Simple Roles for Complex Options*

Su Li
U.S. Securities and Exchange Commission

David K. Musto
Wharton School, University of Pennsylvania

Neil D. Pearson
University of Illinois at Urbana-Champaign and CDI Research Fellow

First Draft: November 18, 2022

This Draft: March 4, 2024

Abstract

Among the many possible complex options trades, just a few dominate the market. Two simple roles largely explain their use. Using a new approach to identify complex trades, we find that vertical, vertical ratio, calendar, and diagonal spreads account for most complex volume, and volatility trades such as straddles and strangles account for a much smaller fraction. Many trades are executed not to obtain the payoffs of the complex packages, but instead to adjust simple options positions by changing either strikes or expiration dates. Others appear intended to make simple bets on price movements with small initial investments.

JEL codes: G13, G23, G24

Keywords: options trading, complex trade protocols, complex options trades, rolling positions

---

* This paper is part of the Division of Economic and Risk Analysis' (DERA's) Working Paper Series. Papers in this series investigate a broad range of issues relevant to the Commission's mission and are preliminary materials disseminated to stimulate discussion and critical comment. Inclusion of a paper in this series does not indicate a Commission determination to take any particular action. References to this paper, and all papers in the DERA Working Paper Series, should indicate that the paper is a "DERA Working Paper." The Securities and Exchange Commission disclaims responsibility for any private publication or statement of any SEC employee or Commissioner. This paper expresses the authors' views and does not necessarily reflect those of the Commission, the Commissioners, or members of the staff.

The authors thank seminar participants at the University of Melbourne, Deakin University, Wharton, the CDI Conference and discussant Ilias Filippou, FINRA and the Chicago Conference on Empirical Finance for useful comments. Contact information: Su Li: Office of Markets, U.S. Securities and Exchange Commission, 100 F Street, NE Washington, DC 20549, email: lisu@sec.gov, voice: 202-551-7006. David K. Musto: Wharton School of the University of Pennsylvania, 3620 Locust Walk, Philadelphia, PA 19104, email: musto@upenn.edu, voice: 215-898-4239. Neil D. Pearson, Department of Finance, University of Illinois at Urbana-Champaign, 1206 S. Sixth Street, Champaign, IL 61820, email: pearson2@illinois.edu, voice: 217-244-0490.

Simple Roles for Complex Options

Abstract

Among the many possible complex options trades, just a few dominate the market. Two simple roles largely explain their use. Using a new approach to identify complex trades, we find that vertical, vertical ratio, calendar, and diagonal spreads account for most complex volume, and volatility trades such as straddles and strangles account for a much smaller fraction. Many trades are executed not to obtain the payoffs of the complex packages, but instead to adjust simple options positions by changing either strikes or expiration dates. Others appear intended to make simple bets on price movements with small initial investments.

JEL codes: G13, G23, G24

Keywords: options trading, complex trade protocols, complex options trades, rolling positions

I. Introduction

Textbooks (e.g., Hull (2009)) show how to combine simple options positions on an underlying into complex ones – vertical, calendar, and diagonal spreads, vertical ratios and butterflies, straddles and strangles, etc. – and their payoff diagrams suggest motives for using them. But what are their actual roles in securities markets? Is it to meet a demand for those payoffs, or is it something else the complex packages can do? We address these questions by identifying complex trades, characterizing the circumstances in which they are used, and establishing their net effect by relating them to changes in open interest. We find that investors use complex options to adjust simple positions and lower the cost of directional bets, and that betting on volatility plays a smaller role.

Trading complex options directly, as opposed to assembling complex positions through simple trades, has become increasingly easy and popular over the years. ISE offered the first complex limit-order book (C-LOB) in 2002, followed by the CBOE in 2005, and today there are 11 U.S. options exchanges with either electronic or manual complex execution protocols. From 2011 to 2021,we find that complex volume grew as a fraction of total volume from 26 to 38 percent through 2019 (see Figure 1), before dropping somewhat during the pandemic boom in retail volume.

What explains this popularity? The payoff diagrams and Greeks of complex options suggest possible motives to trade them, and the internet is full of advice, but the forces that drive complex trading volume remain a mystery. Investors might want those payoffs and Greeks, but they might instead want the changes the complex trades impart to their *existing* simple or complex options positions. That is, a complex trade might be a convenient way to close one position and open another. We gauge the significance of this motive by relating complex volume back to the circumstances that encourage adjustment of simple positions. We also relate it forward to changes in open interest, as legs of complex that close (open) options positions would decrease (increase) open interest.

The first step in this analysis is identifying complex trades in the data. Our sample period is 2016-2018, and our sample consists of all options on the 30 listed firms and 30 exchange-traded products (ETPs) with the most options trading volume, though in fact there are 31 stocks in the sample of listed firms due to Alphabet having two share classes with listed options (ticker symbols GOOG and GOOGL). The Option Price Reporting Authority (OPRA) reports the legs of complex trades together with the simple trades and includes trade condition codes that indicate whether the trade was part of a complex package. The codes do not, however, identify the other trades that were part of the package. We identify the specific complex trades by grouping trades that are flagged as components of complex trades into packages by exploiting the fact that options trades on the same underlying stock that are flagged as legs of complex trades and reported within a few milliseconds are almost certain to be legs of a single complex package. The type of complex trade (vertical spread, calendar spread, etc.) is usually clear, though occasionally ambiguous when there is more than one complex package that involves the same legs. For example, both a straddle and a combination consist of trades in a put and a call with the same strike and expiration, with the difference being that in a straddle the put and call are both purchased or sold, while in a combination, one is purchased and one is sold. While we can estimate the direction of the overall package in the usual way by relating its cost to the prevailing quotes for its component legs, we cannot distinguish between a straddle and a combination by signing the legs because the package is traded with a single net price.

Our main result from this identification is that trading of various spreads dominates complex trading volume, and volatility trading accounts for only a small fraction of complex trading. Spreads are verticals, calendars, diagonals, and vertical ratios. Each of these complex trades involves two legs that are either both calls or both puts, differ in strike and/or expiration, and in the case of vertical ratios, size. They are therefore suited to adjust the strike and/or expiration, and in the case of vertical ratios, also the size, of simple options positions. Volatility trades consist of straddles, strangles, butterflies, and condors that provide exposure to changes in volatility, and are not suited to make adjustments to simple options positions. We find that spreads account for 66% of complex volume in stock options and 57% in ETP options, with verticals the most common. Volatility trades, in contrast, account for at most 16% of complex volume, and this much only if all the trades that could be either combinations or straddles are deemed to be straddles. This low frequency of volatility trades is consistent with the low incidence of such trades in Lakonishok et al. (2007), based on older data. Relating the spread trades to the underlying price, we find that spreads with both legs out-of-the-money (OTM) are most frequent, and spreads with both legs ITM are least frequent.

The next step is to assess the use of spreads to adjust simple positions, and we start by relating spreads to the changes in circumstances likely to encourage adjustment. We consider changes in moneyness, which could encourage adjustment for two reasons. One is simply that the investor prefers the original moneyness and wants to restore it. For example, an investor who insures against market breaks by buying out-of-the-money puts might move strikes up or down as the market moves up or down to maintain the moneyness of the protective put position. The other reason is that an investor might want to close out positions in options to avoid the cash outlay required to exercise an option, or to avoid the need to deliver or accept delivery of the underlying stock. For example, an investor who writes out-of-the-money calls that become in-the-money as the stock price moves and who wants to keep the underlying stock might prefer to realize the paper loss by moving the strike up rather than get assigned to deliver the underlying shares. So the return on the underlying should drive spread trades, especially in the direction of the return: increases in the underlying should drive vertical call spread trades, especially those with just one leg in-the-money that the spread could be closing, and decreases in the underlying should drive vertical put spread trades, again especially those with just one leg in-the-money. The same logic applies to vertical ratio trades that change strike and size, to diagonal spreads that change strike and expiration, and even to calendar spreads that change only expiration since moving the expiration date out ordinarily delays option exercise.

We test for this use of spreads with eight predictive regressions representing the four spread types in both puts and calls. The regressions use recent returns to predict the daily volume of spreads with the near leg in-the-money. The results show that positive returns consistently predict call spread volumes and negative returns consistently predict put spread volumes.

To test for this use of complex options trades from a different direction, we estimate whether they are closing and opening positions by relating them to changes in open interest. If an investor uses a spread to close a position in an option, then aggregate open interest in that option either goes down, if the trader on the other side of the trade is also closing a position, or it stays the same, if the trader on the other side is opening a position. Likewise, open interest in an option in which an investor is opening a position either stays the same or increases. So, we can estimate the tendency for a given leg of a complex type, e.g. the near leg of a calendar spread, to open or close a position by relating the change in an option's open interest to the same-day volume of legs in that circumstance, e.g., the volume of trade in that option occurring in near legs of calendar spreads. We run this test with regressions that explain changes in open interest with explanatory variables that decompose all volume that day into such categories, so that positive coefficients reflect a tendency to open positions and negative coefficients reflect a tendency to close positions.

The results from the open-interest regressions strongly associate diagonals, calendars, verticals, and vertical ratio trades with adjusting simple positions. In the case of diagonals and calendars, the near-expiration leg closes and the far leg opens, and in the case of verticals and vertical ratios, the ITM leg closes and the OTM leg opens. The results for non-ratio vertical spreads with both legs OTM indicate a higher incidence of trades that open positions in both legs, rather than opening a position in one leg and closing a position in the other. We also find evidence that verticals roll down existing short stock call positions (e.g., adjusting covered calls).

The evidence that OTM vertical spreads tend to open positions in both legs raises the question of why investors prefer them over simple options trades. One possible reason is that selling a vertical spread caps the possible loss, whereas a sold call or put does not. Another possible reason is to save money. A long position in a put or call might cost more than an investor wants to spend. The investment can be reduced by selling an option of the same type (put or call) strike, that is, by buying a vertical spread. This consideration can be important in the options market, compared to the stock market, because an investor can buy fewer than 100 shares but cannot buy options on fewer than 100 shares. This minimum, especially for high-priced underlying stocks, could make a vertical spread more desirable than a simple position in near-the-money options to an investor seeking to spend less. We explore this motive to use spreads by using stock splits to examine what happens when the split's exogenous change to the price of the underlying relaxes the budget constraint. We track the stocks that split relative to a control sample, and consistent with this motive, we find that the complex fraction of their option volume drops across the splits, and more so with greater split factors.

Finally, we assess the use of volatility trades with similar regressions explaining changes in open interest, asking whether the initiators tend to be buying or selling volatility.[1] When we break out the volatility trades by whether the initiator was (judging from quotes) buying or selling we find traders more likely to buy volatility with straddles, in that these trades associate

[1]"Buying" ("selling") means entering into an options position that increases (decreases) in value when volatility increases, that is, entering into a position for which vega is positive (negative). While a debit spread is required to volatility using straddles, strangles, iron butterflies, and iron condors, buying volatility through a butterfly spread involves a credit spread, which entails collecting premium on the trade date.

more with increased open interest, and to sell volatility with strangles, butterflies, iron condors and iron butterflies, especially in ETP options.

This paper is in seven sections. Section II covers the institutional background on complex trades and the intraday OPRA data, and Section III describes the sample. Section IV presents an overview of complex options trading and some initial results, including the finding that spread trades account for the majority of complex volume. Section V focuses on the spread trades and their motivations, while Section VI considers the volatility trades. Section VII summarizes and concludes. The variables used in some of the regression analyses are defined in Appendix A. The algorithm to infer complex strategies from the OPRA data is described in Appendix B.

II. Background on Complex Trades and Market Structure

II.A Complex Trades

On a given underlying asset there can be many listed options contracts (series) with different strike prices, expirations, and types. For example, on 6/29/2018 the OptionMetrics data show 6,048 unique options on SPY with 275 distinct strikes and 34 distinct expirations. They can trade separately or as packages, and certain packages are common enough to be named, including vertical, vertical ratio, calendar, and diagonal spreads, butterfly spreads, straddles, strangles, condors, and others.[2] The more common packages and their names are collected for reference in Table 1.

There are multiple possible motives behind a complex trade. The trade's payoff might simply match the bet or hedge the trader desires. It could also save money, for example when a trader finds buying an at- or near-the-money call too expensive and reduces this cost by adding a written out-of-the money call through a call spread. It can implement a bet on volatility by conveying equal exposure to up and down movements in the underlying, as in at-the-money straddles, strangles, butterflies, and condors do. And it can adjust simpler positions if a subset of the legs offsets an existing position and the remainder of the legs opens a new position with different strikes or expirations.

Broker-dealers and options exchanges facilitate complex trades with specific order types containing instructions for multiple legs. Such orders are called complex orders, and quoted in

[2] In these packages, one of the legs can be the underlying itself. For example, a covered call consists of a long position in the underlying stock combined with a written call on the same number of shares.

the form of a net price.[3] Traders could instead assemble the legs separately, but this risks large transaction costs since it can mean crossing the NBBO spread multiple times and it also risks not completing the trade, or getting picked off if the trader tries to economize with non-marketable limit orders. The savings from trading via complex orders can be substantial: for example, Li, Musto and Pearson (2023) show that buying vertical spreads through complex orders rather than two single-leg trades can save up to 83% on transactions costs.

II.B Market Structure

Complex orders have had their own C-LOBs since the ISE introduced the first in 2002. Before that, complex orders could only be executed manually by floor brokers through telephone or hand signals inside the trading pits. C-LOBs operate similarly to simple limit order books (S-LOBs), except that typically market makers do not maintain continuous bid and ask quotes. Instead, it is usually the customers or proprietary trading firms that post liquidity or trading interests, and the market makers respond by trading or not. So usually only one side of a complex package is quoted on a C-LOB. Also, these quotes are not as visible as the quotes for simple trades: the best bids and offers on a market's S-LOB are publicized through the OPRA Securities Information Processor (SIP) but those from the C-LOBs are not, partly due to the very large number of possible complex trades. Market participants also find liquidity for complex trades, especially for larger trades, on the increasingly electronic open-outcry trading floors.

Complex price improvement auctions (C-PIAs) are another important execution protocol for complex orders. Executing brokers can execute both retail and institutional trades through these auctions, instead of routing them to a C-LOB or floor broker. To initiate a C-PIA, the affiliated market maker of an executing broker submits a two-sided order representing a customer's complex order and its own "contra" complex order to the exchange.[4] The C-PIA typically lasts for 100 milliseconds, during which time the exchange exposes and broadcasts the order to other exchange members to allow them to improve the price to the customer. After the C-PIA concludes, the trade is allocated by price across the original contra order and any additional responding orders.

[3] For example, see the Cboe Complex Book Process, available at https://cdn.cboe.com/resources/membership/US-Options-Complex-Book-Process.pdf.
[4] The market maker also must be a member of the exchange that provides the C-PIA protocol.

Some exchanges allow certain complex orders to execute legs through the S-LOBs if the net price is "marketable" against orders in the S-LOBs and there is enough liquidity from the S-LOB for each leg. Li, Musto and Pearson (2023) show that complex orders executed through S-LOBs tend to be the costliest in terms of effective spread among the electric complex execution protocols, controlling for the characteristics of the orders, followed by the orders executed on the C-LOB. The C-PIAs tend to provide the best price improvements relative to the net NBBOs derived from the NBBOs on the S-LOBs.

Recent years have seen a proliferation of market centers listing options, and a proliferation of complex execution protocols introduced by those market centers to compete for the increased complex order flow. At the end of 2012, when complex volume was 27% of total volume, there were eleven options exchanges where market participants could execute trades in listed options, six of which provided complex execution protocols. In 2016, seven of the fourteen U.S. options exchanges offered C-LOBs and/or C-PIAs, and by January 2022 eleven of the sixteen U.S. options exchanges, operated by five exchange groups, offered complex execution protocols.[5] Some exchanges such as the CBOE, AMEX, and ARCA continue to allow floor market makers to execute complex orders through open outcry platforms. The landscape of the complex functionalities across the options exchanges is shown in Table 2, which marks the exchanges launched between 2016 and 2020 with a single star and the exchanges that added electronic complex protocols during the same period with two stars.

II.C OPRA data

Member exchanges are responsible for submitting trades and their best bids and offers (BBOs) in real time to the SIP, which then combines the submissions into a consolidated OPRA feed disseminated in real time to subscribing exchanges and market participants. Our data come from the OPRA feed via MDR, a subsidiary of the CBOE. The data include the information about intraday trades and exchange-level quotes for U.S.-listed options contracts. The quotes include the prices and depth from each exchange's S-LOB for each option series, but do not include information from the C-LOBs which are not included in the OPRA feed. Instead, the OPRA data report each leg of a complex trade separately, similar to single-leg trades, but with trade condition codes that flag the legs of complex orders.

[5] PEARL was launched in February 2017, and EMERALD was launched in March 2019.

Table 3, Panel A shows the codes and the types of orders they are supposed to flag. According to the OPRA plan, "*L*" indicates a leg of a "Spread" order, "*M*" indicates a leg of a "Straddle" order, "*Q*" indicates part of a "Combo" order and "*P*" indicates the option leg of a buy-write trade. Table B1 in Appendix B illustrates examples of 32 *M*, *L*, and *Q* trades executed over the 31 seconds from 11:01:24 AM to 11:01:55 AM on October 18, 2016, for the SPY option class.

It appears that the exchanges' actual use of these codes does not always follow the plan. Panel B of Table 3 summarizes multi-leg trades in December 2016 and shows that while all seven exchanges report "*L*" trades, only ISE, PHLX, and MIAX report "*M*" trades, only CBOE, ISE, MIAX, and PHLX report "*Q*" trades, and only ISE and PHLX report "*P*" trades. This suggests that some options exchanges, such as EDGX, ARCA, and C2, may have labeled all complex trades, regardless of whether they are legs of a straddle, a vertical, or a stock-option trade, as "*L*."

As of November 4, 2019, the OPRA plan (OPRA (2019)) deployed revised codes, collected in Panel C of Table 3, which allow subscribers to identify whether trades were executed on the limit order book, through an auction (crossing) mechanism, or on an open-outcry trading floor.[6] The revision still allows subscribers to indicate whether trades are complex, they no longer differentiate between straddle/strangle trades and combo trades. However, the algorithm based on the original codes can still be applied with minor modifications.

In Table 3, Panel D we report the volume summary and market share across the exchanges that provided complex functionalities in July 2021 for all complex trades, further divided by condition code.

III. Data

We use both the intraday OPRA data and end-of-day Optionmetrics data for listed options for the period running from January 4, 2016 through December 31, 2018. As discussed in Section II.C and Appendix B, we can identify complex trades and infer the associated strategies from the OPRA data. We acquire the OptionMetrics data through Wharton Research Data

---

[6] Please refer to https://assets.website-files.com/5ba40927ac854d8c97bc92d7/610ab429b74c1b692dd2f071_OPRA%20Pillar%20Output%20Specification.pdf for more details.

Services (WRDS). We derive the daily change in open interest for each options series in our sample from the OptionMetrics data.

We limit our analysis to the most liquid stocks and ETPs that are listed for the entire 30-month sample period. Our original intent was to study complex volume in the 30 most liquid stock options and 30 most liquid ETP options based on daily volume, but we then adjusted the sample to reflect the division of trading in Alphabet across its two listed common stock share classes GOOG and GOOGL. Options on neither share class make the top 30 separately, but together they rank 20th. Therefore, the final sample consists of the 31 stock options classes and 30 ETP options classes listed in Table 4.

Our analysis of complex versus simple volume around stock splits uses the entire sample from 2011 to 2021 to assemble the sample of stocks and ETPs experiencing splits with split ratios greater than one. This excludes reverse splits that increase stock prices.

IV. Overview of Complex Options Trading Volume

Table 4 reports the volumes of simple and complex option trades for the options classes in our sample, the stock classes in Panel A and the ETP classes in Panel B. For each class the table shows the call and put volumes that occurred in simple trades, complex option-only trades, and complex option-plus-underlying trades (e.g., buy-write trades). The complex ratio in the last column divides the volume occurring in complex trades by total volume, and shows that complex trades are much more likely among ETPs than stocks, accounting for 36% of ETP volume as opposed to 24% of stock volume, and vary significantly across classes, ranging from 16% for TWTR to 37% for AMZN among stocks, and from 14% for SLV to 65% for HYG among ETPs.

The table reveals a disparate use of puts and calls between stock and ETP options. Options on stocks tend strongly toward calls. Simple call volume exceeds simple put volume for every stock class, and complex call volume exceeds complex put volume for most stock classes. In contrast, options on ETPs tend strongly toward puts, with simple and complex put volumes exceeding call volume for many ETP classes, especially for IWM, HYG, SPY, and QQQ. We do not try to explain this disparity, but it is at least consistent with investors using stock options more for levered upside bets and index options more to lay off downside risk.

In Table 5 we break out the complex option trades into the different types (see Table 1 for definitions of all types), with results for stocks displayed in Panel A and those for ETPs in displayed in Panel B. Simple vertical call and put spreads are the most popular trades, comprising 36% of stock options volume and 34% of ETP options volume. Calendar and diagonal spreads account for another 22% and 12% of stock and ETP options volume. Thus, these simple spread trades account for 58% (= 36% + 22%) and 46% (= 34% + 12%) of stock and ETP complex options volume, respectively.

Vertical ratios comprise another 6% and 9% of stock and ETP options volume, respectively, while vertical rolls comprise another 2% of volume in both stock and ETP options. Thus, the various vertical trades account for 44% (= 36% + 6% + 2%) and 45% (= 34% + 9% + 2%) of stock and ETP complex options volume, respectively. All spread trades, including the vertical ratios and vertical rolls, account for 66% (= 36% + 22% + 6% + 2%) and 57% (= 34% + 12% + 9% + 2%) of complex volume.

