.03
ITM * Auction			0.14***	0.03			0.14***	0.04
OTM * Auction			0.14***	0.02			0.16***	0.02
Single * Auction			-0.56***	0.04			-0.51***	0.04
ITM * Single * Auction			-0.04	0.05			-0.03	0.05
OTM * Single * Auction			0.25***	0.04			0.25***	0.04
Trade Size					0.00**	0.00	0.00**	0.00
Underlying Volatility					-4.65***	0.14	-4.81***	0.14
SPY Volatility					0.18	0.51	0.12	0.51
Realized Spread					0.00***	0.00	0.00***	0.00
Lagged Underlying Return					0.47***	0.09	0.46***	0.10
Lagged SPY Return					-0.66**	0.29	-0.60**	0.29
Limit Order Book Imbalance					0.19***	0.02	0.19***	0.02
Implied Volatility					-0.20***	0.06	-0.22***	0.06
Intercept	0.37***	0.02	0.18***	0.02	0.56***	0.04	0.42***	0.04

Table 6: Underlying Price Impacts: Panel Regression

This table reports the regression coefficient of event time underlying price impacts on option trading characteristics. The dependent variable, *Event Time Underlying Price Impact*, is the log difference between the underlying mid price at the time of the next option trade of the same option contract and that at the time of the current option trade. Column (1) includes no fixed effects. Column (2) includes firm fixed effect. Column (3) includes date fixed effect. Column (4) includes both firm and date fixed effects. *Single-leg* takes the value of 1 if the current option trade contain a trade condition ID specifying a single-leg trade, and 0 otherwise. *Auction* takes the value of 1 if the current option trade contain a trade condition ID indicating that the trade involves a price improvement auction, and 0 otherwise. *Trade Size* is the number of option contracts traded. *Time Difference* is the time taken between the current option trade and the next option trade of the same option contract. *ITM* takes the value of 1 if the moneyness of the option traded was greater than 3%, and 0 otherwise. *OTM* takes the value of 1 if the moneyness of the option traded was less than -3%, and 0 otherwise. Standard errors clustered by firm and date are reported in parentheses. *, **, and *** marks coefficient at 10%, 5%, and 1% respectively.

	(1)	(2)	(3)	(4)
Single-leg	0.0028	0.0068*	-0.0013	0.0130***
	(0.0036)	(0.0037)	(0.0038)	(0.0037)
Auction	-0.0093***	-0.0078**	-0.0204***	-0.0170***
	(0.0036)	(0.0035)	(0.0036)	(0.0035)
Trade Size	0.00004	0.00005**	0.00005***	0.00006***
	(0.00003)	(0.00002)	(0.00002)*	(0.00002)
Time Difference	0.00019***	0.00021***	0.00026***	0.00021***
	(0.00004)	(0.00001)	(0.00001)	(0.00001)
ITM	0.015	0.034***	0.022***	0.032***
	(0.013)	(0.008)	(0.008)	(0.008)
OTM	0.052***	0.058***	0.071***	0.051***
	(0.003)	(0.003)	(0.003)	(0.003)
Intercept	-0.020***	X	X	X
	(0.004)			
Firm FE	N	Y	N	Y
Date FE	N	N	Y	Y

Table 7: Option Price Impacts: Panel Regression

This table reports the regression coefficient of event time option price impacts on option trading characteristics. The dependent variable, *event time option price impact*, is the log difference between the option mid price at the time of the next option trade of the same option contract and that at the time of the current option trade. Column (1) includes no fixed effects. Column (2) includes firm fixed effect. Column (3) includes date fixed effect. Column (4) includes both firm and date fixed effects. *Single-leg* takes the value of 1 if the current option trade contain a trade condition ID specifying a single-leg trade, and 0 otherwise. *Auction* takes the value of 1 if the current option trade contain a trade condition ID indicating that the trade involves a price improvement auction, and 0 otherwise. *Trade Size* is the number of option contracts traded. *Time Difference* is the time taken between the current option trade and the next option trade of the same option contract. *ITM* takes the value of 1 if the moneyness of the option traded was greater than 3%, and 0 otherwise. *OTM* takes the value of 1 if the moneyness of the option traded was less than -3%, and 0 otherwise. Standard errors clustered by firm and date are reported in parentheses. *, **, and *** marks coefficient at 10%, 5%, and 1% respectively.