Straddles, strangles, butterflies, iron butterflies and iron condors together account for only 6% of both stock and ETP options volume, while rolls of straddles and strangles account for another 2.7% and 1.5% of stock and ETP options volumes. Thus, these volatility-associated trades are small components of complex volume.[7] In addition, trades that might be straddles but might instead be combinations account for 9% and 10% of stock and ETP complex volume, respectively.

IV.A Determinants of Daily Complex Ratios

Why do seemingly similar stocks and ETPs have different proportions of complex trading? One possibility is that recent returns, which can differ significantly between similar stocks and ETPs, drive complex trading through their effect on moneyness. We explore this possibility using pooled cross-sectional time-series regressions that explain the daily complex ratio of a class using the past week's return on the underlying, which we separate into the positive part that increases the moneyness of calls and the negative part that increases the moneyness of puts. To account for other time-varying motives to trade complex options, the explanatory variables also include implied volatility and the VIX, which may motivate volatility trades, and the absolute value of the three-month ATM volatility minus the one-month ATM volatility, i.e., the slope of the volatility term structure, which also may motivate volatility trades. Additional variables include turnover and the logs of price and market capitalization. To capture expiration effects, we include indicator variables for the expiration Friday and the expiration

[7] Below we show that many of the butterflies and iron butterflies are executed away from the money, which suggests that many of these are directional trades.

week (variable definitions are in Table A1 of Appendix A). We estimate the regressions with stocks and ETPs both separately and together, and with and without options class and date fixed effects. The VIX and the expiration Friday and expiration week indicator variables are omitted from the specifications with date fixed effects, and the regression estimated using both stocks and ETPs without options class fixed effects includes an indicator variable for ETPs. Table 6 displays the results.

The regressions find that recent returns, which change moneyness, help explain the complex ratio, especially in the absence of date and class fixed effects. The log of price, the volatility term structure, and the expiration date and week indicator variables also enter significantly in the specifications without the fixed effects, whereas the specifications with fixed effects find weaker relations. The stronger result for log(price) without the fixed effects is consistent with prices affecting complex trading through investors' budget constraints, and the elevated volume on expiration days and weeks is consistent with the use of complex trades to roll expiring positions.

IV.B Analysis of Complex Options Trading Before and After Stock Splits

The higher complex ratios for higher-priced underlying stocks suggest an effect of price on the investor's budget constraint through the 100-share option contract size, but it might also reflect other possible endogeneities between firm values and option trading. For cleaner identification of whether stock prices and not firm values drive complex trading, we use the abrupt changes in stock prices but not firm values caused by stock splits. The sample consists of all stock or ETP splits (excluding reverse splits) between July 2011 and December 2021, for which the average daily options volume during the 30 trading days prior to the split was at least 500 contracts. After applying this filter, there are 53 splits in the sample, of which 15 are splits of ETPs. For each split we identify a matching stock or ETP that did not experience a split. The matching stock or ETP must have the same CRSP shrcd and 2-digit SIC code, and must meet the same requirement for average daily options volume before the split. Of the symbols that meet these criteria, we select the one closest in option volume and market capitalization.[8] We track option trading from 30 trading days before to 30 trading days after the split.

[8] Specifically, we pick the symbol i that minimizes $(rcap_i - rcap_s)^2 + (rvol_i - rvol_s)^2$ where $rcap_i$ and $rvol_i$ are the ranked market capitalization and option volume, respectively, for symbol i, and $rcap_s$ and $rvol_s$ are the same for the splitting stock.

The empirical question is whether complex volume drops with the price across the split. Accordingly, we estimate difference-in-difference regression models explaining the complex ratio, and test whether the complex ratio drops more across the split for the treated than control symbols. Specifically, we include an indicator $Treat$ for the split stocks and an indicator $After$ for the days after the split, and test is whether the interaction term $Treat \times After$ enters negatively. We also allow for a larger effect of larger splits by interacting $Treat \times After$ with $\log(Sfactor)$, where $Sfactor$ is the split factor (e.g. 3 for a 3-for-1 split), in which case this triple interaction rather than the double interaction picks up the whole effect. The control variables are turnover (*TO*), implied volatility (*ImplVol*), realized volatility (*Volatility*), the log of the market capitalization of the underlying stock (log(*MCap*)), and option class and date fixed effects. We report the results are in Table 7.

We find that complex trading drops significantly with stock splits. The double interaction enters significantly in column (1) and the triple interaction enters significantly in column (2), and remains significant when the controls are included in column (3). Figure 2 presents this result graphically by tracking the average complex ratio of treated minus control securities across the split dates. We conclude from this evidence that the relation between price and complex trading volume is specifically a price effect, and is not due to an endogeneity with firm value.

V. Spread Trades

The results discussed in the previous section show that the complex ratio varies substantially across options classes, and complex volume concentrates in spread trades. This section focuses on the spread trades. Vertical spreads can adjust the strikes of simple positions by closing out one simple position and opening another, calendar spreads can adjust expirations of simple positions, and diagonal spreads can be used to adjust both strikes and expirations. Vertical ratio spreads can be used to adjust both the strikes and sizes of simple positions. When the near leg of a calendar or diagonal spread is ITM, or when one of the legs of a vertical or vertical ratio is ITM, such trades can avoid the exercise or assignment of the options position being closed out. This section explores the hypothesis that a large fraction of spread trades are executed to adjust simple options positions, with an eye toward these motives.

We begin by categorizing spread trades by the moneyness of their legs, which sheds light on the purposes of the trades. We next provide some illustrative graphical evidence on volumes in complex spread trades in options on a single stock, Apple, Inc. (AAPL). We then test for the role of spread trades of adjusting existing options positions. Specifically, we test whether trades in call spreads follow increases in the underlying price and trades in put spreads follow decreases. Finally, we test whether the spread trades open and close positions by regressing the daily changes in open interest of options series on the trade volumes of the legs of the spread trades that are in the series.

V.A Moneyness of Spread Legs

Investors with ITM options can delay exercise with calendar or diagonal spreads, so a preponderance of near legs ITM among calendars, and near legs ITM and far legs OTM among diagonals, is consistent with the spreads adjusting simple positions. The investors can also avoid execution through verticals and vertical ratios with one leg ITM and one OTM. Investors with options that are OTM but uncomfortably near the money can push strikes out with vertical and vertical ratios that have both legs OTM. So the hypothesis that investors use the spreads to adjust simple positions predicts relatively more spreads with these patterns among their component legs. The patterns we would expect less of, by this hypothesis, include verticals and vertical ratios with both legs ITM, and diagonals with the near leg OTM and the far leg ITM, as these would not help avoid exercise.

We categorize the trades in the original sample used in Tables 4-6 and report the results in Table 8, with Panel A displaying results for stock options and Panel B displaying results for ETP options. The column headings are of the form ITM-ITM, ITM-OTM, etc. For calendars and diagonals this is near leg-far leg and for verticals and vertical ratios it is first leg-second leg, where “first leg” (“second leg”) means the lower (higher) strike for call spreads and the higher (lower) strike for put spreads.

The results bear out the prediction that verticals and vertical ratios rarely have both legs in the money. One leg ITM and one OTM is common, and both legs OTM is most common. Both legs OTM is consistent with pushing out strikes but is also consistent with buying the spread to save money, so the motive is less clear. Both legs ITM is common among calendars and diagonals, and as predicted, OTM-ITM diagonals are rare. So the distribution of spreads across these moneyness buckets supports the hypothesis that spreads are commonly used to avoid execution of simple positions.

V.B Complex Spread Trades in a Single Stock

For some additional insight we follow complex trading in a single stock, AAPL, for a single year, 2018, and focus on the spread trades for which at least one leg is ITM. Figure 3 presents the daily AAPL stock price and the daily complex volume of diagonal, calendar, vertical, and vertical ratio spreads on AAPL with at least one leg in the money. The results for diagonal put spreads in Panel B reveal large volume spikes while AAPL declined from 230 to 150 in the last quarter, and the volume spikes before July also tend to coincide with price drops, consistent with using the spreads to keep the put strikes below the stock price. Volume is instead near zero in July and August and other times when the stock ran *up*. So the patterns in how investors trade these put diagonals with a leg in the money is consistent with investors avoiding execution.

Panel A makes the same point about diagonal call spreads, whose volume was near zero during the fourth-quarter drop in AAPL, and in June and July when AAPL traded in a fairly narrow range. In contrast the volume in diagonal call spreads spikes when AAPL runs up. Similarly, the vertical and vertical ratio spreads in Panels E through H show peaks for call spreads when AAPL rises and troughs otherwise, and the opposite for put spreads, again consistent with investors using the spreads to move strikes out of danger.

The same pattern plays out with the calendar spreads in Panels C and D, which show low call volumes in November and December following the decline in AAPL, and low put volumes in May-June and August following sharp runups in AAPL. The relation is not as tight as with the other spreads, which is consistent with the calendar spreads not moving strikes, and thus not being as effective in staving off exercise.

This illustrative year of trading is consistent the hypothesis that investors use spreads to move strikes away from the price of the stock or ETP. The hypothesis that investors want diagonal spreads because they desire the exposure provided by the spreads themselves does not predict its volume evaporating as it does when stock returns do not raise the threat of exercise.

V.C Relation between volumes of complex spread trades and recent past returns

We generalize from the AAPL example to the full sample to test the hypothesis that changes in moneyness explain the volumes of spread trades. The test for the effect of increasing moneyness is a set of regressions that explain the relevant daily volume for a class, for example, the daily volume of all vertical, calendar, and diagonal spreads on AAPL with at least one leg in the money, using the ten most recent daily returns on the underlying as explanatory variables.

The hypothesis is that volume increases with moneyness, so call-spread volume increases and put-spread volume decreases with recent returns. The regressions separate calls from puts and stocks from ETPs, and include option class and day fixed effects. Standard errors are clustered at the option class and day level. We report the results in Table 9.

The regressions show strong effects in the predicted directions. Call spread volume is positively and put spread volume is negatively related to recent returns. This applies at a smaller scale to calendar spreads that only defer expiration. These results are consistent with many market participants actively maintaining their exposures by moving strikes and/or expirations.

V.D Expiration and Spread Legs

If calendars and diagonals serve to extend expirations on simple positions, it stands to reason that this motive is stronger when expiration is more imminent. To assess this motive, we first sort these spread trades by day of the week, and then for each day and series we calculate the fraction of the traded spreads for which the near leg was expiring on the Friday of that week, and then average across series. Results are reported in Figure 4 whose eight panels sort the trades into calendars vs. diagonals, puts vs. calls, and stocks vs. ETPs. The graphs show a general increase over the week, from about 35 percent on Monday, i.e., 35 percent of calendar and diagonal trades on Mondays have the near leg expiring that Friday, to 50 percent on Friday. This increase over the week is consistent with the hypothesis that many of these spreads are moving out expirations.

V.E Explaining Changes in Open Interest

Do the spread trades open and close existing options positions as these patterns suggest? We cannot see directly in our data whether a trade opens or closes a position, but we can see the change in the open interest of the affected class across the whole day, so we learn what we can from that. We do this with regressions in which the dependent variable is a day's change in the open interest of an option series, and the explanatory variables are the trading volume in that option series disaggregated into its components: how much occurred in simple trades, how much in vertical spreads, calendars, diagonals, straddles, etc. We separate calls from puts and stocks from ETPs and include fixed effects for option classes and weeks to expiration. Standard errors are clustered at the level of option class and weeks to expiration. We report the results in Tables 10, 11, and 12. All variables described in Table A1, Panel C.

Table 10 presents the baseline regression with one row for each type of complex trade. The unit of observation is a combination of option series and date, for example, AAPL calls with strike K and expiration T on date t. The explanatory variables are the same-day trading volume in series j (e.g., AAPL calls with strike K and expiration T) due to each type of complex trade. For example, consider all vertical spread trades on date t for which one of the legs is in series j. If the volume in series j across the vertical spreads that involve series j is x contracts, then for series j the covariate *vv* ("vertical volume") equals x. The other explanatory variables are the volumes due to vertical ratios ("*vr*"), calendars ("*cv*"), diagonals ("*dv*"), straddles ("*straddle*"), strangles ("*strangle*"), butterflies ("*bf*"), iron condors ("*ica*"), vertical rolls ("*vertical_roll*"), straddle rolls ("*straddle_roll*"), buy-write trades ("*bw*"), regular single-leg trades ("*regular*"), and other types of complex trades ("*other*"). The results show that almost all trade types enter positively, that is, the volumes are associated more with increases in open interest, consistent with many options positions terminating through expiration or exercise rather than a closing trade. Many of the coefficients are close to 0.4 or greater, and highly significant. The coefficients on straddles and strangles in the stock option regressions are especially large, close to 0.8.

Interpretation of these coefficients needs to consider that a customer trading to open a position increases open interest only if the trade also increases the magnitude of the counterparty's position. For example, suppose a customer with no position in series j buys one contract via a complex trade, and the counterparty is a market maker who sells one contract to the customer. If the counterparty (market maker) was previously long series j, this trade would not change open interest because it does not change the total number of long positions. That is, it increases the customer's long position by one contract and reduces the market maker's long position by one contract. On the other hand, if the market maker was previously short series j, this trade would increase open interest by one contract. So if complex trades always opened positions and counterparties were equally likely to be long or short, we would expect the regression coefficient to be 0.5. Conversely, if complex trades always closed positions and the counterparty were equally likely to be long or short, we would expect the coefficient to be −0.5.

Unfortunately, we do not know the probability that the counterparty to a complex trade in series j is long or short series j. Nonetheless, the discussion above suggests that coefficients close to or greater than 0.5 should be interpreted as evidence that a large fraction of trades are opening trades. Similarly, coefficients close to or below −0.5 indicate a large fraction of closing trades. For example, if the counterparty were equally likely to be long or short, the coefficients of about 0.4 would imply that 80% of trades are opening trades. The coefficients of about 0.8 are consistent with straddles and strangles always opening positions, with the counterparty having the opposite position 80% of the time.

To test whether legs of spread trades tend to open or close positions, we expand the specification in Table 10 by breaking out the spread trades by whether certain legs are ITM. We divide the calendar and diagonal spreads by whether the near leg is ITM, and then break out the volumes for the near and far legs, i.e., $cv = cniv + cnov + cfiv + cfov$, where *cniv* and *cfiv* are the volumes occurring in the near and far legs of calendar spreads that have the near leg ITM, and *cnov* and *cfov* are the volumes occurring in the near and far legs of calendar spreads have the near leg OTM. For diagonals, $dv = dniv + dnov + dfiv + dfov$, where *dniv* and *dfiv* are the volumes occurring in the near and far legs that have the near leg ITM, and *dnov* and *dfov* are the volumes occurring in the near and far legs that have the near leg OTM. We apply similar logic to categorize the volumes of vertical ratio spreads, with the additional consideration of whether the combination of the first leg (lower strike for calls and higher strike for puts) and second leg is ITM-ITM or ITM-OTM. Therefore, the explanatory variables of interest are *vrfiov* and *vrsiov*, where *vrfiov* and *vrsiov* are the volumes occurring in the first and second legs of vertical ratio spreads that have the first leg ITM and the second leg OTM. For vertical spread trades, we go one step further by breaking out the trades that have both legs OTM into buyer-initiated volume and seller-initiated volume. Therefore, the variables of interest are *vfov_s, vsov_s, vfov_b* and *vsov_b*, where *vfov_s* and *vsov_s* are the volume occurring in the first leg and second leg of the vertical trades where both legs are OTM and the whole package is seller-initiated, and *vfov_b* and *vsov_b* are the volume occurring in the first leg and second leg of the vertical trades where both legs are OTM and the whole package is buyer-initiated. The definitions of all variables are collected in Table A1, Panel C. We report the regression results in Table 11.

The results in Table 11 are clearest for diagonal and calendar spreads, and show that these trades tend to close the near leg and open the far leg. The coefficient on *dniv*, the volume occurring in the near legs of diagonal spreads with the near leg ITM, is strongly negative in three of the four specifications, and *dnov* is small in all specifications. In contrast, the coefficients on *dfiv* and *dfov* are positive and large, and we find similar results for calendar spreads.

The results for vertical ratio spreads are less strong. Although *vrsiov*, the volume in the second legs of vertical ratio trades where the first leg is ITM and the second leg is OTM, enters positively and significantly in all four specifications, *vrfiov*, the volume of the first legs of such trades, enters negatively and significantly in two out of four specifications. In the other two specifications, *vrfiov* enters negatively but not significantly. The results for vertical spreads are similar. In two out of four specifications, we find that for trades with one leg ITM and the other leg OTM, the rolling volume dominates volume from other motives. It is also worth pointing out that *vfov_s* enters positively and *vsov_s* enters negatively for vertical call trades. This suggests that the majority of seller-initiated vertical call trades, where both legs are OTM, are used to roll down existing short call options (e.g., covered call).

The results displayed in Figure 4 show that the use of calendars and diagonals to move expiration out grows as expiration approaches. We use the regression models to further explore that finding by breaking out the calendar and diagonal trades by whether they fall in the expiration week. The results are in Table 12 where, for example, *cnev* is the volume occurring in the near legs of calendar spreads expiring that week and *cnnv* is the volume occurring in the near legs of calendar spreads *not* expiring that week. The hypothesis is that the calendar and diagonal spread trades during expiration weeks are more likely to be closing their near legs. This is what we see: the coefficients are generally much more negative for the near legs in expiration weeks. The coefficient on volume in the far legs are generally quite positive. Thus, the calendar and diagonal spread trades that can move expirations appear to be serving this purpose.

The specifications in Table 12 also test whether four-leg trades, specifically trades we identify from their legs as rolls of verticals, straddles and strangles, are indeed used to roll out existing complex positions. A vertical roll trade is a four-leg trade consisting of two verticals with different expirations, where one vertical is purchased and the other is sold. Straddle and strangle rolls are analogous. As rolling is the obvious and widely accepted motive for these trades (and the source of the name), this regression is to some extent testing whether our methodology succeeds in detecting rolls. We break out the vertical roll and straddle/strangle roll trades into two near legs and two far legs so that *vertical_roll_n* is the volume occurring in the near legs of the vertical roll spreads, and we find that it does indeed enter negatively in all four specifications and significantly in three of the four. The variables *vertical_roll_f*, which is the volume occurring in the far legs, enters positively and also significantly. The pattern also holds for the straddle and strangle rolls. Thus, as expected, the regression results identify the four-leg trades that appear from their construction to be roll trades as roll trades.

VI. Volatility Trades

This section focuses on the trades that are often considered volatility trades. We begin by presenting some statistics characterizing the long and short straddle, strangle, butterfly, iron butterfly, and iron condor trades. Specifically, we examine three measures: (a) *MK/S*, a measure of the symmetry of the trade relative to the underlying, defined as the average of the strikes divided by the underlying price; (b) *DK/S*, which is only applicable to strangles and iron condors, is defined as the absolute difference between the two inner strikes divided by the underlying price, i.e., it is the width at the inner strikes relative to the underlying; and (c) *DK2/S*, which is applicable to the three- and four-leg trades, defined as the absolute difference between the outer strike on either side and the next inner strike, divided by the underlying price. We report the results in Table 13.

According to conventional wisdom (e.g., Hull (2009)), market participants use straddles, strangles, butterflies, and condors to obtain exposure to changes in volatility with little or no exposure to changes in the underlying price. If this conventional wisdom is correct, then *MK*/*S* should typically be close to one, because if it is then the trade is centered or approximately centered on the current underlying price, and provides exposure to changes in volatility but not to changes in the underlying price. However, the results we report in Table 13 show that the some of these trade types are not typically centered around the underlying price. While the average value of *MK/S* is close to one for strangles, iron butterflies, and iron condors, this is not the case for the other trade types. The strikes of straddles tend to be above the current underlying price, with the average values of *MK*/*S* being 1.045 (1.037) and 1.019 (1.096) for sold and bought straddles, respectively, on stocks (ETPs). Butterfly call and put spreads tend to be centered on prices above and below the current stock price, respectively. For example, the results in Panel A show that for sold call butterfly spreads on stocks the average value of *MK*/*S* is 1.076, that is, the middle strike is on average 7.6% above the current stock price. The average value of *MK*/*S* is even larger, 1.10, for sold call butterfly spreads on ETPs. These results indicate that there is an important directional component to many straddles and butterflies, and thus that the conventional wisdom that views them as pure volatility bets is not correct.

The width *DK/S* is relevant only for strangles and iron condors. It is similar for both trade types on both stocks and ETPs, ranging only from 10.9% to 12.3% of the underlying. The averages of the width *DK2/S* display more variability, ranging from 4.4% for sold iron condors on stocks to 8.5% for bought iron butterflies on ETPs.

To gauge which trades tend to open or close positions, we adapt the baseline open-interest regression by breaking out the volatility trades by whether we estimate them to be buyer- or seller-originated, using the quote rule. Specifically, the unit of observation is a combination of option series and date, and the explanatory variables for series *j* and date *t* are the date *t* volumes in that series due to trades of a specified type. For example, for series *j* and date *t*, *bf_buy_volume* (long volatility) is the volume in that series on that date from the legs of butterfly spreads for trades classified as buys. Similarly, *bf_sell_volume* is olume in that series on that date from the legs of butterfly spreads for trades classified as sells. The variables *straddle_buy_volume*, *straddle_sell_volume*, *strangle_buy_volume*, *strangle_sell_volume*, *ica_buy_volume*, and *ica_sell_volume*, where "ica" standards for iron condors and iron butterflies, as we combine the iron butterfly and iron condor volumes into the same set of variables. Table A1 includes the definitions of both these and the other variables included in the regression models.

We report the results in Table 14. They show that *bf_buy_volume* enters negatively, but with scant statistical significance. By contrast, the analogous measure *bf_sell_volume* for sales of butterflies enters positively in all four specifications, and significantly in three of the four. The same pattern holds for the combined iron condor and iron butterfly buy volumes (*ica_buy_volume*) and sell volumes (*ica_sell_volume*), consistent with market participants using butterflies and iron condors to short volatility and using the buy trades to close out existing short volatility positions rather than to open new positions to buy volatility.

We do not find the same patterns among straddles and strangles. Instead, we find that straddle sales appear to open positions less often than straddle buys, in that the coefficient on *straddle_sell_volume* is less than that on *straddle_buy_volume* three times out of four, and strangle sales open positions *more* than strangle buys do, in that the coefficient of *strangle_sell_volume* is always greater than that of *strangle_buy_volume*. These results are consistent with traders using straddles to go long volatility and strangles to go short.

VI. Summary and Conclusion

Complex trades command a large fraction of options volume but little has been documented about the roles they play. We explore these roles with a new database of complex orders, assembled from the OPRA feed using the time stamps and condition codes of individual legs. We find that spread trades dominate the complex market, and that an important use of spread trades is to adjust existing simple positions, rather than to acquire the spread's payoff. Spreads in particular help investors avoid the frictions of execution and assignment, as is apparent from the use of call spreads after the underlying goes up and put spreads after it goes down and the use of calendar, diagonal and vertical ratio spreads that close the leg expiring sooner and open the leg expiring later. Spreads also help option investors economize when high-priced underlying makes the 100-share minimum expensive, as is apparent in the effect of stock splits that relax this constraint.

Options textbooks, training materials and commentary tend to focus on the payoffs of complex options, rather than the other problems they can solve. We show that the payoffs are often incidental to complex volume, which instead reflects the desire to move strikes and expirations at low transactions costs. So where it can be a strain to explain the appeal of a diagonal spread's payoff viewed in a vacuum, it is easy to understand the diagonal as a means to an end in its actual context. The complex trades reduce the net cost of carrying out simple-option strategies and this helps explain why exchanges have been so eager to facilitate complex trading with limit-order books and price improvement auctions. What might from a distance appear to be increasing complexity is in fact an increasingly efficient solution to a simple need.

References

Battalio, Robert, Hatch, B. and Jennings, R., 2004. Toward a national market system for US exchange–listed equity options. *The Journal of Finance* 59, No. 2, pp. 933-962.

Hu, Jianfeng, 2014. Does option trading convey stock price information? *Journal of Financial Economics* 111, No. 3, pp. 625-645.

Hull, J., 2009. Options, Futures and Other Derivatives. Upper Saddle River, NJ: Prentice Hall.

Lakonishok, Josef, Inmoo Lee, Neil D. Pearson, and Allen M. Poteshman. 2007. Option market activity. *Review of Financial Studies* 20, No. 3, pp. 813-857.

Lee, C.M. and Ready, M.J., 1991. Inferring trade direction from intraday data. *Journal of Finance*, 46, No. 2, pp.733-746.

Li, Su, David K. Musto, and Neil D. Pearson, 2022. Costs of executing complex options trades. DERA Working paper.

Li, Su, David K. Musto, Neil D. Pearson, and Li Wang. 2023. Reversals, complex trades, and earning announcements. Working paper.

OPRA, New OPRA Trade Message Codes, 2019.

Securities and Exchange Commission, Release No. 34-43086, Order approving option market intermarket linkage plan submitted by the American Stock Exchange LLC, the Chicago Board Options Exchange, Inc., and the International Stock Exchange, LLC, 2000.

Securities and Exchange Commission, Release No. 34-43591, Firm quote and trade-through disclosure rules for options, 2000.

Figure 1
Monthly Simple (Single-Leg) Options Volume and Complex (Multi-Leg) Options Volume over the Period between 2011/07 and 2021/12.

The figure shows the monthly volume of single-leg options (simple volume) and complex options (complex volume) traded between July 2011 and December 2021. The vertical axis to the left represents the volume of options traded, while the horizontal axis represents the time period from July 2011 to December 2022. The dashed line represents the complex ratio which is defined as the complex volume divided by the total volume and correspond to the vertical axis to the right.

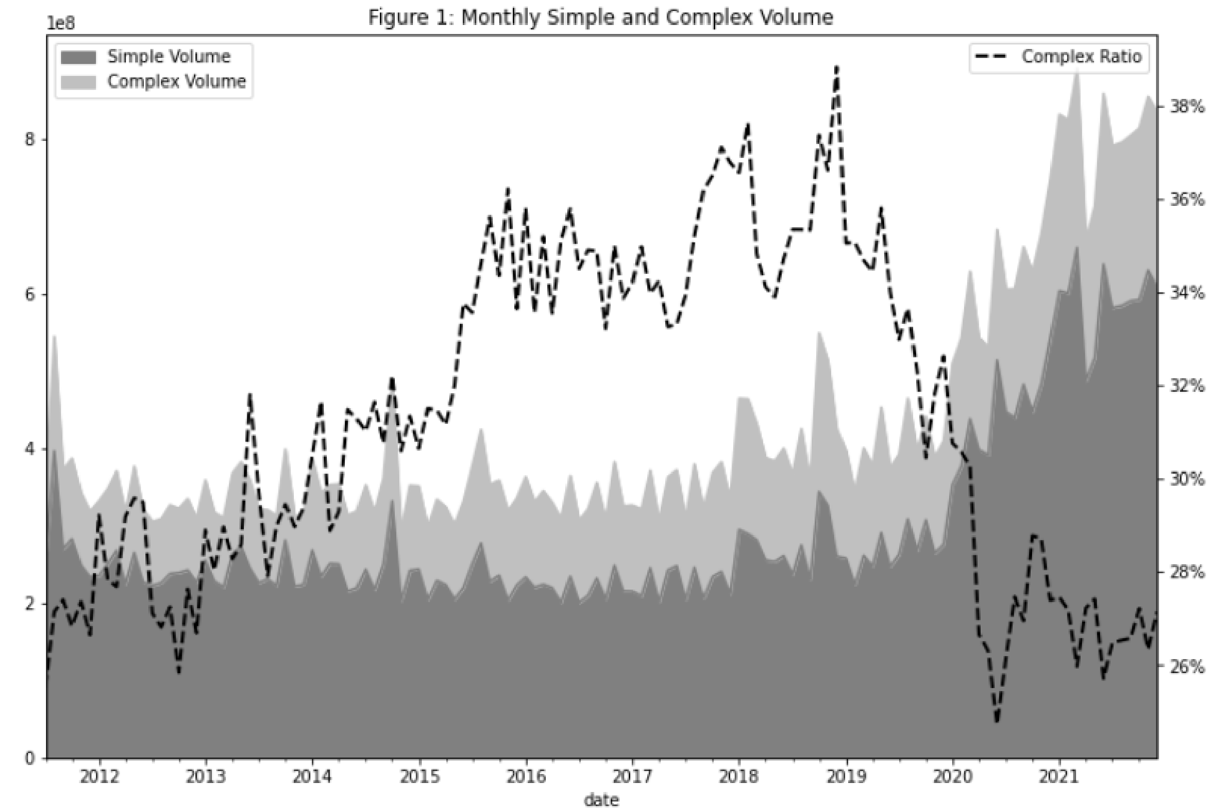


Figure 2
Complex ratio and scaled volume of the split and control stocks and ETPs.
Panels A and B, respectively, display the complex ratio and scaled volume of the split and control stocks and ETPs. The complex ratio is the ratio of complex volume to total options volume, and scaled volume is the ratio of option volume (in contracts) to stock trading volume (CRSP variable shrout, which is in thousands). The sample consists of the 53 split events between July 2011 and December 2021, for which the average daily options volume during the 30 trading days prior to the split was at least 500 contracts. For each split we identify a matching symbol that did not experience a split. The matching symbol must have the same CRSP shrcd and 2-digit SIC code. Of the symbols that match on these two criteria, we select the one that has the smallest Euclidean distance from the splitting symbol in the space spanned by option volume and underlying market capitalization, after both variables are first transformed to their ordinal ranks.

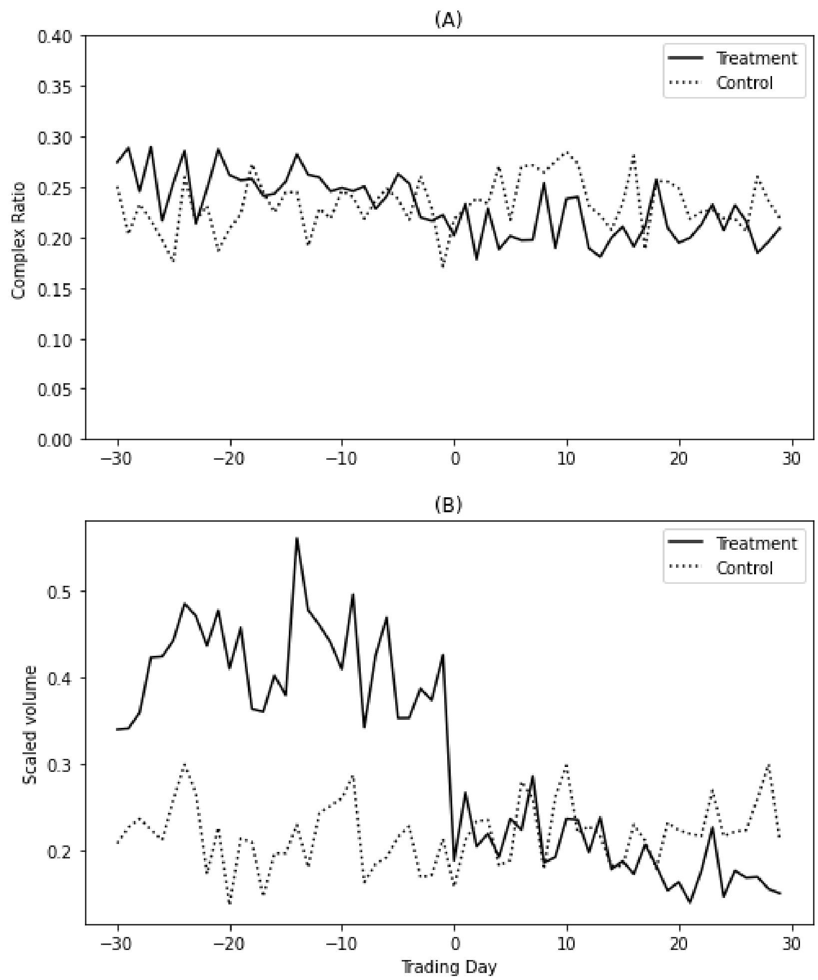
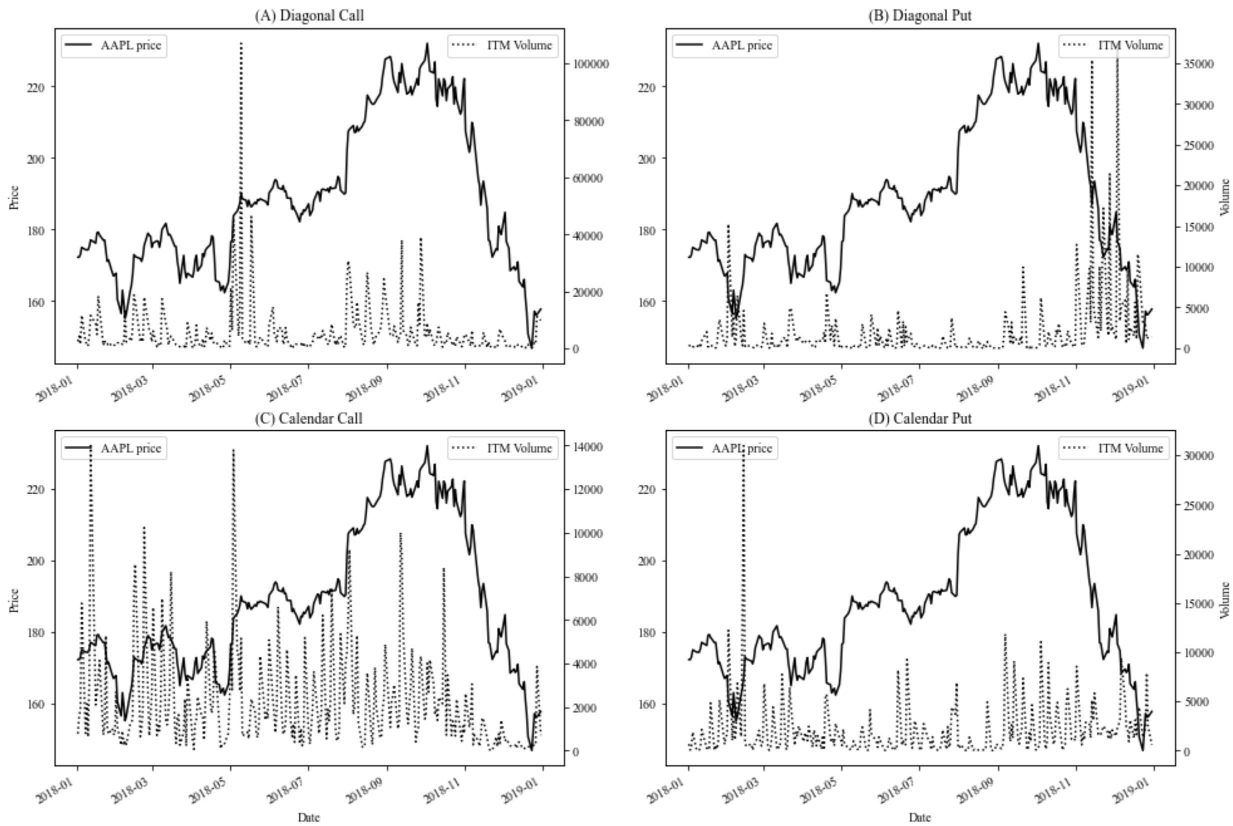


Figure 3
Diagonal, calendar and vertical trades with at least one leg ITM and stock price movement
The charts in this figure illustrate the relation between the ITM volumes of various multi-leg trades and the underlying price movement for AAPL stock during the period running from January through December 2018. For Vertical Call (Put) and Vertical Ratio Call (Put), the ITM refers to the trades where the lower (higher) strike is ITM. For Calendar and Diagonal, it refers to the trades where the near leg is ITM.



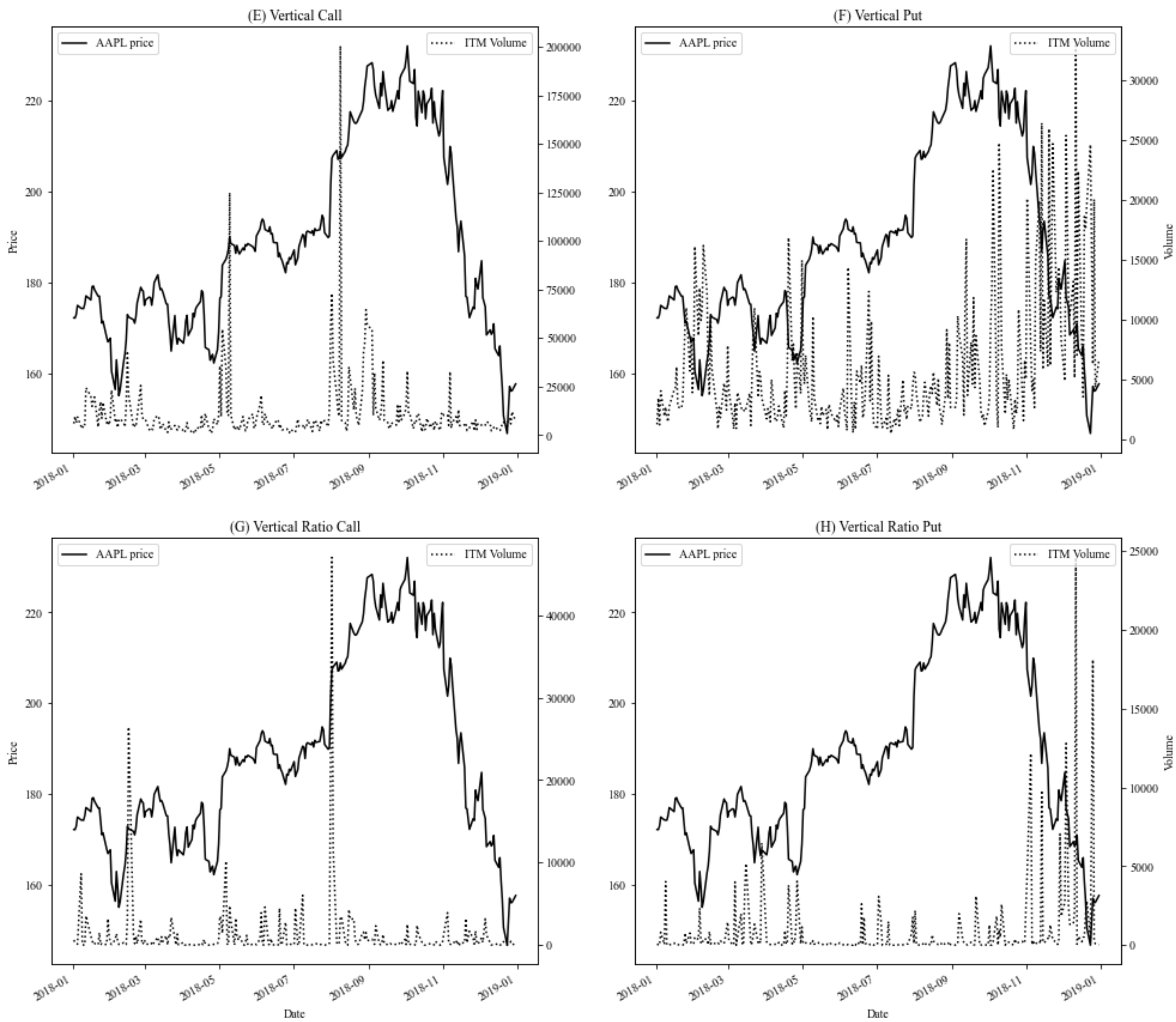
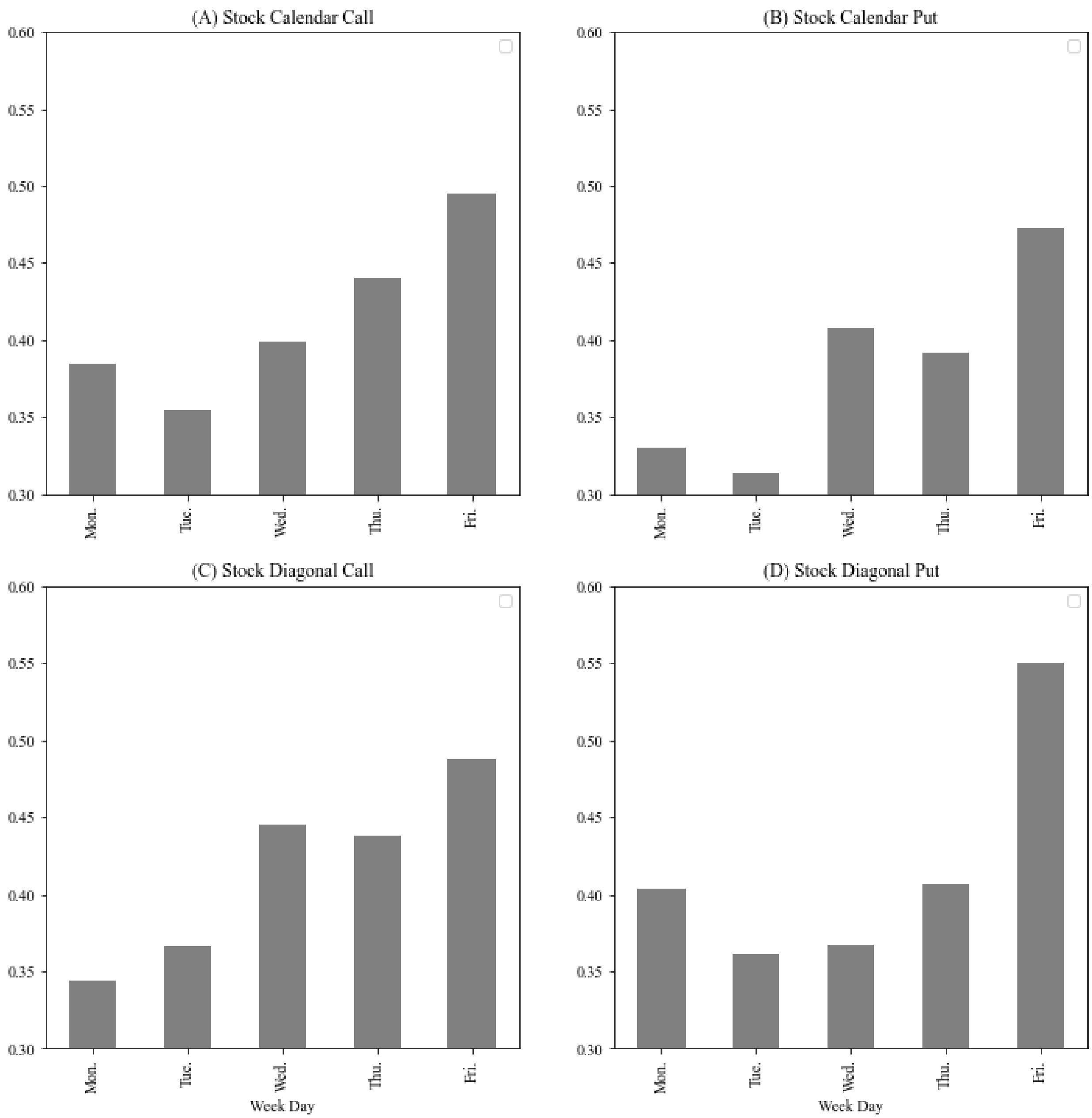

(E) Vertical Call
AAPL price
ITM Volume
Price
(F) Vertical Put
AAPL price
ITM Volume
Volume
(G) Vertical Ratio Call
AAPL price
ITM Volume
Price
Date
(H) Vertical Ratio Put
AAPL price
ITM Volume
Volume
Date

Figure 4
Calendar and diagonal spreads trades with near leg expiring on each weekday
The bar charts illustrate the fraction of the calendar and diagonal spread trades that have the near leg expiring at the end of the week on each weekday. Panels A through D are for the stock options, and Panels E through H are for the ETP options sample. We first derive the statistics for each options class, and then average across the classes.