	(1)	(2)	(3)	(4)
Single-leg	2.06***	1.95***	1.41***	1.69***
	(0.094)	(0.098)	(0.099)	(0.098)
Auction	-1.11***	-1.29***	-1.29***	-1.31***
	(0.086)	(0.093)	(0.094)	(0.093)
Trade Size	0.0035***	0.0028***	0.0029***	0.0030***
	(0.0007)	(0.0005)	(0.0005)	(0.0005)
Time Difference	-0.062***	-0.062***	-0.061***	-0.062***
	(0.0007)	(0.0002)	(0.0002)	(0.0002)
ITM	4.47***	3.41***	1.39***	2.59***
	(0.135)	(0.210)	(0.217)	(0.211)
OTM	3.77***	3.58***	2.60***	2.74***
	(0.084)	(0.083)	(0.087)	(0.085)
Intercept	-4.37***	X	X	X
	(0.097)			
Firm FE	N	Y	N	Y
Date FE	N	N	Y	Y

Table 8: Option Trade Type and Impact Directions

This tables shows the direction of implied volatility impact and underlying price impact upon a hypothetical informed option trade.

	Implied Volatility Impact	Underlying Price Impact
Buy Call	+	+
Sell Call	-	+
Buy Put	+	-
Sell Put	-	+

Table 9: Implied Volatility Impacts: Summary

This table reports the summary statistics of the implied volatility impact of option trades. The total number of observations is 138,180,628. The subsample frequencies are reported in percentages of the total sample size. Option trades are classified into *Auction* or *order book* based on OPRA trade condition ID as specified in table A2. *ITM*, *OTM*, and *ATM* are defined as option trades with moneyness greater than 3%, less than -3%, and in between -3% and 3% respectively, where moneyness is the difference between the option mid price and the strike price divided by the strike price, multiplied by -1 for put options. An option is classified as *Bullish* if it is a buyer-initiated call trade or seller-initiated put trade. An option is classified as *Bearish* if it is a buyer-initiated put trade or seller-initiated call trade.

				Percentiles				
	Frequency	Mean	SD	5%	25%	50%	75%	95%
Full Sample	100.00%	2.12	267.28	-209.51	-37.80	0.02	39.79	217.27
Auction	24.02%	1.29	247.96	-199.95	-35.34	0.06	39.01	201.94
order book	75.98%	2.38	273.10	-215.02	-39.03	0.01	40.48	224.12
ITM	4.00%	4.65	754.66	-792.35	-106.41	0.01	110.42	810.02
OTM	45.79%	2.22	135.82	-145.37	-27.92	0.01	30.02	153.66
ATM	50.21%	1.82	283.04	-248.12	-47.77	0.03	49.49	255.38
Call	68.22%	1.77	240.75	-202.96	-37.08	0.02	38.60	209.69
Put	31.78%	2.87	316.80	-225.17	-39.66	0.02	42.81	235.15
Bullish	49.97%	-8.57	267.29	-227.79	-46.51	0.00	31.81	200.55
Bearish	50.03%	12.79	266.85	-192.04	-29.03	0.10	47.81	235.72
BuyCall	33.76%	33.51	242.34	-142.32	-15.17	3.30	59.75	266.47
BuyPut	15.51%	37.56	322.96	-164.14	-18.46	4.77	65.47	298.55
SellCall	34.45%	-29.34	235.07	-254.31	-58.00	-2.27	17.88	147.42
SellPut	16.27%	-30.20	307.19	-277.12	-59.67	-0.94	21.93	168.15

Table 10: Event Time Implied Volatility Impacts

This table reports the regression coefficient of event time implied volatility impacts on option trading characteristics. The dependent variable, *event time implied volatility impact*, is the log difference between the implied volatility at the time of the next option trade of the same option contract and that at the time of the current option trade. Column (1) includes no fixed effects. Column (2) includes firm fixed effect. Column (3) includes date fixed effect. Column (4) includes both firm and date fixed effects. *Single-leg* takes the value of 1 if the current option trade contain a trade condition ID specifying a single-leg trade, and 0 otherwise. *Auction* takes the value of 1 if the current option trade contain a trade condition ID indicating that the trade involves a price improvement auction, and 0 otherwise. *Trade Size* is the number of option contracts traded. *Time Difference* is the time taken between the current option trade and the next option trade of the same option contract. *ITM* takes the value of 1 if the moneyness of the option traded was greater than 3%, and 0 otherwise. *OTM* takes the value of 1 if the moneyness of the option traded was less than -3%, and 0 otherwise. Standard errors clustered by firm and date are reported in parentheses. *, **, and *** marks coefficient at 10%, 5%, and 1% respectively.