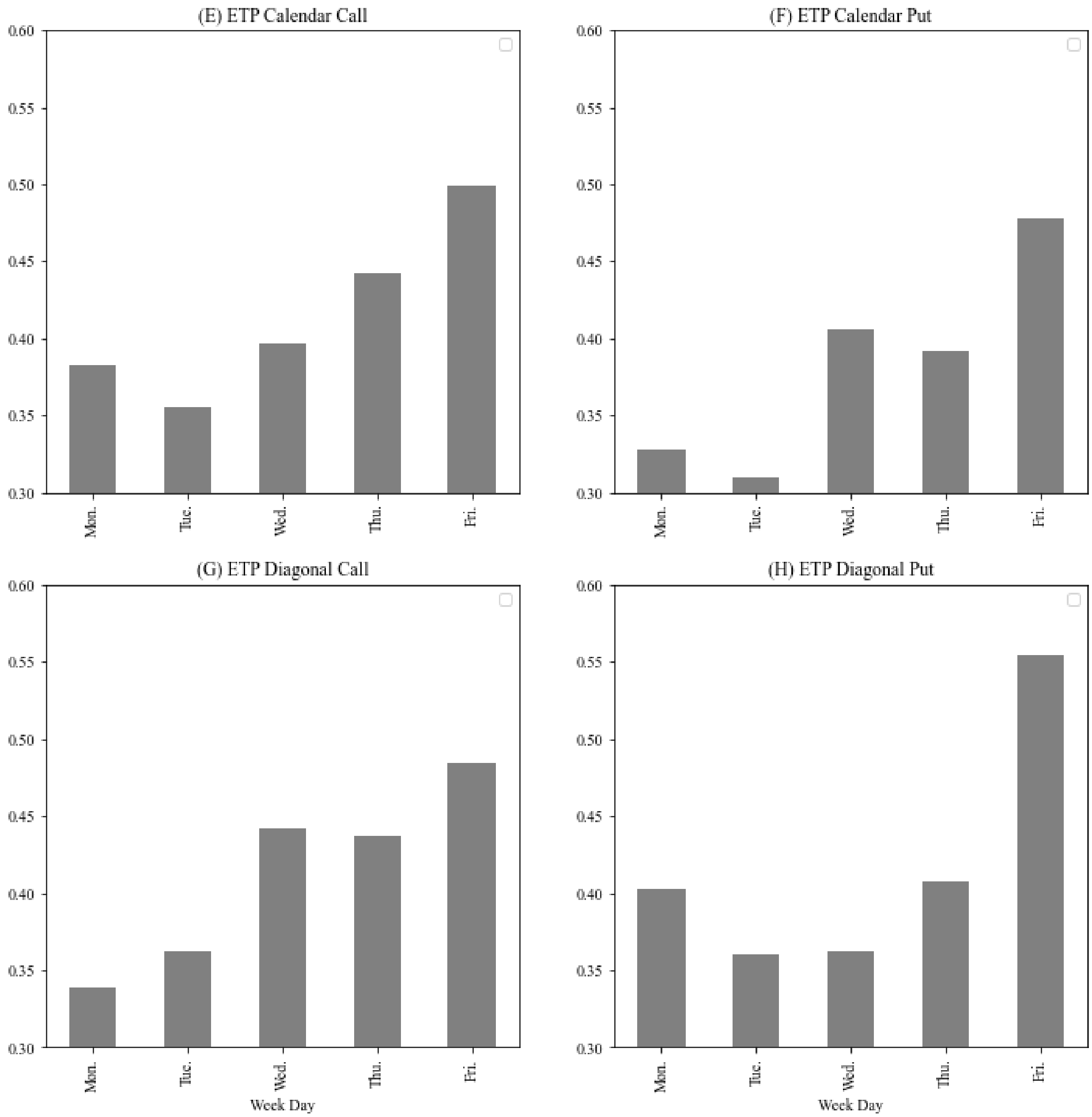

(E) ETP Calendar Call
(F) ETP Calendar Put
(G) ETP Diagonal Call
(H) ETP Diagonal Put
0.60
0.55
0.50
0.45
0.40
0.35
0.30
Mon.
Tue.
Wed.
Thu.
Fri.
Week Day

Table 1
Definition of the side of the complex trade
This table indicates how we define the side of trade for certain complex strategies.

Spread Type	Side of trade for each leg when buying the package
Vertical Call	Long the lower strike and sell the higher strike.
Vertical Put	Long the higher strike and sell the lower strike.
Vertical Call Ratio	Long the lower strike and sell the higher strike.
Vertical Put Ratio	Long the higher strike and sell the lower strike.
Calendar Call and Put	Long the near leg and sell the far leg.
Diagonal Call and Put	Long the near leg and sell the far leg.
Straddle	Long both the call and put legs.
Strangle	Long both the call and put legs.
Butterfly Call and Put	Long the middle strike and sell the two outer strikes.[9]
Iron Butterfly	Long the middle two strikes and sell the two outer strikes.
Iron Condor	Long the middle two strikes and sell the two outer strikes.

[9] This is essentially a credit spread. We define the direction of butterfly trade in this manner to align it with its volatility exposure.

Table 2
Complex Execution Protocols in Options Exchanges
This table provides an overview of the institutional background on the availability of complex order execution functionalities across different options exchanges. The exchanges marked with * are those that were launched between 2016 and 2020, while the exchanges marked with ** are those that added electronic complex execution protocols during the same period.

Group	Exchange	OPRA Code	Complex	C-LOB	C-Auction	C-Floor
NYSE	AMEX	A	Y	Y	Y	Y
	ARCA	N	Y	Y	N	Y
NASDAQ	NOM	Q	N	N	N	N
	BX	T	N	N	N	N
	PHLX	X	Y	Y	Y	Y
	ISE	I	Y	Y	Y	N
	GEMINI	H	N	N	N	N
	MRX **	J	Y	Y	Y	N
CBOE	CBOE	C	Y	Y	Y	Y
	C2	W	Y	Y	N	N
	BZX	Z	N	N	N	N
	EDGX **	E	Y	Y	Y	N
MIAX	MIAX **	M	Y	Y	Y	N
	PEARL *	P	N	N	N	N
	EMERALD *,**	D	Y	Y	N	N
BOX	BOX	B	Y	Y	Y	Y

Table 3
Historical and Recent Message Codes for Complex Options Trades from OPRA Plan and Summary Statistics

The table provides an overview of the historical and recent trade message codes used to tag complex trades in the options market according to the OPRA plan. The table is divided into four panels. Panel A describes the trade message codes used to tag complex trades before 11/04/2019.Panel B reports the volume summary and market share across the exchanges that provided complex functionalities in 2016/12 for all complex trades, further divided into each trade message code bucket. Panel C provides the description of the trade message codes used to tag complex trades after 11/04/2019. Panel D reports the volume summary and market share across the exchanges that provided complex functionalities in 2021/07 for all complex trades, further divided into each trade message code bucket.

Panel A: Message codes for the pre-2019 OPRA plan

Code	Value	Description
L	SPRD	Transaction represents a trade in two options in the same option class (a buy and sell in the same class).
M	STDL	Transaction represents a trade in two options in the same option class (a buy and sell in a put and a call).
Q	CMBO	Transaction represents the buying of a call and the selling of a put for the same underlying stock or index. This prefix appears solely for information; process as a regular transaction
P	BWRT	Transaction represents the option portion of an order involving a single option leg (buy or sell of a call or put) and stock. The prefix appears solely for information; process as a regular transaction.

Panel B: Frequency of pre-2019 OPRA plan message codes for trades reported on options exchanges

Exchange	Volume	Market Share	L	M	Q	P
AMEX (A)	18,894,098	11.80%	11.80%	0.00%	0.00%	0.00%
BOX (B)	1,789,512	1.12%	1.12%	0.00%	0.00%	0.00%
CBOE (C)	69,038,611	43.12%	39.23%	0.00%	3.88%	0.00%
EDGX (E)	2,894,525	1.81%	1.81%	0.00%	0.00%	0.00%
ISE (I)	23,637,998	14.76%	11.32%	0.73%	0.25%	2.47%
MIAX (M)	2,538,604	1.59%	1.56%	0.03%	0.00%	0.00%
ARCA (N)	13,387,675	8.36%	8.36%	0.00%	0.00%	0.00%
C2 (W)	1,255,173	0.78%	0.78%	0.00%	0.00%	0.00%
PHLX (X)	26,685,879	16.67%	9.75%	0.26%	0.07%	6.58%
Total	160,122,075	100.00%	85.73%	1.01%	4.21%	9.05%

Panel C: Message codes for the updated OPRA plan

Multi-leg code	Multi-leg Type	Execution Protocol	Value
f	Other than Buy Write	Limit order book	Multi-leg electronic
j			Multi-leg electronic against single leg(s)
g		Auction/Cross	Multi-leg auction
l			Multi-leg auction against single leg(s)
h			Multi-leg cross
i		Floor	Multi-leg floor trade
m			Multi-leg floor trade against single leg(s)
t			CBOE Combo trade
n	Stock options (Buy Write)	Limit order book	Stock options electronic trade
q			Stock options electronic trade against single leg(s)
k		Auction/Cross	Stock options auction
r			Stock options auction against single leg(s)
o			Stock options cross
p		Floor	Stock options floor trade
s			Stock options floor trade against single leg(s)

Panel D: Volume summary and market share across the exchanges that provided complex functionalities in July 2021

			Other Multi-leg Trades								Stock-options Trades (buy write)						
			LOB		Auction/Cross			Floor			LOB		Auction/Cross			Floor	
Exchange	Complex Volume	Market Share	LOB	LOB Against Single	Auction	Auction Against Single	Cross	Floor	Floor Against Single	CBOE Combo	LOB	LOB Against Single	Auction	Auction Against Single	Cross	Floor	Floor Against Single
AMEX (A)	22096152	10.67%	4.07%	0.14%	0.33%	0.00%	1.61%	4.22%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.30%	0.00%
BOX (B)	6603560	3.19%	0.29%	0.05%	0.12%	0.00%	0.76%	1.97%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
CBOE (C)	58862797	28.43%	12.98%	0.20%	8.60%	0.00%	0.23%	0.00%	4.07%	1.50%	0.40%	0.00%	0.03%	0.00%	0.01%	0.43%	0.00%
EMERALD (D)	10611090	5.12%	4.23%	0.89%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
EDGX (E)	7636028	3.69%	2.76%	0.10%	0.74%	0.00%	0.02%	0.00%	0.00%	0.00%	0.07%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
ISE (I)	36714939	17.73%	13.59%	1.44%	0.56%	0.00%	0.12%	0.00%	0.00%	0.00%	0.62%	0.00%	0.01%	0.00%	1.39%	0.00%	0.00%
MRX (J)	1975029	0.95%	0.27%	0.01%	0.66%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
MIAX (M)	19790873	9.56%	5.91%	0.23%	3.40%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
ARCA (N)	12185061	5.88%	1.19%	1.70%	0.00%	0.00%	0.07%	2.82%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.11%	0.00%
C2 (W)	5780115	2.79%	1.98%	0.81%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
PHLX (X)	24808803	11.98%	2.72%	0.40%	0.55%	0.00%	0.00%	3.59%	0.00%	0.00%	0.08%	0.00%	0.02%	0.00%	4.18%	0.45%	0.00%
All	207064447	100.00%	49.99%	5.96%	14.95%	0.00%	2.81%	12.60%	4.07%	1.50%	1.19%	0.00%	0.06%	0.00%	5.58%	1.29%	0.00%

Table 4

Sample ticker symbols, complex volumes, and complex ratios

This table presents the options trading volume for various types and complex ratio for each options class included in the sample for stock options and ETP options in (A) and (B) respectively over the period: 2016/01/04 through 2018/12/31. The sample includes the most liquid stock options and ETP options. To be included in the sample, the options class need to be listed across the whole sample period. The complex volume is the summation of the multi-leg volume and stock-option volume. The complex ratio is defined as the complex volume divided by total trading volume for each trading day and then average across the sample period for each options class.

Panel A: Equity options

Class	Singe-leg put	Single-leg call	Multi-leg put	Multi-leg call	Stock-option put	Stock-option call	Complex	Total	Complex ratio
AAPL	113,557,022	193,560,775	31,494,012	46,644,823	2,583,759	2,904,622	83,627,216	390,745,013	20.80%
AMD	37,059,976	68,887,256	7,196,168	9,227,956	1,399,046	858,195	18,681,365	124,628,597	13.34%
AMZN	25,099,141	36,709,102	19,130,729	17,225,336	98,111	147,125	36,601,301	98,409,544	37.42%
BABA	29,856,364	57,058,110	12,482,361	20,475,366	1,235,925	1,616,058	35,809,710	122,724,184	27.55%
BAC	66,451,957	156,917,516	12,646,789	35,329,180	2,416,602	3,999,961	54,392,532	277,762,005	17.65%
C	18,780,565	32,769,395	5,963,292	9,162,735	872,215	726,813	16,725,055	68,275,015	22.85%
CSCO	11,181,722	19,557,757	2,764,487	5,622,031	587,506	618,900	9,592,924	40,332,403	21.81%
DIS	8,697,451	14,448,688	4,359,095	4,223,208	730,736	632,866	9,945,905	33,092,044	28.23%
F	14,776,861	23,187,874	2,974,327	4,501,544	1,440,413	1,211,776	10,128,060	48,092,795	18.82%
FB	58,986,267	98,359,547	19,489,701	27,939,788	1,212,195	1,691,696	50,333,380	207,679,194	23.64%
FCX	15,719,547	25,513,367	3,383,476	4,056,904	969,330	674,076	9,083,786	50,316,700	16.46%
GE	35,429,945	51,956,099	13,405,402	11,093,400	3,549,525	2,311,627	30,359,954	117,745,998	23.28%
GILD	7,676,236	16,481,982	3,062,828	3,953,297	273,345	203,472	7,492,942	31,651,160	22.12%
GM	9,743,513	15,980,133	3,193,008	5,228,546	901,697	1,143,475	10,466,726	36,190,372	25.31%
GOOG	4,449,783	7,087,213	2,796,294	2,644,480	43,159	128,252	5,612,185	17,149,181	32.53%
GOOGL	5,612,120	10,519,936	4,492,023	4,542,603	35,528	54,114	9,124,268	25,256,324	35.83%
INTC	18,118,834	30,330,715	5,065,244	8,359,185	1,247,752	1,250,029	15,922,210	64,371,759	23.13%
JD	8,946,467	18,385,202	3,188,813	5,770,698	1,413,836	896,413	11,269,760	38,601,429	26.04%
JPM	15,868,007	27,247,895	5,215,454	6,500,737	514,167	532,980	12,763,338	55,879,240	22.11%
MSFT	25,535,301	45,789,060	7,280,915	12,533,049	842,716	1,340,235	21,996,915	93,321,276	22.09%
MU	27,060,115	63,620,003	8,420,143	14,170,757	1,506,044	2,003,497	26,100,441	116,780,559	20.16%
NFLX	32,930,264	45,936,625	13,342,228	13,492,553	303,249	349,804	27,487,834	106,354,723	25.50%
NVDA	25,026,535	40,274,779	9,510,173	9,451,920	318,049	268,629	19,548,771	84,850,085	21.69%
PBR	10,478,733	17,003,725	2,262,815	4,262,756	1,751,231	2,197,594	10,474,396	37,956,854	20.08%
QCOM	7,245,878	15,599,055	3,406,712	5,331,479	1,059,209	1,115,648	10,913,048	33,757,981	26.99%
T	18,260,404	29,142,037	7,467,351	10,147,700	1,883,403	1,110,276	20,608,730	68,011,171	26.47%
TWTR	20,451,023	46,577,509	4,712,677	7,459,888	704,322	883,404	13,760,291	80,788,823	15.76%
WFC	12,976,303	18,291,132	5,215,140	5,694,726	799,457	706,789	12,416,112	43,683,547	26.34%
WMT	8,396,750	15,033,952	3,136,432	4,707,334	343,814	280,328	8,467,908	31,898,610	24.45%
X	13,445,651	19,600,437	3,877,062	3,631,808	469,646	317,514	8,296,030	41,342,118	19.35%
XOM	12,055,097	16,789,811	3,832,370	4,167,976	385,709	342,533	8,728,588	37,573,496	22.15%

Panel B: ETP options

Class	Singe-leg put	Single-leg call	Multi-leg put	Multi-leg call	Stock-option put	Stock-option call	Complex	Total	Complex ratio
DIA	13,060,506	14,222,368	3,607,167	4,298,964	49,484	57,761	8,013,376	35,296,250	23.19%
EEM	60,784,121	47,396,579	77,396,029	51,434,124	25,275,786	22,358,055	176,463,994	284,644,694	59.93%
EFA	17,108,785	14,764,021	22,681,264	12,591,582	7,197,378	5,277,955	47,748,179	79,620,985	56.58%
EWZ	20,658,576	17,964,280	13,765,158	14,998,174	8,080,368	7,840,144	44,683,844	83,306,700	49.77%
FXI	17,389,686	17,859,811	13,266,595	14,928,332	6,116,914	5,302,509	39,614,350	74,863,847	49.13%
GDX	26,832,170	37,994,084	7,475,211	9,322,365	1,884,698	1,777,712	20,459,986	85,286,240	22.43%
GLD	27,715,153	43,701,283	10,722,270	23,638,082	1,971,233	4,756,389	41,087,974	112,504,410	33.47%
HYG	19,138,490	7,534,428	34,751,827	9,596,433	9,142,790	4,252,035	57,743,085	84,416,003	64.59%
IWM	129,253,795	78,190,476	124,793,811	36,481,636	9,994,872	5,165,895	176,436,214	383,880,485	44.06%
IYR	6,645,949	4,850,529	4,872,820	2,557,612	339,265	424,528	8,194,225	19,690,703	34.91%
KRE	5,687,254	5,229,165	5,526,141	3,123,497	585,216	888,788	10,123,642	21,040,061	35.91%
NUGT	6,329,912	14,113,686	1,536,574	2,013,653	6,700	45,163	3,602,090	24,045,688	14.80%
OIH	3,040,095	5,386,466	2,100,931	2,428,805	362,100	315,824	5,207,660	13,634,221	32.22%
QQQ	216,006,412	139,840,957	99,359,143	40,440,015	13,632,643	7,780,829	161,212,630	517,059,999	28.98%
SLV	10,983,295	28,903,016	2,221,228	4,387,363	181,867	492,513	7,282,971	47,169,282	14.35%
SMH	8,148,873	3,513,668	5,905,899	2,093,815	1,285,653	918,673	10,204,040	21,866,581	36.43%
SPY	863,038,657	617,842,004	375,057,904	228,733,122	18,238,028	12,964,886	634,993,940	2,115,874,601	29.50%
USO	38,328,988	40,322,283	18,790,770	12,848,745	2,916,301	3,255,167	37,810,983	116,462,254	29.36%
UVXY	25,773,352	44,878,016	6,289,777	8,572,639	53,791	149,212	15,065,419	85,716,787	17.87%
VXX	68,272,904	85,782,045	33,943,743	34,216,346	1,260,919	1,403,495	70,824,503	224,879,452	30.68%
XBI	6,925,577	6,219,982	7,775,215	3,401,601	202,254	63,418	11,442,488	24,588,047	35.80%
XLE	12,503,117	13,771,627	7,012,401	6,264,721	1,189,688	897,152	15,363,962	41,638,706	34.05%
XLF	33,982,296	43,573,422	24,179,183	22,086,223	4,955,539	4,590,557	55,811,502	133,367,220	36.29%
XLK	4,760,212	4,485,781	3,029,461	1,280,116	315,149	170,447	4,795,173	14,041,166	25.64%
XLP	5,034,113	2,814,434	3,683,178	1,176,749	544,993	162,060	5,566,980	13,415,527	30.66%
XLU	10,854,390	7,284,256	6,576,299	3,775,799	583,467	603,695	11,539,260	29,677,906	33.23%
XLV	4,345,290	4,039,566	2,729,832	2,269,065	332,935	259,877	5,591,709	13,976,565	32.27%
XME	4,117,906	2,720,062	2,557,312	1,513,759	182,886	145,763	4,399,720	11,237,688	27.91%
XOP	24,089,550	16,300,593	26,452,462	12,796,947	3,222,055	1,838,172	44,309,636	84,699,779	45.61%
XRT	5,670,339	3,178,768	5,104,260	1,923,706	396,577	196,740	7,621,283	16,470,390	37.17%

Table 5
Trades and trading volumes of complex strategies
This table presents the number of trades, number of packages, trade volumes, and other statistics for the various complex strategies for both the stock (Panel A) and ETP (Panel B) options classes. Trades is the total number of leg traded for each strategy, excluding the stock-option trades. Packages and Volume are the total number of packages and the total contract volume associated with each strategy. The columns headed % purchases and % sales indicate the percentages of the complex packages that are classified as buyer and seller-initiated, respectively. The last column reports the percentage of total complex volume due to each strategy.

Panel A: Stock options

Complex strategy	Trades	Packages	Volume	% purchases	% sales	% of complex volume
Calendar (Call)	2,168,312	1,007,150	28,680,756	49.22%	45.55%	5.10%
Calendar (Put)	1,876,218	872,651	21,963,756	54.49%	40.43%	3.91%
Diagonal (Call)	4,275,222	1,971,988	52,837,524	58.02%	37.02%	9.40%
Diagonal (Put)	1,650,341	761,209	18,653,070	51.87%	43.19%	3.32%
Vertical (Call)	10,982,350	5,085,900	114,505,810	50.15%	43.77%	20.36%
Vertical (Put)	11,151,454	5,161,475	88,696,236	41.84%	52.36%	15.77%
Combo	311,967	145,164	3,565,672	48.03%	46.97%	0.63%
Combo/straddle	2,161,780	1,011,810	51,195,750	46.74%	45.12%	9.10%
Straddle	542,982	256,436	6,377,584	42.11%	48.41%	1.13%
Strangle	987,412	468,321	8,846,852	40.05%	50.51%	1.57%
Butterfly (Call)	769,018	234,200	5,034,448	33.59%	64.91%	0.90%
Butterfly (Put)	411,400	126,463	2,676,088	34.60%	64.01%	0.48%
Iron Butterfly	683,672	163,067	2,559,260	35.82%	62.11%	0.46%
Iron Condor	2,840,706	674,079	10,911,896	32.52%	65.11%	1.94%
Vertical (Call|Ratio)	843,330	352,136	23,475,043	60.94%	37.80%	4.17%
Vertical (Put|Ratio)	479,135	204,134	10,818,708	54.27%	43.55%	1.92%
Vertical Roll (Call)	299,757	143,013	3,897,360	49.62%	49.45%	0.69%
Vertical Roll (Put)	486,975	226,372	6,664,412	22.11%	75.95%	1.19%
Straddle/strangle roll	243,188	117,190	15,215,040	47.73%	51.45%	2.71%
Other	6,617,694		85,746,019			15.25%

Panel B: ETP options

Spread Type	Trades	Packages	Volume	% purchases	% sales	% of complex volume
Calendar (Call)	1,430,017	669,381	32,394,196	43.50%	51.38%	2.12%
Calendar (Put)	1,346,042	635,194	40,781,032	34.24%	60.74%	2.67%
Diagonal (Call)	2,416,036	1,140,343	47,527,190	44.40%	50.03%	3.11%
Diagonal (Put)	1,847,018	871,144	60,869,522	37.97%	56.93%	3.98%
Vertical (Call)	7,782,145	3,683,614	172,367,442	50.13%	43.01%	11.28%
Vertical (Put)	9,859,357	4,617,747	347,050,046	51.64%	42.14%	22.71%
Combo	176,372	82,464	5,314,530	49.41%	45.88%	0.35%
Combo/Straddle	2,052,659	958,519	149,111,456	44.88%	46.42%	9.76%
Straddle	597,758	284,339	15,509,458	40.08%	48.88%	1.01%
Strangle	899,334	428,218	21,348,616	36.13%	52.39%	1.40%
Butterfly (Call)	346,734	108,471	8,788,112	33.78%	64.73%	0.58%
Butterfly (Put)	494,923	153,874	30,516,940	30.73%	65.03%	2.00%
Iron Butterfly	773,237	186,601	4,320,252	35.71%	61.82%	0.28%
Iron Condor	2,154,591	521,409	12,767,736	36.31%	61.56%	0.84%
Vertical (Call|Ratio)	513,492	220,482	45,820,081	54.49%	43.54%	3.00%
Vertical (Put|Ratio)	908,800	387,870	93,862,861	55.22%	42.32%	6.14%
Vertical Roll (Call)	391,496	188,448	10,532,608	36.96%	60.41%	0.69%

Vertical Roll (Put)	336,631	161,530	24,536,084	49.85%	49.57%	1.61%
Straddle/Strangle Roll	283,663	137,308	22,870,196	40.68%	56.93%	1.50%
Other	7,519,425		382,069,497			25.00%

Table 6
Complex Ratio Regression Analysis

This table presents regression results for the complex ratio of stock and ETP options. Our sample period is 2016/01/04 through 2018/12/31. The dependent variable is the complex ratio defined as daily complex volume divided by the daily total volume for each options class. All of the explanatory variables are defined in Table A1, Panel A. Standard errors are clustered at the options class and day level. * $p < 0.1$, * *$p < 0.05$, and *** $p < 0.01$.

	Stock options			ETP options			Stock and ETP options		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log(Price)	0.070***	0.035***	0.044	0.043***	-0.003	0.033***	0.033***	0.023**	0.022**
	(2.784)	(6.715)	(1.545)	(4.814)	(-0.082)	(4.390)	(3.280)	(2.175)	(2.190)
Log(Mcap)	-0.035	-0.005	-0.031	0.028	0.012	0.013	0.008	0.003	-0.002
	(-1.470)	(-0.743)	(-1.207)	(1.467)	(0.755)	(0.619)	(0.572)	(0.255)	(-0.155)
Turnover	-0.229**	-0.493*	-0.083	0.002	0.057**	0.001	0.001	0.025	0.001
	(-2.671)	(-1.902)	(-0.960)	(0.263)	(2.092)	(0.132)	(0.175)	(1.277)	(0.063)
VIX	-0.108*			0.102			-0.006		
	(-1.740)			(1.179)			(-0.118)		
Impl. Vol.	0.055*	-0.023	0.044	0.011	-0.204***	0.061**	0.029	-0.147***	0.065***
	(1.742)	(-0.526)	(1.352)	(0.544)	(-4.250)	(2.067)	(1.615)	(-3.660)	(3.206)
Dif. Impl. Vol.	0.149**	0.027	0.121*	0.109***	0.067	0.103**	0.127***	0.095	0.091**
	(2.688)	(0.276)	(1.828)	(3.003)	(0.468)	(2.142)	(3.487)	(0.878)	(2.397)
Return5_P	0.049*	0.041	0.030	0.069*	0.079	0.027	0.055***	0.056	0.033
	(1.792)	(1.009)	(1.101)	(1.816)	(1.468)	(0.658)	(2.765)	(1.554)	(1.642)
Return5_N	0.146***	0.100*	0.058	0.314**	0.210*	0.275**	0.201***	0.128**	0.149***
	(3.221)	(2.020)	(1.222)	(2.724)	(1.699)	(2.486)	(3.947)	(2.466)	(3.367)
Friday	0.011***			0.002			0.007**		
	(3.465)			(0.600)			(2.525)		
Third Week	0.018***			0.004			0.011***		
	(5.467)			(1.320)			(4.298)		
ETP								0.118***	
								(2.788)	
Constant	0.347*	0.164*	0.401**	-0.093	0.299**	0.078	0.061	0.150	0.199*
	(1.957)	(1.973)	(2.045)	(-0.528)	(2.269)	(0.403)	(0.558)	(1.518)	(1.765)
Class fxd. effs.	Y	N	Y	Y	N	Y	Y	N	Y
Day fxd. effs.	N	Y	Y	N	Y	Y	N	Y	Y
Observations	23,002	23,002	23,002	22,453	22,453	22,453	45,455	45,455	45,455
R-squared	0.170	0.169	0.219	0.307	0.083	0.335	0.334	0.166	0.353

Table 7

Regressions analyzing the impact of splits on complex trading volume

This table displays the results of regressions that analyze the impact of splits on the complex ratio and scaled volume. The complex ratio is the ratio of complex volume to total options volume, and scaled volume is the ratio of option volume (in contracts) to stock trading volume (CRSP variable shrout, which is in thousands). The sample consists of the 53 split events between July 2011 and December 2021, for which the average daily options volume during the 30 trading days prior to the split was at least 500 contracts. For each split we identify a matching symbol that did not experience a split. The matching symbol must have the same CRSP shrcd and 2-digit SIC code. Of the symbols that match on these two criteria, we select the one that has the smallest Euclidean distance from the splitting symbol in the space spanned by option volume and underlying market capitalization, after both variables are first transformed to their ordinal ranks. We include the dates from 30 trading days before to 30 trading days after the split.

	Complex ratio			Scaled volume		
	(1)	(2)	(3)	(4)	(5)	(6)
treat*after	-0.057***	-0.007	-0.004	-0.233***	-0.040	0.017
	(-3.995)	(-0.430)	(-0.244)	(-3.653)	(-0.372)	(0.183)
treat*after*log(sfactor)		-0.012***	-0.013***		-0.048*	-0.059**
		(-4.395)	(-4.428)		(-1.757)	(-2.459)
to			0.000			0.001***
			(1.466)			(2.922)
Impl. Vol.			-0.011			-0.359**
			(-0.238)			(-2.199)
volatility			-1.100**			7.166***
			(-2.113)			(3.072)
log(Mcap)			0.010			0.397**
			(0.448)			(2.162)
Constant	0.246***	0.246***	0.100	0.320***	0.320***	-6.058**
	(77.081)	(84.931)	(0.286)	(20.729)	(21.805)	(-2.087)
Class fixed effect	Y	Y	Y	Y	Y	Y
Day fixed effect	Y	Y	Y	Y	Y	Y
Obs	6,360	6,360	6,360	6,360	6,360	6,360
R-squared	0.207	0.211	0.212	0.728	0.733	0.761

Table 8

Characterization of vertical, calendar, diagonal and vertical ratio spreads

This table presents the break-down of volume of Vertical, Calendar, Diagonal and Vertical Ratio Spreads into four categories: (1) ITM-ITM, (2) ITM-OTM, (3) OTM-ITM, and (4) OTM-OTM for stock options in (A) and ETP options in (B) respectively. The explicit categorization is defined in Table A1. We first derive the break-down for each options class and then average across the classes.

(A) Stock options

Spread Type	ITM-ITM	ITM-OTM	OTM-ITM	OTM-OTM
Calendar (Call)	39.46%	N/A	N/A	60.54%
Calendar (Put)	53.29%	N/A	N/A	46.71%
Diagonal (Call)	23.65%	32.75%	5.73%	37.87%
Diagonal (Put)	21.98%	26.29%	6.11%	45.62%
Vertical (Call)	18.18%	24.34%	N/A	57.48%
Vertical (Put)	9.84%	22.63%	N/A	67.53%
Vertical (Call|Ratio)	7.49%	29.57%	N/A	62.94%
Vertical (Put|Ratio)	14.11%	24.52%	N/A	61.38%

(B) ETP options

Spread Type	ITM-ITM	ITM-OTM	OTM-ITM	OTM-OTM
Calendar (Call)	42.15%	N/A	N/A	57.85%
Calendar (Put)	29.56%	N/A	N/A	70.44%
Diagonal (Call)	21.08%	33.77%	5.22%	39.93%
Diagonal (Put)	13.95%	28.21%	5.35%	52.50%
Vertical (Call)	12.32%	24.38%	N/A	63.30%
Vertical (Put)	5.72%	22.19%	N/A	72.09%
Vertical (Call|Ratio)	3.65%	30.97%	N/A	65.38%
Vertical (Put|Ratio)	5.10%	26.85%	N/A	68.04%

Table 9

Relation between complex spread volumes and recent returns

This table presents regression results for the daily ITM volumes for the stock options sample in (A) and the ETP options sample in (B). For the Vertical trades, it is the volume of certain trades where the lower (higher) strike is ITM for calls (puts). For the Calendar and Diagonal trades, it is the volume of certain trades where the near leg is ITM. The explanatory variables are the net returns at day t, t-1… t-10. In all specifications, we include both the options class and day fixed effects, and the standard errors are clustered at both the options class and day levels. * $p < 0.1$, * *$p < 0.05$, and *** $p < 0.01$.

Panel A: Stock options

	Call				Put			
	Vertical	Ratio	Calendar	Diagonal	Vertical	Ratio	Calendar	Diagonal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$return_t$	0.028***	0.003**	0.003***	0.015***	-0.020***	-0.009	-0.005***	-0.007***
	(4.386)	(2.373)	(4.398)	(4.887)	(-3.507)	(-1.610)	(-3.199)	(-3.445)
$return_{t-1}$	0.024***	0.003***	0.002***	0.012***	-0.010***	-0.002**	-0.006***	-0.006***
	(4.930)	(3.774)	(4.340)	(5.956)	(-3.285)	(-2.484)	(-3.226)	(-4.204)
$return_{t-2}$	0.015***	0.003**	0.002***	0.009***	-0.008***	-0.002**	-0.004***	-0.003***
	(5.610)	(2.452)	(3.190)	(4.243)	(-2.823)	(-2.211)	(-3.547)	(-4.260)
$return_{t-3}$	0.010***	0.002*	0.002***	0.007***	-0.004***	-0.002	-0.002***	-0.003***
	(4.123)	(1.756)	(4.183)	(5.271)	(-2.763)	(-1.439)	(-3.649)	(-4.067)
$return_{t-4}$	0.007***	-0.000	0.001***	0.006***	-0.004***	-0.001**	-0.003***	-0.002***
	(3.390)	(-0.644)	(5.324)	(5.202)	(-3.284)	(-2.299)	(-3.238)	(-2.980)
$return_{t-5}$	0.006***	-0.001	0.001***	0.004***	-0.005***	-0.002	-0.003***	-0.002***
	(3.799)	(-0.884)	(3.207)	(4.540)	(-3.000)	(-1.572)	(-3.505)	(-4.954)
$return_{t-6}$	0.005***	0.001	0.001***	0.002***	-0.002***	-0.000	-0.003***	-0.001**
	(3.210)	(0.938)	(2.803)	(3.100)	(-2.795)	(-0.948)	(-2.772)	(-2.486)
$return_{t-7}$	0.006***	0.001*	0.001***	0.004***	-0.002	-0.001	-0.001**	-0.001*
	(3.662)	(1.802)	(3.621)	(4.385)	(-1.430)	(-1.098)	(-2.389)	(-1.763)
$return_{t-8}$	0.004**	0.002	0.001**	0.003***	-0.001	-0.002	-0.001	-0.001***
	(2.201)	(1.407)	(2.394)	(2.785)	(-1.132)	(-1.577)	(-1.294)	(-3.125)
$return_{t-9}$	0.003	0.001**	0.001**	0.002***	-0.000	-0.001	-0.001***	-0.001**
	(1.538)	(2.365)	(2.526)	(2.943)	(-0.162)	(-1.273)	(-3.053)	(-2.291)
$return_{t-10}$	0.003**	0.000	0.000	0.001*	-0.001	-0.000	-0.003*	-0.001
	(2.135)	(0.593)	(0.918)	(1.752)	(-1.047)	(-1.321)	(-1.971)	(-1.333)
Constant	0.001***	0.000***	0.000***	0.001***	0.001***	0.000***	0.000***	0.000***
	(48.248)	(46.738)	(41.196)	(39.541)	(47.814)	(13.942)	(28.287)	(34.016)
Obs.	23,033	23,064	23,033	23,033	23,033	23,033	23,033	23,033
R-squared	0.272	0.018	0.103	0.154	0.167	0.023	0.069	0.087

Panel B: ETP options

	Call				Put			
	Vertical (1)	Ratio (2)	Calendar (3)	Diagonal (4)	Vertical (5)	Ratio (6)	Calendar (7)	Diagonal (8)
$return_t$	0.405**	0.060**	0.050**	0.233***	-0.241*	-0.159*	-0.031***	-0.131***
	(2.318)	(2.558)	(2.565)	(3.041)	(-1.865)	(-1.862)	(-3.219)	(-2.960)
$return_{t-1}$	0.369**	0.043	0.093**	0.223***	-0.219	-0.110**	-0.029**	-0.082***
	(2.725)	(1.597)	(2.408)	(3.357)	(-1.692)	(-2.056)	(-2.478)	(-3.216)
$return_{t-2}$	0.292**	0.026**	0.053**	0.170***	-0.119**	-0.095*	-0.034**	-0.069***
	(2.179)	(2.642)	(2.427)	(4.049)	(-2.185)	(-1.876)	(-2.621)	(-2.773)
$return_{t-3}$	0.209**	0.045**	0.057*	0.090***	-0.066	-0.065	-0.031	-0.061*
	(2.390)	(2.306)	(1.937)	(3.666)	(-1.083)	(-1.604)	(-1.630)	(-1.955)
$return_{t-4}$	0.225**	0.040**	0.051**	0.102**	-0.088	-0.055	-0.029***	-0.041
	(2.327)	(2.124)	(2.345)	(2.384)	(-1.175)	(-1.245)	(-3.485)	(-1.517)
$return_{t-5}$	0.083	-0.003	0.053*	0.053**	-0.036	-0.025	-0.016	-0.032
	(1.673)	(-0.223)	(1.923)	(2.342)	(-0.788)	(-1.073)	(-1.428)	(-1.211)
$return_{t-6}$	0.172**	0.021	0.053**	0.065***	-0.005	-0.028	-0.002	-0.009
	(2.552)	(1.455)	(2.268)	(3.007)	(-0.124)	(-0.925)	(-0.275)	(-0.483)
$return_{t-7}$	0.134**	-0.004	0.051**	0.068***	-0.043	-0.003	0.003	-0.015
	(2.141)	(-0.241)	(2.093)	(6.280)	(-0.903)	(-0.140)	(0.309)	(-0.898)
$return_{t-8}$	0.134*	0.030	0.031**	0.073**	0.028	0.005	-0.004	-0.012
	(1.850)	(1.427)	(2.123)	(2.354)	(0.819)	(0.386)	(-0.715)	(-0.588)
$return_{t-9}$	0.110*	0.060	0.020*	0.098**	-0.016	0.044	-0.015**	-0.019
	(1.721)	(1.331)	(1.796)	(2.501)	(-0.414)	(0.858)	(-2.255)	(-0.803)
$return_{t-10}$	0.046	0.029	0.025**	0.040	0.000	-0.008	-0.002	0.005
	(1.262)	(1.601)	(2.121)	(1.447)	(0.011)	(-0.202)	(-0.351)	(0.240)
Constant	0.017***	0.003***	0.003***	0.006***	0.020***	0.008***	0.003***	0.006***
	(42.231)	(44.027)	(29.261)	(32.888)	(37.468)	(38.022)	(60.002)	(52.289)
Obs.	22,280	22,312	22,280	22,280	22,280	22,280	22,280	22,280
R-squared	0.221	0.017	0.085	0.108	0.088	0.021	0.048	0.042

Table 10
Open interest Regression Analysis: Baseline
This table presents the baseline results of the open interest regression analysis for the stock options sample and ETP options sample during the sample period 2016/01/04 through 2018/12/31. The dependent variable is the change of open interest between day t and day t-1 on the options series level. The explanatory variables include the various components of the total trading volume for the same series on day t. All the explanatory variables are defined in Appendix A. All the specifications include options class and the number of weeks until expiration fixed effects. Standard errors are clustered at the options class and the number of weeks till expiration. * $p < 0.1$, * *$p < 0.05$, and *** $p < 0.01$.

	Stock options		ETP options	
Covariate	Call	Put	Call	Put
vv	0.103**	0.287***	0.602***	0.450***
	(2.306)	(4.740)	(6.540)	(7.138)
vr	0.388***	0.440***	0.468***	0.481***
	(6.472)	(7.252)	(7.035)	(8.681)
cv	0.217	0.269	0.358**	0.470***
	(1.196)	(1.574)	(2.626)	(6.081)
dv	0.196**	0.403***	0.107	0.319***
	(2.349)	(3.607)	(1.002)	(3.976)
straddle	0.818***	0.787***	0.496	0.497*
	(11.069)	(13.698)	(1.487)	(1.832)
strangle	0.807***	0.825***	0.450***	0.499***
	(3.438)	(6.090)	(3.338)	(3.723)
bf	0.191	0.490**	0.571***	0.273***
	(0.652)	(2.500)	(5.534)	(3.779)
ica	0.501*	-0.138	0.146	0.146
	(1.890)	(-0.535)	(0.410)	(0.531)
vertical_roll	0.149	0.184	0.081	0.183**
	(0.564)	(0.787)	(0.423)	(2.665)
straddle_roll	0.405**	0.455**	0.277	0.286
	(2.191)	(2.304)	(1.573)	(1.658)
bw	0.694***	0.497***	0.469***	0.456***
	(15.098)	(4.167)	(5.119)	(8.138)
regular	0.166***	0.319***	0.175**	0.195***
	(4.221)	(4.844)	(2.435)	(3.476)
other	0.335***	0.408***	0.425***	0.394***
	(2.988)	(4.058)	(4.098)	(7.850)
Constant	6.044	-1.178	3.762	6.087
	(1.298)	(-0.239)	(0.406)	(0.793)
Observations	12,315,192	12,314,950	15,343,947	15,343,856
R-squared	0.179	0.126	0.263	0.281

Table 11
Open Interest Regression Analysis: Rolling 1

This table presents the results of the open interest regression analysis to examine whether market participants roll simple or complex options positions. The dependent variable is the change of open interest between day t and day t-1 on the options series level. The explanatory variables include the various components of the total trading volume for the same series on day t. All the explanatory variables are defined in Appendix A. All the specifications include options class and the number of weeks till expiration fixed effects. Standard errors are clustered at the options class and the number of weeks till expiration. * $p < 0.1$, * *$p < 0.05$, and *** $p < 0.01$.