	(1)	(2)	(3)	(4)
Single-leg	-0.108*	0.088	-0.070	0.009
	(0.0581)	(0.0570)	(0.0557)	(0.0560)
Auction	-1.35***	-1.25***	-1.27***	-1.23***
	(0.0512)	(0.0538)	(0.0534)	(0.0536)
Trade Size	0.0007*	0.0008***	0.0006**	0.0008***
	(0.0004)	(0.0003)	(0.0003)	(0.0003)
Time Difference	0.025***	0.025***	0.025***	0.025***
	(0.0003)	(0.0001)	(0.0001)	(0.0001)
ITM	1.22***	1.89***	1.32***	1.75***
	(0.3227)	(0.1244)	(0.1191)	(0.1206)
OTM	-0.36***	0.035	-0.080*	0.033
	(0.0396)	(0.0498)	(0.0476)	(0.0484)
Intercept	0.83***	X	X	X
	(0.0612)			
Firm FE	N	Y	N	Y
Date FE	N	N	Y	Y

Table 11: Underlying Price Impact On Option Trade: Instrumental Variable approach

This table reports the instrument variable regression analysis of the underlying price impact on option price impact, and implied volatility impact as endogenous explanatory variables, with option trade characteristics as exogenous explanatory variables. Model (1) reports the OLS regression results. Model (2) reports the instrumental variable regression results. Model (3) reports the first stage regression of option price impact on exogenous explanatory variables. Model (4) reports the first stage regression of implied volatility impact on exogenous explanatory variables. Standard errors clustered by firm and date are reported in parentheses. *, **, and *** marks coefficient at 10%, 5%, and 1% respectively.

	(1) OLS	(2) Inst. Var.	(3) FS (Option Price Impact)	(4) FS (IV Impact)
Option Price Impact	1.173***	1.035***		
	(0.103)	(0.152)		
Implied Volatility Impact	-2.094***	-2.136***		
	(0.155)	(0.252)		
Implied Volatility	-0.933**	-2.947	-70.280***	-45.851***
	(0.405)	(7.633)	(3.949)	(2.378)
Single-leg			2.714	1.613
			(3.071)	(1.927)
Auction			3.339	2.186
			(3.073)	(1.930)
Trade Size			-0.002	-0.001
			(0.003)	(0.002)
Time Difference			0.000	0.000
			(0.000)	(0.000)
ITM			-0.018	0.949***
			(0.074)	(0.117)
OTM			0.908***	0.113
			(0.111)	(0.072)
Intercept	0.488***	2.028	23.457***	15.124***
	(0.183)	(2.726)	(3.266)	(1.999)

References

Andersen, Torben, Ilya Archakov, Leon Grund, Nikolaus Hautsch, Yifan Li, Sergey Nasekin, Ingmar Nolte, Manh Cuong Pham, Stephen Taylor, and Viktor Todorov, 2021, A descriptive study of high-frequency trade and quote option data, *Journal of Financial Econometrics* 19, 128–177.

Back, Kerry, 1993, Asymmetric information and options, *The Review of Financial Studies* 6, 435–472.

Black, Fischer, and Myron Scholes, 1973, The pricing of options and corporate liabilities, *Journal of political economy* 81, 637–654.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan, 2019, Price discovery without trading: Evidence from limit orders, *The Journal of Finance* 74, 1621–1658.

Bryzgalova, Svetlana, Anna Pavlova, and Taisiya Sikorskaya, 2023, Retail trading in options and the rise of the big three wholesalers, *The Journal of Finance* 78, 3465–3514.

Busch, Thomas, Bent Jesper Christensen, and Morten Ørregaard Nielsen, 2011, The role of implied volatility in forecasting future realized volatility and jumps in foreign exchange, stock, and bond markets, *Journal of Econometrics* 160, 48–57.