	Stock options		ETP options	
Covariate	Call	Put	Call	Put
vfiov	-0.322*	-0.559**	-0.578***	-0.425***
	(-1.977)	(-2.741)	(-3.380)	(-6.401)
vfiiv	-1.346***	-0.587***	-1.419***	-0.564***
	(-6.383)	(-3.308)	(-4.309)	(-10.737)
vfov_b	0.601***	0.982***	0.928***	0.755***
	(3.725)	(5.368)	(32.768)	(18.656)
vfov_s	0.857***	0.285**	0.515**	0.049
	(8.379)	(2.426)	(2.496)	(0.400)
vsiov	0.499***	-0.245	-0.185	0.353**
	(3.594)	(-1.373)	(-0.525)	(2.158)
vsiiv	-1.195***	-0.556***	-1.279***	-0.088
	(-3.671)	(-3.168)	(-3.619)	(-0.306)
vsov_b	0.212	0.856***	0.832***	0.661***
	(1.411)	(3.901)	(11.208)	(14.725)
vsov_s	-0.547***	0.183	0.592***	0.141
	(-6.699)	(1.674)	(2.763)	(1.119)
vrfiiv	-0.196	-0.251**	-0.386	-0.368**
	(-0.866)	(-2.127)	(-0.560)	(-2.688)
vrfiov	-0.657***	-0.324	-0.408***	-0.222
	(-3.456)	(-1.426)	(-3.308)	(-1.432)
vrfov	0.677***	0.264	0.538***	0.537***
	(4.265)	(1.023)	(6.288)	(8.506)
vrsiiv	0.474***	0.038	-0.857***	0.155
	(3.379)	(0.278)	(-5.788)	(0.809)
vrsiov	0.563***	0.832***	0.597***	0.597***
	(8.678)	(6.732)	(5.183)	(3.949)
vrsov	0.388***	0.590***	0.519***	0.505***
	(2.921)	(3.565)	(6.595)	(7.740)
cniv	-0.629**	-0.505***	-0.189	-0.460**
	(-2.398)	(-9.251)	(-0.447)	(-2.249)
cnov	-0.107	-0.047	-0.014	0.179
	(-0.665)	(-0.176)	(-0.061)	(1.103)
cfiv	0.760***	0.783***	0.560***	0.626***
	(6.284)	(7.397)	(3.318)	(8.019)
cfov	0.468*	0.835***	0.858***	0.807***
	(1.923)	(12.941)	(8.325)	(14.999)
dniv	-0.699***	-0.158	-1.068***	-0.493***
	(-8.898)	(-0.747)	(-5.648)	(-3.527)
dnov	0.019	-0.070	-0.196*	-0.151**
	(0.086)	(-0.559)	(-1.842)	(-2.117)
dfiv	0.788***	1.007***	0.912***	0.833***
	(11.306)	(8.117)	(11.509)	(13.562)
dfov	0.518***	0.694***	0.769***	0.822***
	(3.036)	(2.985)	(8.813)	(25.153)
straddle	0.814***	0.801***	0.519	0.541**
	(10.445)	(10.891)	(1.658)	(2.139)
strangle	0.726***	0.819***	0.430***	0.453***
	(3.083)	(5.914)	(3.024)	(2.987)

bf	0.174	0.487***	0.565***	0.276***
	(0.576)	(3.521)	(5.600)	(3.794)
ica	0.306	-0.225	0.130	0.146
	(1.387)	(-0.792)	(0.435)	(0.518)
vertical_roll	0.159	0.117	0.103	0.189**
	(0.643)	(0.811)	(0.503)	(2.177)
straddle_roll	0.409**	0.456**	0.280	0.288
	(2.136)	(2.283)	(1.541)	(1.691)
bw	0.685***	0.498***	0.466***	0.458***
	(13.026)	(4.119)	(5.059)	(8.194)
regular	0.180***	0.321***	0.200***	0.203***
	(4.910)	(5.049)	(3.176)	(3.971)
other	0.351***	0.405***	0.428***	0.395***
	(3.323)	(3.853)	(4.139)	(7.774)
Constant	5.813	-1.199	4.102	6.688
	(1.406)	(-0.267)	(0.482)	(0.903)
Observations	12,315,192	12,314,950	15,343,947	15,343,856
R-squared	0.290	0.147	0.342	0.330

Table 12
Open Interest Regression Analysis: Rolling 2
This table presents the results of the open interest regression analysis to examine whether market participants roll options which are close to the expiration dates. The dependent variable is the change of open interest between day t and day t-1 on the options series level. The explanatory variables include the various components of the total trading volume for the same series on day t. All the explanatory variables are defined in Appendix A. All the specifications include options class and the number of weeks till expiration fixed effects. Standard errors are clustered at the options class and the number of weeks till expiration. * $p < 0.1$, * *$p < 0.05$, and *** $p < 0.01$.

	Stock options		ETP options	
Covariate	Call	Put	Call	Put
vv	0.116**	0.285***	0.606***	0.452***
	(2.243)	(4.219)	(6.542)	(7.188)
vr	0.386***	0.444***	0.470***	0.481***
	(6.351)	(6.976)	(7.133)	(8.424)
cnev	-0.489**	-0.541***	-1.409***	-0.952***
	(-2.512)	(-10.736)	(-3.302)	(-4.741)
cnnv	-0.087	-0.065	0.189	0.258*
	(-0.446)	(-0.285)	(1.056)	(1.924)
cfev	0.205	0.745***	0.268*	0.604***
	(0.505)	(6.584)	(1.784)	(10.075)
cfnv	0.876***	0.850***	0.891***	0.811***
	(15.692)	(20.337)	(7.398)	(18.482)
dnev	-0.131	-0.564***	-1.268***	-0.676***
	(-0.980)	(-4.172)	(-5.476)	(-4.159)
dnnv	-0.423***	-0.019	-0.481***	-0.127**
	(-3.310)	(-0.131)	(-2.829)	(-2.449)
dfev	0.719***	0.827***	0.741***	0.780***
	(6.637)	(7.871)	(6.934)	(11.521)
dfnv	0.622***	0.869***	0.871***	0.831***
	(10.528)	(4.248)	(16.140)	(32.875)
straddle	0.824***	0.788***	0.516*	0.513**
	(11.061)	(12.002)	(1.876)	(2.087)
strangle	0.798***	0.832***	0.451***	0.497***
	(3.267)	(6.128)	(3.318)	(3.799)
bf	0.195	0.513***	0.574***	0.272***
	(0.668)	(3.356)	(5.676)	(3.731)
ica	0.544**	-0.084	0.171	0.155
	(2.241)	(-0.341)	(0.533)	(0.576)
vertical_roll_n	-0.530	-0.452	-0.791***	-0.415***
	(-1.518)	(-1.242)	(-6.937)	(-5.228)
vertical_roll_f	0.863***	0.755***	0.949***	0.768***
	(8.792)	(8.113)	(36.598)	(5.795)
straddle_roll_n	-0.027	-0.031	-0.375***	-0.375***
	(-0.273)	(-0.336)	(-3.294)	(-3.358)
straddle_roll_f	0.779***	0.874***	0.930***	0.944***
	(10.091)	(21.384)	(20.262)	(25.801)
bw	0.691***	0.497***	0.466***	0.456***
	(14.011)	(4.137)	(5.158)	(8.247)
regular	0.169***	0.323***	0.187**	0.204***
	(4.272)	(4.960)	(2.677)	(3.764)
other	0.333***	0.404***	0.423***	0.396***
	(2.968)	(4.028)	(4.062)	(7.866)
Constant	5.572	-1.506	3.043	5.426
	(1.287)	(-0.346)	(0.393)	(0.796)
Observations	12,315,192	12,314,950	15,343,947	15,343,856
R-squared	0.205	0.136	0.296	0.309

Table 13
Characteristics of Straddle, Strangle, Butterfly, Iron Butterfly and Iron Condor Trades
This table presents statistics describing the characteristics of straddle, strangle, butterfly, iron butterfly, and iron condor trades for stock options in Panel A and ETP options in Panel B, respectively. Specifically, it displays the average values of the following ratios: *MK/S*, defined as the average of the strikes divided by the underlying price; *DK/S*, which is only applicable to strangles and iron condors, is defined as the absolute difference between the two inner strikes divided by the underlying price and *DK2/S*, which is applicable to the three- and four-leg trades, defined as the absolute difference between the outer strike on either side and the next inner strike, divided by the underlying price. We compute the averages displayed in the table by first computing the measures for each options class and then averaging across the classes.

Panel A: Stock options

Spread Type	Side	MK/S	DK/S	DK2/S
Straddle	Sell	1.045	N/A	N/A
Straddle	Buy	1.019	N/A	N/A
Strangle	Sell	1.001	0.123	N/A
Strangle	Buy	1.000	0.112	N/A
Butterfly (Call)	Sell	1.076	N/A	0.060
Butterfly (Call)	Buy	1.047	N/A	0.063
Butterfly (Put)	Sell	0.961	N/A	0.053
Butterfly (Put)	Buy	0.965	N/A	0.063
Iron Butterfly	Sell	1.002	N/A	0.064
Iron Butterfly	Buy	1.002	N/A	0.070
Iron Condor	Sell	1.002	0.117	0.044
Iron Condor	Buy	1.001	0.117	0.046

Panel B: ETP options

Spread Type	Side	MK/S	DK/S	DK2/S
Straddle	Sell	1.037	N/A	N/A
Straddle	Buy	1.096	N/A	N/A
Strangle	Sell	1.004	0.109	N/A
Strangle	Buy	1.004	0.114	N/A
Butterfly (Call)	Sell	1.100	N/A	0.060
Butterfly (Call)	Buy	1.047	N/A	0.065
Butterfly (Put)	Sell	0.929	N/A	0.057
Butterfly (Put)	Buy	0.956	N/A	0.059
Iron Butterfly	Sell	1.001	N/A	0.083
Iron Butterfly	Buy	0.999	N/A	0.085
Iron Condor	Sell	0.999	0.111	0.048
Iron Condor	Buy	0.999	0.111	0.050

Table 14
Open Interest Regression Analysis: Volatility Spreads

This table presents the results of the open interest regression analysis to examine whether market participants would long certain volatility spreads to close out the previous short complex positions on volatility. The dependent variable is the change of open interest between day t and day t-1 on the options series level. The explanatory variables include the various components of the total trading volume for the same series on day t. All the explanatory variables are defined in Appendix A. All the specifications include options class and the number of weeks till expiration fixed effects. Standard errors are clustered at the options class and the number of weeks till expiration. * $p < 0.1$, * *$p < 0.05$, and *** $p < 0.01$.

	Stock options		ETP options	
	Call	Put	Call	Put
vv	0.103**	0.288***	0.451***	0.601***
	(2.265)	(4.673)	(7.077)	(6.515)
vr	0.389***	0.440***	0.481***	0.468***
	(6.444)	(7.235)	(8.553)	(7.005)
cv	0.217	0.269	0.470***	0.357**
	(1.197)	(1.579)	(6.096)	(2.619)
dv	0.196**	0.403***	0.318***	0.107
	(2.329)	(3.606)	(3.932)	(1.003)
straddle_buy_volume	0.867***	0.724***	0.807***	0.805***
	(10.051)	(6.336)	(4.003)	(4.016)
straddle_sell_volume	0.757***	0.783***	0.236	0.272
	(8.008)	(10.406)	(0.595)	(0.485)
strangle_buy_volume	0.572	0.743***	0.086	-0.052
	(1.673)	(3.648)	(0.261)	(-0.162)
strangle_sell_volume	0.981***	0.864***	0.741***	0.773***
	(4.872)	(6.166)	(9.278)	(9.710)
bf_buy_volume	-0.323	-0.339	-0.126	-0.089
	(-1.127)	(-0.871)	(-0.791)	(-0.556)
bf_sell_volume	0.304	0.699***	0.336***	0.781***
	(0.862)	(4.086)	(5.273)	(5.853)
ica_buy_volume	-0.240	-0.901	-0.137	-0.285
	(-0.356)	(-1.142)	(-0.212)	(-0.359)
ica_sell_volume	0.838**	0.213	0.404***	0.551***
	(2.645)	(0.644)	(3.496)	(3.110)
vertical_roll	0.152	0.185	0.183**	0.081
	(0.576)	(0.811)	(2.663)	(0.417)
straddle_roll	0.406**	0.455**	0.286	0.277
	(2.187)	(2.303)	(1.657)	(1.583)
bw	0.694***	0.497***	0.456***	0.469***
	(15.038)	(4.161)	(8.075)	(5.114)
regular	0.166***	0.319***	0.195***	0.175**
	(4.236)	(4.872)	(3.467)	(2.429)
other	0.338***	0.410***	0.397***	0.425***
	(3.022)	(4.087)	(7.964)	(4.108)
Constant	6.072	-1.143	6.039	3.715
	(1.301)	(-0.232)	(0.788)	(0.402)
Observations	12,315,192	12,314,950	15,343,856	15,343,947
R-squared	0.179	0.126	0.283	0.264

Appendix A. Variables Used in the Regression Models

Table A1
Definition of the variables used in the regression models

This table provides the definitions of the variables used in the regression models. Panel A defines the variables used in the regressions for which results are reported in Tables 6-7, and Panel C defines the variables used in Tables 10-12 and 14. Panel B defines the classifications used in Table 8.

Panel A

Variable	Definition
Log(Price)	The logarithm of the daily opening price of the underlying security.
Log(Mcap)	The logrithm of the daily market cap of the underlying security.
Turnover	The daily turnover ratio of the underlying security.
VIX	The opening VIX value divided by 100.
Impl. Vol.	The previous eod 30-day ATM implied volaility derived from the volatility surface table in Optionmetrics data.
Dif. Impl. Vol.	The absolute difference between the 30-day ATM and 91-day ATM implied volatilities from the volatility surface table in Optionmetrics data.
Return5_P	This variable is set to be net return from t-5 to t-1 if it is positive, and 0 otherwise.
Return5_N	This variable is set to be the absolute value of the net return from t-5 to t-1 if it is negative, and 0 otherwise.
Friday	Indicator variable which equals 1 if day t is a Friday and 0 otherwise.
Third Week	Indicator variable which equals 1 if day t is at the third week of the month and 0 otherwise.
ETP	Indicator variable which equals 1 if the underlying security is an ETP and 0 otherwise.

Panel B

Variable	Definition
ITM	A particular trade of an option series is deemed as ITM if the strike is strictly less (greater) than underlying price at the time of the trade for Call (Put).
OTM	A particular trade of an option series is deemed as OTM if the strike is greater (less) than or equal to the underlying price at the time of the trade for Call (Put).
ITM-OTM	For Vertical and Vertical Ratio trades, it means that one leg is ITM and the other leg is OTM. For Calendar and Diagonal trades, it means that the near leg is ITM and the far leg is OTM.
ITM-ITM	Both legs of the trades are ITM.
OTM-OTM	Both legs of the trades are OTM.
OTM-ITM	Only applicable to the Diagonal trades. It means that the near leg is OTM and the far leg is ITM.

Panel C

Variable	Definition
vv	The Vertical volume associated with the option series on day t.
vfiiv	The volume of the lower (higher) strike of the Vertical Call (Put) when both legs are ITM associated with the option series on day t.
vfiov	The volume of the lower (higher) strike of the Vertical Call (Put) when one leg is ITM and the other is OTM associated with the option series on day t.
vfov_s	The volume of the lower (higher) strike of the Vertical Call (Put) when both legs are OTM and the whole package is seller-initiated (i.e., credit spread) associated with the option series on day t.
vfov_b	The volume of the lower (higher) strike of the Vertical Call (Put) when both legs are OTM and the whole package is buyer-initiated (i.e., debit spread) associated with the option series on day t.
vsiiv	The volume of the higher (lower) strike of the Vertical Call (Put) when both legs are ITM associated with the option series on day t.
vsiov	The volume of the higher (lower) strike of the Vertical Call (Put) when one leg is ITM and the other is OTM associated with the option series on day t.
vsov_s	The volume of the higher (lower) strike of the Vertical Call (Put) when both legs are OTM and the whole package is seller-initiated (i.e., credit spread) associated with the option series on day t.
vsov_b	The volume of the higher (lower) strike of the Vertical Call (Put) when both legs are OTM and the whole package is buyer-initiated (i.e., debit spread) associated with the option series on day t.
vrfiiv	The volume of the lower (higher) strike of the Vertical Ratio Call (Put) when both legs are ITM associated with the option series on day t.
vrfiov	The volume of the lower (higher) strike of the Vertical Ratio Call (Put) when one leg is ITM and the other is OTM associated with the option series on day t.
vrfov	The volume of the lower (higher) strike of the Vertical Ratio Call (Put) when both legs are OTM associated with the option series on day t.
vrsiiv	The volume of the higher (lower) strike of the Vertical Ratio Call (Put) when both legs are ITM associated with the option series on day t.
vrsiov	The volume of the higher (lower) strike of the Vertical Ratio Call (Put) when one leg is ITM and the other is OTM associated with the option series on day t.
vrsov	The volume of the higher (lower) strike of the Vertical Ratio Call (Put) when both legs are OTM associated with the option series on day t.
cnov	The volume of the far leg of the Calendar when the near leg is OTM associated with the option series on day t.
cv	The Calendar volume associated with the option series on day t.
cniv	The volume of the near leg of the Calendar when the near leg is ITM associated with the option series on day t.
cnov	The volume of the near leg of the Calendar when the near leg is OTM associated with the option series on day t.
cfiv	The volume of the far leg of the Calendar when the near leg is ITM associated with the option series on day t.
cfov	The volume of the far leg of the Calendar when the near leg is OTM associated with the option series on day t.
dv	The Diagonal volume associated with the option series on day t.
dniv	The volume of the near leg of the Diagonal when the near leg is ITM associated with the option series on day t.
dnov	The volume of the near leg of the Diagonal when the near leg is OTM associated with the option series on day t.
dfiv	The volume of the far leg of the Diagonal when the near leg is ITM associated with the option series on day t.
dfov	The volume of the far leg of the Diagonal when the near leg is OTM associated with the option series on day t.
cnev	The volume of the near leg of the Calendar when the near leg will expire within a week.
cnnv	The volume of the near leg of the Calendar when the near leg will not expire within a week.
cfev	The volume of the far leg of the Calendar when the near leg will expire within a week.
cfnv	The volume of the far leg of the Calendar when the near leg will not expire within a week.
dnev	The volume of the near leg of the Diagonal when the near leg will expire within a week.
dnnv	The volume of the near leg of the Diagonal when the near leg will not expire within a week.
dfev	The volume of the far leg of the Diagonal when the near leg will expire within a week.
dfnv	The volume of the far leg of the Diagonal when the near leg will not expire within a week.
straddle	The Straddle volume associated with the option series on day t.
straddle_buy	The long Straddle volume associated with the option series on day t.
straddle_sell	The short Straddle volume associated with the option series on day t.
strangle	The Strangle volume associated with the option series on day t.
strangle_buy	The long Strangle volume associated with the option series on day t.
strangle_sell	The short Strangle volume associated with the option series on day t.
bf	The Butterfly volume associated with the option series on day t.
bf_buy	The long Butterfly (credit spread) volume associated with the option series on day t.
bf_sell	The short Butterfly (debit spread) volume associated with the option series on day t.
ica	The Iron Butterfly and Iron Condor volume associated with the option series on day t.

ica_buy	The long Iron Butterfly and Iron Condor volume associated with the option series on day t.
ica_sell	The short Iron Butterfly and Iron Condor volume associated with the option series on day t.
vr	The Vertical Ratio volume associated with the option series on day t.
vr_buy	The long Vertical Ratio volume associated with the option series on day t.
vr_sell	The short Vertical Ratio volume associated with the option series on day t.
vertical_roll	The Vertical Roll volume associated with the option series on day t.
vertical_roll_n	The near two legs of the Vertical Roll volume associated with the option series on day t
vertical_roll_f	The far two legs of the Vertical Roll volume associated with the option series on day t
straddle_roll	The Straddle (Strangle) Roll volume associated with the option series on day t.
straddle_roll_n	The near two legs of the Straddle (Strangle) Roll volume associated with the option series on day t.
straddle_roll_f	The far two legs of the Straddle (Strangle) Roll volume associated with the option series on day t.

Appendix B. Algorithm to Infer Complex Strategies from the Trades Reported in the OPRA Data

This appendix describes the algorithm we use to infer complex strategies from the trades reported in the OPRA data. We begin by describing a manual exercise we carried out, and then describe the automatic algorithm.

A. Manual Exercise

In this section, we outline the process of manually grouping and identifying complex trades based on a sample data set, which was used to calibrate a systematic algorithm for efficient and accurate analysis of larger sets of complex trade records. The sample data includes all the complex trades (except for the stock-option trades) for the SPY option class on October 18, 2016, including 10,657 (423,609 contracts) *L* trades, 667 (7,906 contracts) *M* trades, and 1,972 (64,537 contracts) *Q* trades. The SPY option class was selected as the sample data set for this exercise because it is the most liquid option class, with a large number of trades and high contract volume, making it well representative of diverse market participants and complex strategies.

The manual grouping and matching of complex trades was performed using the following principles. First, complex trade records with identical or nearly identical timestamps, executed on the same options exchange, and containing the same trade message code, were placed in the same group. The difference between the maximum and minimum timestamps of the trades in the same group should be no more than a few milliseconds. Table B2 provides an illustration of the grouping process using a sample set of complex trades shown in Table B1. Within the same group, complex trade records possess the same timestamps, except for the 25th and 26th trades, which have a difference of 1 millisecond. In total, the 32 complex trades executed on six exchanges were grouped into 13 groups. It is important to note that while the 5th, 6th, 7th, 8th, and 9th trades have the same timestamp and are executed on the same exchange, the 8th and 9th trades were placed in separate groups because they have different trade message codes.