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

Christensen, Bent J, and Nagpurnanand R Prabhala, 1998, The relation between implied and realized volatility, *Journal of financial economics* 50, 125–150.

Cox, John C, Stephen A Ross, and Mark Rubinstein, 1979, Option pricing: A simplified approach, *Journal of financial Economics* 7, 229–263.

Easley, David, Maureen O'hara, and Pulle Subrahmanya Srinivas, 1998, Option volume and stock prices: Evidence on where informed traders trade, *The Journal of Finance* 53, 431–465.

Glosten, Lawrence R, and Paul R Milgrom, 1985, Bid, ask and transaction prices in a specialist market with heterogeneously informed traders, *Journal of financial economics* 14, 71–100.

Hasbrouck, Joel, 1991a, Measuring the information content of stock trades, *The Journal of Finance* 46, 179–207.

Hasbrouck, Joel, 1991b, The summary informativeness of stock trades: An econometric analysis, *The Review of Financial Studies* 4, 571–595.

Hasbrouck, Joel, 1995, One security, many markets: Determining the contributions to price discovery, *The journal of Finance* 50, 1175–1199.

Hendershott, Terrence, Charles M Jones, and Albert J Menkveld, 2011, Does algorithmic trading improve liquidity?, *The Journal of finance* 66, 1–33.

Hendershott, Terrence, Saad Khan, and Ryan Riordan, 2022, Option auctions, *Available at SSRN 4110516* .

Holowczak, Richard, Yusif E Simaan, and Liuren Wu, 2006, Price discovery in the us stock and stock options markets: A portfolio approach, *Review of Derivatives Research* 9, 37–65.

Hsieh, Shu-Fan, Chia-Ying Chan, and Ming-Chun Wang, 2020, Retail investor attention and herding behavior, *Journal of Empirical Finance* 59, 109–132.

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

Kyle, Albert S, 1985, Continuous auctions and insider trading, *Econometrica: Journal of the Econometric Society* 1315–1335.

Lee, Charles MC, and Mark J Ready, 1991, Inferring trade direction from intraday data, *The Journal of Finance* 46, 733–746.

Lipson, Marc L, Davide Tomio, and Jiang Zhang, 2023, A real cost of free trades: Retail option trading increases the volatility of underlying securities, *Available at SSRN 4383463* .

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

Ni, Sophie X, Jun Pan, and Allen M Poteshman, 2008, Volatility information trading in the option market, *The Journal of Finance* 63, 1059–1091.

Ni, Sophie X, Neil D Pearson, Allen M Poteshman, and Joshua White, 2021, Does option trading have a pervasive impact on underlying stock prices?, *The Review of Financial Studies* 34, 1952–1986.

Appendix

Table A1: OPRA Trade Condition ID Description
This table reports the descriptions for the OPRA trade conditional IDs.

Trade Condition ID	Condition Name	Condition Description
18	AutoExecution	Transaction was executed electronically. Prefix appears solely for information; process as a regular transaction.
21	Reopen	Transaction is a reopening of an option contract in which trading has been previously halted. Prefix appears solely for information; process as a regular transaction.
40	Cancel	Transaction previously reported (other than as the last or opening report for the particular option contract) is now to be cancelled.
41	CANCLAST	Transaction is the last reported for the particular option contract and is now cancelled.
42	CANCOPEN	Transaction was the first one (opening) reported this day for the particular option contract. Although later transactions have been reported, this transaction is now to be cancelled.
43	CANCONLY	Transaction was the only one reported this day for the particular option contract and is now to be cancelled.
95	IntermarketSweep	Transaction was the execution of an order identified as an Intermarket Sweep Order. Process like normal transaction.
108	Trade through Exempt	Transaction is Trade Through Exempt. The transaction should be treated like a regular sale.
114	SingLegAuctNon-ISO	Transaction was the execution of an electronic order which was "stopped" at a price and traded in a two sided auction mechanism that goes through an exposure period. Such auctions mechanisms include and not limited to Price Improvement, Facilitation or Soliciation Mechanism.

115	SingLegAuctISO	Transaction was the execution of an Intermarket Sweep electronic order which was "stopped" at a price and traded in a two sided auction mechanism that goes through an exposure period. Such auctions mechanisms include and not limited to Price Improvement, Facilitation or Solicitation Mechanism marked as ISO.
116	SingLegCross-NonISO	Transaction was the execution of an electronic order which was "stopped" at a price and traded in a two sided crossing mechanism that does not go through an exposure period. Such crossing mechanisms include and not limited to Customer to Customer Cross and QCC with a single option leg.
118	SingLegFlr	Transaction represents a non-electronic trade executed on a trading floor. Execution of Paired and Non-Paired Auctions and Cross orders on an exchange floor are also included in this category.
119	MultLegAutoEx	Transaction represents an electronic execution of a multi leg order traded in a complex order book.
120	MultLegAuct	Transaction was the execution of an electronic multi leg order which was "stopped" at a price and traded in a two sided auction mechanism that goes through an exposure period in a complex order book. Such auctions mechanisms include and not limited to Price Improvement, Facilitation or Solicitation Mechanism.
121	MultLegCross	Transaction was the execution of an electronic multi leg order which was "stopped" at a price and traded in a two sided crossing mechanism that does not go through an exposure period. Such crossing mechanisms include and not limited to Customer to Customer Cross and QCC with two or more options legs.
122	MultLegFlr	Transaction represents a non-electronic multi leg order trade executed against other multi-leg order(s) on a trading floor. Execution of Paired and Non-Paired Auctions and Cross orders on an exchange floor are also included in this category.
123	MultLegAutoSin-gLeg	Transaction represents an electronic execution of a multi Leg order traded against single leg orders/ quotes.

124	StkOptAuct	Transaction was the execution of an electronic multi leg stock/options order which was "stopped" at a price and traded in a two sided auction mechanism that goes through an exposure period in a complex order book. Such auctions mechanisms include and not limited to Price Improvement, Facilitation or Solicitation Mechanism.
125	MultLegAuctSingLeg	Transaction was the execution of an electronic multi leg order which was "stopped" at a price and traded in a two sided auction mechanism that goes through an exposure period and trades against single leg orders/ quotes. Such auctions mechanisms include and not limited to Price Improvement, Facilitation or Solicitation Mechanism.
126	MultLegFlrSingLeg	Transaction represents a non-electronic multi leg order trade executed on a trading floor against single leg orders/ quotes. Execution of Paired and Non-Paired Auctions on an exchange floor are also included in this category.
127	StkOptAutoEx	Transaction represents an electronic execution of a multi leg stock/options order traded in a complex order book.
128	StkOptCross	Transaction was the execution of an electronic multi leg stock/options order which was "stopped" at a price and traded in a two sided crossing mechanism that does not go through an exposure period. Such crossing mechanisms include and not limited to Customer to Customer Cross.
129	StkOptFlr	Transaction represents a non-electronic multi leg order stock/options trade executed on a trading floor in a Complex order book. Execution of Paired and Non-Paired Auctions and Cross orders on an exchange floor are also included in this category.

Table A2: Trade Characteristics and Trade Condition ID mapping

This table presents the classification for trade characteristics: Auction and order book for pricing mechanism. Single-leg and Multi-leg for trading strategy. Canceled trades (Trade Condition ID 40, 41, 42, 43) and floor trades (Trade Condition ID 118, 122, 129) are removed from the sample.

TradeConditionID	Condition Name	Cancel	Auction	Orderbook	Single-leg	Multi-leg
18	AutoExecution					
21	Reopen					
40	Cancel	X				
41	CANCLAST	X				
42	CANCOPEN	X				
43	CANCONLY	X				
95	IntermarketSweep			X	X	
108	Trade through Exempt			X		
114	SingLegAuctNonISO		X		X	
115	SingLegAuctISO		X		X	
116	SingLegCrossNonISO			X	X	
118	SingLegFlr				X	
119	MultLegAutoEx			X		X
120	MultLegAuct		X			X
121	MultLegCross			X		X
122	MultLegFlr					X
123	MultLegAutoSingLeg			X		X
124	StkOptAuct		X			X
125	MultLegAuctSingLeg		X			X
126	MultLegFlrSingLeg		X			X
127	StkOptAutoEx			X		X
128	StkOptCross			X		X
129	StkOptFlr					X