After the trades are grouped, the next step is to match the records that appear to form a complex package in each group. Table 1 provides a list of commonly used complex strategies, along with the number of legs associated with each strategy and how to set it up. To give an example, for the first group of trades in our sample in Table B2, it is clear that the two legs are both call options, with the same strike price but different expiration dates. Thus, these two legs form a calendar call spread. For the two trades belonging to the second group, as they have the same expiration date but different strike prices, one being a call and the other being a put, and the condition code is set to "*M*", they form a strangle. The trades in the seventh and eighth groups appear to form strangles, but there is a possibility that they may actually form combo spreads (long/short call positions and short/long call positions to mimic a long or short position of the underlying) as they are executed on ARCA with the trade message code set to "*L*."[10] The four trades belonging to the twelfth group may form an unbalanced iron condor spread, as they have two calls and two puts with the same expiration date but different strike prices, where the strike prices of the puts are lower than that of the calls, and the difference between the strike prices of the puts is not equal to the difference between the strike prices of the calls.

Table B3 presents the result of the manual matching of the 13 groups of complex trades previously grouped in Table B2. The final manual identifications resulted in the identification of 1 calendar call spread that had 8 contracts, 1 strangle spread with 2 contracts, 1 butterfly put spread totaling 20 contracts, 2 diagonal put spreads with 34 contracts, 3 ladder put spreads with a total of 174 contracts, 2 vertical put spreads with a total of 12 contracts, 1 unbalanced iron condor spread with 20 contracts, and two trades that were either strangles or combo spreads with a total of 14 contracts.

It is important to keep in mind that there may be instances where a customer's multi-leg order is filled by multiple liquidity providers simultaneously, such as in a price improvement auction, or is traded against single-leg resting limit orders or market maker quotes of individual option series, which can result in multiple "footprints" for a single leg. Therefore, sometimes it is optimal to combine trade records that belong to the same option series before attempting to match the legs and identify the strategy of a complex trade. Three examples of these scenarios are provided in Table B4. In Panel A of Table B4, the first and second trades belong to the same option series, and when combined, they constitute one leg of a vertical Put spread, with the third trade being the other leg within the same trade package. In Panel B of Table B4, each leg of a diagonal call spread is evenly split into two trades with equal trade sizes. In Panel C of Table B4, each leg of a vertical put spread consists of two trades with unequal trade sizes. Therefore, in

[10] We mentioned in the main text that the trade message codes for all complex trades are set to be "*L*", and therefore whenever we see a call and a put with the same expiration in a group, we cannot tell whether the two legs constitute a straddle (strangle) or a combo.

these cases, it is necessary to combine the smaller trades of the same option series into a single leg trade before identifying the strategy for the given complex trade.

However, it is also possible that two separate complex packages are executed at the same exchange almost simultaneously with similar or very close timestamps and the same message code reported in the OPRA data. This could happen if both packages contain legs from the same option series. In such cases, it would not make sense to combine trades within the same series before identifying the complex strategy. An example of this is shown in Table B5. From the table, it can be seen that eight SPY options complex trades were executed within a one-millisecond window between 9:30:00.924 AM and 9:30:00.925 AM on ISE. Additionally, the second and sixth trades belong to the same call option series with a strike of 219 and expiration date of 2016/11/18. If we were to combine these two trades into one leg, we might mistakenly consider the group of trades as the completion of a seven-leg unknown complex strategy, as shown in panel C of the table. However, by taking into account the possibility of multiple packages existing for this group and not combining the trades, we can see that the first four trades make up an iron condor strategy and the last four trades make up a straddle (strangle) Roll strategy.

In some rare situations, trade records may contain more than 4 distinct options series, indicating that a single market participant executed a complex package with more than 4 legs or multiple complex packages were executed by different market participants simultaneously. Unfortunately, the OPRA data does not have the information needed to resolve this issue with certainty. To address this, a simple rule is proposed. When there are 4 or more unique legs within the same group of trade records, a ratio is calculated. The ratio is defined as the maximum leg contract volume divided by the minimum leg contract volume. If the ratio is below 3, the trades are considered to be one complex package with more than 4 legs. If the ratio is greater than 3, it is assumed that the group contains multiple packages and each is identified.[11] Table B6 provides two examples to illustrate this process. In Panel A, there are 8 trades with the maximum leg contract volume being 40 and the minimum leg contract volume being 1, leading to the conclusion that there are multiple complex packages. The trades are then identified as two vertical calls and two vertical puts. In Panel B, 8 leg trades are very close in terms of timestamps, but the ratio of 2 suggests that we are treating it as a complex package with 8 legs.

[11] CBOE specifies that the ratio needs to be between 1/3 and 3 for electronic complex orders.

Additionally, there may be a small number of trades, particularly large ones that are not matched in the previous steps of the process. This could occur due to technical issues with the complex order book matching process that take longer than expected, causing a leg of the complex order to be filled or reported later than the other legs. The discrepancy in execution time for different legs of the same package could also be caused by manual processes when floor market makers report the trades via handheld device or handwritten ticket. To address this issue, we have relaxed the time tolerance threshold for grouping trade records. Trades can still be considered as belonging to the same group, even if they are tens of milliseconds or even seconds apart, as long as they are likely part of a common complex strategy, such as a vertical, calendar, butterfly, etc.

Table B7 provides three examples, one in Panel A, one in Panel B and one in Panel C. In the first example, each leg of the trade has a size of 700 contracts and the two legs are 155 milliseconds apart. In the second example, each leg of the trade has a size of 5 contracts and the two legs are 13 milliseconds apart. The strategies in both cases are diagonal puts. Sometimes, it may be necessary to compare unmatched trade records with adjacent trade records, even if they have already been matched. An example of this is provided in Panel C. Initially, one might think that the first, second, and third trades comprise a diagonal ratio spread and that the fourth trade was left unmatched because of a 28 millisecond gap between its timestamp and the timestamps of the first three trades. After careful consideration, it is reasonable to believe that the four trades together comprise a diagonal spread, with one trade of the far leg being reported late to the consolidated tape for unknown reasons.

After a thorough examination of all 13,296 complex trades, we grouped and identified them based on various scenarios as well as the corner cases as seen in Tables B3 to B7. The results are summarized in Table B8, which shows the number of trades, number of packages identified, and total contract volume for each complex strategy category. The most popular strategy, as measured by contract volume, was the vertical put spread, with 2,542 trades, 1,221 packages, and 121,582 contracts. The second most common strategy was the vertical call spread, with 2,291 trades, 1,090 packages, and 61,752 contracts.

The key to creating a computerized algorithm to mimic the manual process of identifying complex trades is finding an efficient way to group the trade records in the first place. To achieve this, we need to find a way to characterize the "closeness" of the time interval when measured by the difference between the timestamps of the legs that belong to the same package and the average time interval between adjacent packages. To start, we create two time interval measures. The first one is "*Time1*", which is the difference between the maximum and minimum timestamps of the trade records within the same package. The other interval is "*Time2*", which represents the difference between the minimum timestamp in the package and the maximum timestamp from the previous package executed on the same exchange with the same trade message code.

Panels A and B of Table B9 summarize the first and second interval measure respectively. From Panel A, we can tell that the average timestamp difference between different legs within the same package is less than 1 millisecond for all the exchanges, except CBOE, which an upward mean due to a “large” gap observed in one package executed on the floor. At the 99th percentile, the difference is less than or equal to 3 milliseconds, except for CBOE, which is around 4 milliseconds. Panel B reports the time interval between adjacent complex packages and shows that the average interval is greater than 22 seconds across all exchanges.

B. Automated Approach

An efficient automated process for matching and identifying complex trades should aim to reduce false positives, i.e., to minimize the number of incorrectly matched trades. Given what has been learned in the manual exercise, to achieve this goal, the time tolerance threshold used to group trades has to be set in such a way that trades that are too “far apart” from each other as measured by timestamps are not placed in the same group. At the same time, the algorithm also needs to be flexible enough to be able to handle the following scenarios: (1) group floor multi-leg trades which tend to have much wider gaps measured by the difference between the timestamps across the legs, and (2) separate multiple packages that are executed very close with each other in terms of timestamp. Therefore, a multi-step procedure with varying time tolerance levels is needed to group the trades. The first step involves filtering trades with modest tolerance thresholds and grouping them to identify complex strategies for groups containing 4 or fewer legs. The second step involves regrouping some trades with stricter time thresholds and identifying complex strategies for groups that may contain multiple packages. Finally, the third step involves relaxing the time thresholds from step 1 to group and identify any remaining unmatched trades.

C. First Step of the Automated Approach

This section introduces a very simple group filtering process to bucket adjacent trade records. The technique involves sweeping through all the complex trade records executed on the same options exchange with the same condition code on the same day for the same options class, and bucketing trades into different groups to be considered for possible matching in the later process based on the execution timestamps. Suppose we already have a group of trades, to determine whether the next successive trade with timestamp T can join the current group for possible matching requires two predetermined time interval thresholds (or time interval tolerances): *Threshold1* and *Threshold2* with *Threshold2* less than or equal to *Threshold1*, both in the magnitude of milliseconds, and the following variables:

1. *Base_trade_time*: the minimum timestamp of the trades in the current group.
2. *Pre_trade_time*: the maximum timestamp of the trades in the current group.
3. *Base_diff*: $T - Base\_trade\_time$.
4. *Delta_diff*: $T - Pre\_trade\_time$.

The principle of the grouping filter process mandates that trades within the same group must meet either one of the following two conditions: (1) the last trade in a given group must have been executed within *Threshold1* of the earliest trade within the same group, and (2) any two successive trades within the same group must have been executed within Threshold2 of each other.

To begin, we consider the first two trades with timestamps t_1 and t_2 respectively. We then set both *Base_trade_time* and *Pre_trade_time* to be t_1, and both *Base_diff* and *Delta_diff* to be $t_2 - t_1$. Next, we compare *Base_diff* to *Threshold1* and compare *Delta_diff* to *Threshold2*. If either $Base\_diff <= Threshold1$ or $Delta\_diff <= Threshold2$ holds, then the second trade belongs to the same group with the first trade, and *Pre_trade_time* is then set to be t_2 while *Base_diff* remains to be t_1. We then move to the third trade with the timestamp equal to t_3. By definition, now *Base_diff* is set to be $t_3 - t_1$, and *Delta_diff* is set to be $t_3 - t_2$. If either $Base\_diff \leq Threshold1$ or $Delta\_diff \leq Threshold2$ holds, then the third trade belongs to the group comprised of the first trade and second trade. The algorithm iterates until the nth trade with the timestamp equal to t_n, and it happens that both $t_n - t_1 > Threshold1$ and $t_n - t_{n-1} > Threshold2$ hold, and therefore the nth trade does not belong to the group comprised of the previous $n - 1$ trades. Then we reset *Base_trade_time* and *Pre_trade_time* to be t_n, and the algorithm continues until the last complex trade record of the day for the same option class for a given exchange and trade message code.

Apparently, the smaller the *Threshold1* or *Threshold2*, the more groups we derive from the same set of complex trades. Clearly, if both *Threshold1* and *Threshold2* are set to be zero, then complex trades only executed on the same millisecond-level timestamp on the same exchange with the same trade condition code can be connected to form a group.

Next, we use an example to illustrate the group filtering technique, and in particular, to show how the outcomes may vary and depend on the choices of the two sets of time threshold parameters. The example consists of five SPY options trades executed on CBOE over a four-millisecond time interval ranging from 9:30:02.678 AM to 9:30:02.682 AM as shown in Panel A of Table B10. If we set *Threshold1* to be 4 milliseconds and *Threshold2* to be 3 milliseconds, given that the second, third, and fourth trades are executed within 4 milliseconds of the first trade and the fifth trade is executed within 3 milliseconds of the fourth trade, all five trades belong to the same group as illustrated in Panel B. When we set both *Threshold1* and *Threshold2* to be 1 millisecond instead, clearly the fourth and fifth trades form a different group from the group formed by the first three trades since the gap between the third trade execution time and fourth trade execution time is 2 milliseconds which is greater than the time tolerance thresholds. The grouping based on the second set of parameters are illustrated in panel C of Table B10.

We need to strike a balance when setting the time threshold parameters in this step. At first, we don't have to start with either super sensitive parameters (i.e., both *Threshold1* and *Threshold2* set to be 1 millisecond) to take care of the scenarios in which multiple packages are executed extremely closely with each other, or somewhat non-sensitive parameters (i.e., both *Threshold1* and *Threshold2* set to be 1 second) to take care of the scenarios in which different legs within the same package are reported with a wide time gap. The two types of scenarios are rare and can be addressed in the later steps of the process. Therefore, we set *Threshold1* and *Threshold2* in the first step to be 3 milliseconds and 2 milliseconds respectively.

Next, we combine the trades that belong to the same option series within each group. This is necessary because a complex trade could be executed in smaller pieces and result in multiple trade records. To handle this scenario, the matching algorithm is based on the legs of the complex trade rather than individual trades.

To ensure the validity of the grouping, we create a variable called "ratio" to track the ratio of the maximum leg volume to the minimum leg volume for each group. The groups with two legs, three legs, and four legs are kept, while the four-leg groups must have a "ratio" of less than

3. The matching and identification of the remaining complex trade records for groups with five or more legs, four-leg groups with a ratio greater than or equal to 3, and groups with only one leg will be discussed in later steps.

The process of identifying complex trade strategies for two-leg, three-leg, and four-leg bundles can be achieved by comparing the characteristics of each leg within the same group, including the type of contract (call or put), expiration dates, strike prices, and execution volumes. For a group with two legs, the comparison starts by examining the type of contract, i.e., both legs being calls or both legs being puts. If both legs are calls, the strike prices, expiration dates, and contract volumes are compared. If the strike prices are different and the expiration dates and contract volumes are the same, the two legs form a vertical call. Similar logic could be applied to infer whether the complex instrument belongs to other two-leg options such as calendar and diagonal, or three-leg and four-leg options such as butterfly and iron condor.

D. Second Step of the Automated Approach

There are some groups that contain four or more legs with *ratio* greater than three. It is possible that each group may actually contain multiple complex packages, i.e., executions of different complex orders are too close to be separated using the filtering process in step 1. By using a set of stricter set of parameters (e.g., setting both *Threshold1* and *Threshold2* to be 1 millisecond) than the one used in step 1, it may be possible to separate and regroup such complex trades. However, it is important to note that this method is not a complete solution as there may still be instances where multiple complex orders are executed at the same millisecond-level timestamp. In these cases, additional techniques are required to accurately identify and distinguish between different packages.

We propose two sub-steps. In step 2-1, we first decompose any "legs" that are previously combined from trades for the same series within the same group. We then regroup the trades which belong the groups derived in the step 1 with four or more than four legs and with ratio greater than 3 using a different set of parameters: both *Threshold1* and *Threshold2* set to 1 millisecond. After this exercise, some of the previously four-leg and *n*-leg (n>4) groups are decomposed into groups with smaller number of legs, allowing us to use the same technique described in step 1 to match up the legs and identify the complex strategies. As can be seen from the example in Panel A of Table B11, the thresholds set in step 1 placed the five trades into one group, but the stricter thresholds set in this step regrouped the trades into 2 groups with the first group consisting of an uneven butterfly call spread and the second group consisting of a diagonal call spread. Of course, there are still some groups which still contain four legs and more than four legs with ratio greater than 3, in which case we are going to use a pecking order approach as illustrated later in step 2-2 to match up the legs and identify multiple packages within one group.

In step 2-2, we first decompose any "legs" that are combined from trades for the same options series within the same group in step 2-1. We also mandate that trades can only be matched up with each other if they are executed within 1 millisecond. The principle can be illustrated using the example in Panel B of Table B11. The first trade within the group which is executed at 9:30:30.066 cannot be matched up with the ninth trade which is executed at 9:30:30.068 to form a calendar ratio spread, but the fifth and sixth trades, which are executed at 9:30:30.067, can be matched up with the eleventh and twelfth trades, which are executed at 9:30:30.068, to form an iron condor spread.

The principle of the pecking order approach to match up trades when there are multiple packages within the same group that cannot be further broken out into smaller groups using the time thresholds is to scan for the more complicated patterns first such as iron condor, iron butterfly, condor, and butterfly, which consist of either three or four legs, before then identifying the simpler cases such as vertical, calendar, diagonal, straddle, and strangle for the previously unmatched records. Specifically, given such a group of trade records, trade records are matched up in sequence as follows:

Step 2-2-1: Identify four-leg spreads with two calls and two puts: iron Butterfly, iron condor, box, straddle roll, and strangle roll.

Step 2-2-2: Identify four-leg spreads with all calls or all puts: condor, vertical roll.

Step 2-2-3: Identify three-leg spreads with all calls or all puts: butterfly and ladder.

Step 2-2-4: Identify two-leg spreads: vertical, calendar, diagonal, straddle, strangle, combo, and various ratio spreads.

We again use the example in Panel B of Table B11 to illustrate the pecking order process. The sample consists of two trades with timestamps equal to 9:30:02.066, six trades with timestamps equal to 9:30:02.067, and four trades with timestamps equal to 9:30:02.068. Even though the difference between the maximum timestamp and minimum timestamp is 2 milliseconds, which is greater than *Threshold1*, since *Threshold2* is set to be 1 millisecond, the algorithm would place successive trades executed within 1 millisecond in the same group. We then use a segment of codes to scan the group of trades to infer whether any four trades would comprise an iron condor, iron butterfly, box, or straddle/strangle Roll spread first. The algorithm finds two iron condors. We then scan the rest of the trades for possible complex packages following step 2-2-1 through step 2-2-5. Finally, we find two additional packages with one being a vertical call and the other being a strangle or combo.

E. Third Step of the Automated Approach

In the last step, we focus on the trade records which are not matched up in the first step of the process. There are some groups that consist of only one leg, which could be due to the fact that the leg is part of a complex package, but when it is reported to the tape, their timestamps differ from the rest of the legs by larger than the time thresholds set in step 1.

To do this, we start by combining these trades with their closest adjacent trades, even if they had already been matched in the previous step. Next, we regroup these trades using the same algorithm described in step 1, but with less restrictive time tolerance thresholds, such as setting *Threshold1* to 5 seconds and *Threshold2* to 3 seconds. Finally, we use the same method described in step 1 to identify the complex packages based on the combination of option types, expiration dates, strike prices, and transaction sizes within each newly formed group.

F. Performance of the Algorithm

In Table B12, the results from the systematic and algorithmic approach (through step 1 to step 3) are compared to the results from the manual process. The time thresholds Threshold1 and Threshold2 are set to 3 milliseconds and 2 milliseconds in step 1, 1 millisecond and 1 millisecond in step 2-1, and 5 seconds and 3 seconds in step 3, respectively.

The results show that the algorithmic approach can effectively identify complex packages and match the manual process well. In step 1, the algorithm classified 92.28% of the volume (12,456 trades) accurately and matched 100% of the results from the manual process. In step 2-1, the algorithmic approach matched 82.26% of the trades (93.09% by volume) and only 1.40% of the trades (0.49% by volume) were regrouped. In step 2-2, the algorithmic approach matched 88.82% of the trades (93.52% by volume) and only 3.43% of the trades (2.33% by volume) were regrouped. Finally, in step 3, the algorithmic approach matched 100% of the results from the manual process for the remaining trades. These results suggest that the algorithmic approach can provide a reliable and efficient alternative to manual processes.

Table B1
Examples of Complex Trades

This table provides examples of 32 *M, L* and *Q* trades that were executed over a 31-second time interval: 11:01:24 AM through 11:01:55 AM on October 18, 2016 for the SPY options class across six options exchanges. Information regarding the characteristics of each of the leg trades and the timestamp of each trade at millisecond level granularity are also included.

Exchange	Class	Trade Date	Expiration	Put/Call	Strike	Price	Size	Timestamp	Condition
C	SPY	20161018	20161118	C	218	1.33	4	11:01:24.333	L
C	SPY	20161018	20161021	C	218	0.03	4	11:01:24.333	L
X	SPY	20161018	20161028	P	210	0.64	1	11:01:26.007	M
X	SPY	20161018	20161028	C	215	1.09	1	11:01:26.007	M
I	SPY	20161018	20161021	P	210	0.17	5	11:01:27.200	Q
I	SPY	20161018	20161021	P	212	0.48	10	11:01:27.200	Q
I	SPY	20161018	20161021	P	214	1.14	5	11:01:27.200	Q
I	SPY	20161018	20161021	P	214	1.13	8	11:01:27.200	L
I	SPY	20161018	20161019	P	214.5	0.98	8	11:01:27.200	L
A	SPY	20161018	20161021	P	209	0.1	8	11:01:27.488	L
A	SPY	20161018	20161021	P	214	1.14	8	11:01:27.488	L
A	SPY	20161018	20161021	P	212.5	0.61	8	11:01:27.488	L
N	SPY	20161018	20161021	P	212.5	0.61	5	11:01:27.490	L
N	SPY	20161018	20161021	P	209	0.1	5	11:01:27.490	L
N	SPY	20161018	20161021	P	214	1.14	5	11:01:27.490	L
N	SPY	20161018	20161021	P	210	0.18	3	11:01:28.727	L
N	SPY	20161018	20161021	C	217.5	0.05	3	11:01:28.727	L
B	SPY	20161018	20161028	P	210	0.64	4	11:01:29.176	L
B	SPY	20161018	20161028	C	215	1.09	4	11:01:29.176	L
N	SPY	20161018	20161021	P	212.5	0.61	45	11:01:30.543	L
N	SPY	20161018	20161021	P	209	0.1	45	11:01:30.543	L
N	SPY	20161018	20161021	P	214	1.14	45	11:01:30.543	L
I	SPY	20161018	20161028	P	218	4.35	9	11:01:39.137	L
I	SPY	20161018	20161021	P	219	5.21	9	11:01:39.137	L
C	SPY	20161018	20161021	P	209	0.1	5	11:01:39.137	L
C	SPY	20161018	20161021	P	211	0.3	5	11:01:39.138	L
I	SPY	20161018	20161028	C	222	0.02	5	11:01:54.091	Q
I	SPY	20161018	20161028	C	218	0.17	5	11:01:54.091	Q
I	SPY	20161018	20161028	P	204	0.13	5	11:01:54.091	Q
I	SPY	20161018	20161028	P	209	0.48	5	11:01:54.091	Q
I	SPY	20161018	20161021	P	209	0.1	1	11:01:54.157	L
I	SPY	20161018	20161021	P	214	1.13	1	11:01:54.157	L

Table B2
Manual Grouping of the complex trades in Table B1.
In this table, we manually group complex trades based on whether the trades were executed at identical or nearly identical timestamps, on the same options exchange and flagged with the same trade message code. We are able to derive 13 groups of complex trades from 32 execution records.

ID	Exchange	Class	Trade Date	Expiration	Put/Call	Strike	Price	Size	Timestamp	Condition	Group
1	C	SPY	20161018	20161118	C	218	1.33	4	11:01:24.333	L	1
2	C	SPY	20161018	20161021	C	218	0.03	4	11:01:24.333	L	1
3	X	SPY	20161018	20161028	P	210	0.64	1	11:01:26.007	M	2
4	X	SPY	20161018	20161028	C	215	1.09	1	11:01:26.007	M	2
5	I	SPY	20161018	20161021	P	210	0.17	5	11:01:27.200	Q	3
6	I	SPY	20161018	20161021	P	212	0.48	10	11:01:27.200	Q	3
7	I	SPY	20161018	20161021	P	214	1.14	5	11:01:27.200	Q	3
8	I	SPY	20161018	20161021	P	214	1.13	8	11:01:27.200	L	4
9	I	SPY	20161018	20161019	P	214.5	0.98	8	11:01:27.200	L	4
10	A	SPY	20161018	20161021	P	209	0.1	8	11:01:27.488	L	5
11	A	SPY	20161018	20161021	P	214	1.14	8	11:01:27.488	L	5
12	A	SPY	20161018	20161021	P	212.5	0.61	8	11:01:27.488	L	5
13	N	SPY	20161018	20161021	P	212.5	0.61	5	11:01:27.490	L	6
14	N	SPY	20161018	20161021	P	209	0.1	5	11:01:27.490	L	6
15	N	SPY	20161018	20161021	P	214	1.14	5	11:01:27.490	L	6
16	N	SPY	20161018	20161021	P	210	0.18	3	11:01:28.727	L	7
17	N	SPY	20161018	20161021	C	217.5	0.05	3	11:01:28.727	L	7
18	B	SPY	20161018	20161028	P	210	0.64	4	11:01:29.176	L	8
19	B	SPY	20161018	20161028	C	215	1.09	4	11:01:29.176	L	8
20	N	SPY	20161018	20161021	P	212.5	0.61	45	11:01:30.543	L	9
21	N	SPY	20161018	20161021	P	209	0.1	45	11:01:30.543	L	9
22	N	SPY	20161018	20161021	P	214	1.14	45	11:01:30.543	L	9
23	I	SPY	20161018	20161028	P	218	4.35	9	11:01:39.137	L	10
24	I	SPY	20161018	20161021	P	219	5.21	9	11:01:39.137	L	10
25	C	SPY	20161018	20161021	P	209	0.1	5	11:01:39.137	L	11
26	C	SPY	20161018	20161021	P	211	0.3	5	11:01:39.138	L	11
27	I	SPY	20161018	20161028	C	222	0.02	5	11:01:54.091	Q	12
28	I	SPY	20161018	20161028	C	218	0.17	5	11:01:54.091	Q	12
29	I	SPY	20161018	20161028	P	204	0.13	5	11:01:54.091	Q	12
30	I	SPY	20161018	20161028	P	209	0.48	5	11:01:54.091	Q	12
31	I	SPY	20161018	20161021	P	209	0.1	1	11:01:54.157	L	13
32	I	SPY	20161018	20161021	P	214	1.13	1	11:01:54.157	L	13

Table B3

Manually Identify the Complex Strategies from the Complex Trades Grouped in Table B2.

In this table, we identify the 13 groups of complex trades resulting in 1 calendar call, 1 strangle, 1 butterfly put, 2 diagonal put, 3 ladder put, 2 vertical put and 1 unbalanced iron condor.

ID	Exchange	Class	Trade Date	Expiration	Put/Call	Strike	Price	Size	Timestamp	Condition	Group	Spread
1	C	SPY	20161018	20161118	C	218	1.33	4	11:01:24.333	L	1	Calendar (Call)
2	C	SPY	20161018	20161021	C	218	0.03	4	11:01:24.333	L	1	Calendar (Call)
3	X	SPY	20161018	20161028	P	210	0.64	1	11:01:26.007	M	2	Strangle
4	X	SPY	20161018	20161028	C	215	1.09	1	11:01:26.007	M	2	Strangle
5	I	SPY	20161018	20161021	P	210	0.17	5	11:01:27.200	Q	3	Butterfly (Put)
6	I	SPY	20161018	20161021	P	212	0.48	10	11:01:27.200	Q	3	Butterfly (Put)
7	I	SPY	20161018	20161021	P	214	1.14	5	11:01:27.200	Q	3	Butterfly (Put)
8	I	SPY	20161018	20161021	P	214	1.13	8	11:01:27.200	L	4	Diagonal (Put)
9	I	SPY	20161018	20161019	P	214.5	0.98	8	11:01:27.200	L	4	Diagonal (Put)
10	A	SPY	20161018	20161021	P	209	0.1	8	11:01:27.488	L	5	Ladder (Put)
11	A	SPY	20161018	20161021	P	214	1.14	8	11:01:27.488	L	5	Ladder (Put)
12	A	SPY	20161018	20161021	P	212.5	0.61	8	11:01:27.488	L	5	Ladder (Put)
13	N	SPY	20161018	20161021	P	212.5	0.61	5	11:01:27.490	L	6	Ladder (Put)
14	N	SPY	20161018	20161021	P	209	0.1	5	11:01:27.490	L	6	Ladder (Put)
15	N	SPY	20161018	20161021	P	214	1.14	5	11:01:27.490	L	6	Ladder (Put)
16	N	SPY	20161018	20161021	P	210	0.18	3	11:01:28.727	L	7	Strangle (Combo)
17	N	SPY	20161018	20161021	C	217.5	0.05	3	11:01:28.727	L	7	Strangle (Combo)
18	B	SPY	20161018	20161028	P	210	0.64	4	11:01:29.176	L	8	Strangle (Combo)
19	B	SPY	20161018	20161028	C	215	1.09	4	11:01:29.176	L	8	Strangle (Combo)
20	N	SPY	20161018	20161021	P	212.5	0.61	45	11:01:30.543	L	9	Ladder (Put)
21	N	SPY	20161018	20161021	P	209	0.1	45	11:01:30.543	L	9	Ladder (Put)
22	N	SPY	20161018	20161021	P	214	1.14	45	11:01:30.543	L	9	Ladder (Put)
23	I	SPY	20161018	20161028	P	218	4.35	9	11:01:39.137	L	10	Diagonal (Put)
24	I	SPY	20161018	20161021	P	219	5.21	9	11:01:39.137	L	10	Diagonal (Put)
25	C	SPY	20161018	20161021	P	209	0.1	5	11:01:39.137	L	11	Vertical (Put)
26	C	SPY	20161018	20161021	P	211	0.3	5	11:01:39.138	L	11	Vertical (Put)
27	I	SPY	20161018	20161028	C	222	0.02	5	11:01:54.091	Q	12	Iron Condor (Other)
28	I	SPY	20161018	20161028	C	218	0.17	5	11:01:54.091	Q	12	Iron Condor (Other)
29	I	SPY	20161018	20161028	P	204	0.13	5	11:01:54.091	Q	12	Iron Condor (Other)
30	I	SPY	20161018	20161028	P	209	0.48	5	11:01:54.091	Q	12	Iron Condor (Other)
31	I	SPY	20161018	20161021	P	209	0.1	1	11:01:54.157	L	13	Vertical (Put)
32	I	SPY	20161018	20161021	P	214	1.13	1	11:01:54.157	L	13	Vertical (Put)

Table B4

Examples of Complex Trades Being "Shredded".

In this table, we present cases where the number of reported executions could be larger than the number of legs within a given complex order. In Panel A, the first and second trades belong to the same option series, and when combined, they constitute one leg of a vertical put, with the third trade. In Panel B, each leg of a diagonal call is evenly split into two executions. In Panel C, each leg of a vertical put consists of two executions with unequal trade sizes. The examples suggest that sometimes we may need to combine trades belonging to the same options series before matching the legs.

Panel A

ID	Exchange	Class	Trade Date	Expiration	Put/Call	Strike	Price	Size	Timestamp	Condition	Spread
1	N	SPY	20161018	20161028	P	211	1.01	91	10:23:28.694	L	Vertical (Put)
2	N	SPY	20161018	20161028	P	211	1.01	64	10:23:28.694	L	Vertical (Put)
3	N	SPY	20161018	20161028	P	206	0.26	155	10:23:28.694	L	Vertical (Put)

Panel B

ID	Exchange	Class	Trade Date	Expiration	Put/Call	Strike	Price	Size	Timestamp	Condition	Spread
1	N	SPY	20161018	20161028	C	219	0.14	1	9:30:01.901	L	Diagonal (Call)
2	N	SPY	20161018	20161021	C	216	0.32	1	9:30:01.901	L	Diagonal (Call)
3	N	SPY	20161018	20161021	C	216	0.32	1	9:30:01.901	L	Diagonal (Call)
4	N	SPY	20161018	20161028	C	219	0.14	1	9:30:01.901	L	Diagonal (Call)

Panel C

ID	Exchange	Class	Trade Date	Expiration	Put/Call	Strike	Price	Size	Timestamp	Condition	Spread
1	N	SPY	20161018	20161019	P	211.5	0.13	34	9:39:17.930	L	Vertical (Put)
2	N	SPY	20161018	20161019	P	211.5	0.13	103	9:39:17.930	L	Vertical (Put)
3	N	SPY	20161018	20161019	P	213	0.36	3	9:39:17.930	L	Vertical (Put)
4	N	SPY	20161018	20161019	P	213	0.36	134	9:39:17.930	L	Vertical (Put)

Table B5

An Example to Illustrate a Scenario Where Combining the Trades is not Optimal.

This table provides an example to illustrate that it is not always optimal to combine executions belonging to the same option series before matching the legs to identify the complex strategy. Panel A shows a set of 8 *Q* trades executed on ISE over a one millisecond interval. The second and the sixth trades belong to the same options series. Panel B shows the matching result if we don't combine the second and sixth trades before matching the legs. Panel C shows the matching result if otherwise.

Panel A

ID	Exchange	Class	Trade Date	Expiration	Put/ Call	Strike	Price	Size	Timestamp	Condition
1	I	SPY	20161118	20161118	C	221	0.55	10	9:30:00.924	Q
2	I	SPY	20161118	20161118	C	219	1.1	10	9:30:00.924	Q
3	I	SPY	20161118	20161118	P	202	0.85	10	9:30:00.924	Q
4	I	SPY	20161118	20161118	P	204	1.04	10	9:30:00.924	Q
5	I	SPY	20161021	20161021	C	219	0.03	2	9:30:00.924	Q
6	I	SPY	20161118	20161118	C	219	1.08	2	9:30:00.924	Q
7	I	SPY	20161021	20161021	P	213	0.67	2	9:30:00.925	Q
8	I	SPY	20161118	20161118	P	213	2.95	2	9:30:00.925	Q

Panel B

ID	Exchange	Class	Trade Date	Expiration	Put/ Call	Strike	Price	Size	Timestamp	Condition	Spread
1	I	SPY	20161118	20161118	C	221	0.55	10	9:30:00.924	Q	Iron Condor
2	I	SPY	20161118	20161118	C	219	1.1	10	9:30:00.924	Q	Iron Condor
3	I	SPY	20161118	20161118	P	202	0.85	10	9:30:00.924	Q	Iron Condor
4	I	SPY	20161118	20161118	P	204	1.04	10	9:30:00.924	Q	Iron Condor
5	I	SPY	20161021	20161021	C	219	0.03	2	9:30:00.924	Q	Straddle (Strangle) Roll
6	I	SPY	20161118	20161118	C	219	1.08	2	9:30:00.924	Q	Straddle (Strangle) Roll
7	I	SPY	20161021	20161021	P	213	0.67	2	9:30:00.925	Q	Straddle (Strangle) Roll
8	I	SPY	20161118	20161118	P	213	2.95	2	9:30:00.925	Q	Straddle (Strangle) Roll

Panel C

ID	Exchange	Class	Trade Date	Expiration	Put/ Call	Strike	Price	Size	Timestamp	Condition	Spread
1	I	SPY	20161118	20161118	C	221	0.55	10	9:30:00.924	Q	Unknown 7-leg
2	I	SPY	20161118	20161118	C	219	1.097	12	9:30:00.924	Q	Unknown 7-leg
3	I	SPY	20161118	20161118	P	202	0.85	10	9:30:00.924	Q	Unknown 7-leg
4	I	SPY	20161118	20161118	P	204	1.04	10	9:30:00.924	Q	Unknown 7-leg
5	I	SPY	20161021	20161021	C	219	0.03	2	9:30:00.924	Q	Unknown 7-leg
7	I	SPY	20161021	20161021	P	213	0.67	2	9:30:00.925	Q	Unknown 7-leg
8	I	SPY	20161118	20161118	P	213	2.95	2	9:30:00.925	Q	Unknown 7-leg

Table B6

Examples of the Simple Rule Based on the Ratio

In this table, we provide two examples to illustrate the ratio rule. We first calculate the ratio of each group of complex trades as the maximum leg volume divided by the minimum leg volume. If the ratio is less than or equal to (greater than) 3, we assume that this group contains one (multiple) complex instruments. In Panel A, there are 8 trades which consist of two vertical calls and vertical puts. In Panel B, the 8 trades are treated as a single complex package.

Panel A

ID	Exchange	Class	Trade Date	Expiration	Put/Call	Strike	Price	Size	Timestamp	Condition	Spread
1	I	SPY	20161018	20161102	C	217	0.74	1	9:30:01.132	L	Vertical (Call)
2	I	SPY	20161018	20161102	C	216	1.12	1	9:30:01.132	L	Vertical (Call)
3	I	SPY	20161018	20161104	C	220	0.18	10	9:30:01.133	L	Vertical (Call)
4	I	SPY	20161018	20161104	C	217	0.86	10	9:30:01.133	L	Vertical (Call)
5	I	SPY	20161018	20161021	P	203	0.01	1	9:30:01.135	L	Vertical (Put)
6	I	SPY	20161018	20161021	P	205	0.03	1	9:30:01.135	L	Vertical (Put)
7	I	SPY	20161018	20161021	P	206	0.02	40	9:30:01.135	L	Vertical (Put)
8	I	SPY	20161018	20161021	P	210	0.16	40	9:30:01.135	L	Vertical (Put)

Panel B

ID	Exchange	Class	Trade Date	Expiration	Put/Call	Strike	Price	Size	Timestamp	Condition	Spread
1	C	SPY	20161018	20161021	C	215	0.54	2	14:30:00.085	L	Eight-leg
2	C	SPY	20161018	20161021	C	217	0.08	2	14:30:00.085	L	Eight-leg
3	C	SPY	20161018	20161021	C	216	0.22	2	14:30:00.086	L	Eight-leg
4	C	SPY	20161018	20161021	C	218	0.03	2	14:30:00.086	L	Eight-leg
5	C	SPY	20161018	20161028	C	223	0.02	1	14:30:00.086	L	Eight-leg
6	C	SPY	20161018	20161028	C	227	0.01	1	14:30:00.086	L	Eight-leg
7	C	SPY	20161018	20161028	P	205	0.12	1	14:30:00.086	L	Eight-leg
8	C	SPY	20161018	20161028	P	209	0.4	1	14:30:00.086	L	Eight-leg

Table B7

Discrepancies in Execution Time for Different Legs within the Same Instrument.

In this table, we provide three examples in which the execution timestamps for different legs are not identical or nearly identical. In Panel A, the two legs are 155 milliseconds apart. In Panel B, the two leg are 15 milliseconds apart. In Panel C, the first three legs have the same timestamp and are 28 milliseconds away from the fourth leg. The examples illustrate the importance of comparing some of the unmatched trade records with adjacent ones (whether matched or not) to examine whether they can form a complex trade or not.

Panel A

ID	Exchange	Class	Trade Date	Expiration	Put/Call	Strike	Price	Size	Timestamp	Condition	Spread
1	C	SPY	20161118	20161104	P	218	5.57	700	10:45:01.596	L	Diagonal (Put)
2	C	SPY	20161216	20161021	P	212	4.64	700	10:45:01.751	L	Diagonal (Put)

Panel B

ID	Exchange	Class	Trade Date	Expiration	Put/Call	Strike	Price	Size	Timestamp	Condition	Spread
1	C	SPY	20161018	20161104	P	215	2.86	5	10:49:22.633	L	Diagonal (Put)
2	C	SPY	20161018	20161021	P	216	2.56	5	10:48:22.646	L	Diagonal (Put)

Panel C

ID	Exchange	Class	Trade Date	Expiration	Put/Call	Strike	Price	Size	Timestamp	Condition	Spread (step1)	Spread (Step2)
1	C	SPY	20161018	20161216	P	204	2.54	1200	10:29:37.302	L	Diagonal (Call|Ratio)	Diagonal (Call)
2	C	SPY	20161018	20161118	P	205	1.34	800	10:29:37.302	L	Diagonal (Call|Ratio)	Diagonal (Call)
3	C	SPY	20161018	20161118	P	205	1.34	1200	10:29:37.302	L	Diagonal (Call|Ratio)	Diagonal (Call)
4	C	SPY	20161018	20161216	P	204	2.54	800	10:29:37.430	L	Unmatched	Diagonal (Call)

Table B8
Summary of Complex Strategies from the Manual Exercise
This table summarizes the results of the manual identification of complex strategies in a sample of SPY options multi-leg trades executed on 2016/10/18. It includes the number of trades, number of complex packages, and the total contract volume for each identified strategy.

Number of Legs	Strategy	Number of Trades	Number of Packages	Total Contract Volume
2	Calendar (Call)	279	134	5732
2	Calendar (Call|Ratio)	17	7	124
2	Calendar (Put)	448	204	7714
2	Calendar (Put|Ratio)	4	2	3011
2	Combo	11	5	25
2	Diagonal (Call)	684	326	11554
2	Diagonal (Call|Ratio)	207	83	20651
2	Diagonal (Put)	577	277	20486
2	Diagonal (Put|Ratio)	21	10	15200
2	Other Straddle/Strangle	38	19	710
2	Straddle	284	137	5142
2	Straddle/Combo	370	182	21963
2	Straddle/Strangle Roll	132	33	676
2	Strangle	290	141	1888
2	Vertical (Call)	2291	1090	61752
2	Vertical (Call|Ratio)	66	30	6774
2	Vertical (Put)	2542	1221	121582
2	Vertical (Put|Ratio)	351	151	33228
3	Butterfly (Call)	64	21	1560
3	Butterfly (Put)	160	51	41600
3	Ladder (Call)	63	21	444
3	Ladder (Put)	81	26	1878
3	Other Butterfly	375	121	28928
4	Box	4	1	100
4	Condor (Call)	32	8	204
4	Condor (Put)	32	8	224
4	Iron Butterfly	226	56	1616
4	Iron Condor	1021	252	7840
4	Other Box/IC	748	186	2790
4	Other Condor	130	33	628
4	Vertical Roll (Call)	200	50	2844
4	Vertical Roll (Put)	92	25	8956
5+	5 or more Leg	244	56	587
3 or 4	All Other 3/4-Leg	1376	336	58482
	Unmatched	192	192	1007

Table B9
Summary Statistics of the Interval Measures for Complex Packages.
This table provides summary statistics for two measures related to the time interval measures of complex trades within a given exchange. Panel A reports statistics for the first interval measure, defined as the difference between the maximum and minimum timestamps of trades within a complex package. Panel B reports statistics for the time interval between adjacent complex packages for each exchange. The statistics are based on the sample of SPY multi-leg options trades in Table B8.

Panel A

Exchange	Number of Packages	min	p1	p10	p25	p50	mean	p75	p90	p99	max	Stdv
AMEX	183	0	0	0	0	0	0.00016	0	0.001	0.002	0.003	0.00
BOX	66	0	0	0	0	0	0.00014	0	0	0.003	0.003	0.00
CBOE	2135	0	0	0	0	0	0.0028	0	0.001	0.004	4.479	0.10
ISE	3353	0	0	0	0	0	0.00093	0	0	0.002	0.516	0.01
ARCA	576	0	0	0	0	0	0.00016	0	0.001	0.002	0.009	0.00
C2	420	0	0	0	0	0	0.0002	0	0.001	0.001	0.005	0.00
PHLX	1022	0	0	0	0	0	6.16E-05	0	0	0.001	0.02	0.00

Panel B

Exchange	Number of Packages	min	p1	p10	p25	p50	mean	p75	p90	p99	max	Stdv
AMEX	181	0	0	0.087	4.12	83.52	262.56	313.04	707.63	1994.83	3723.97	452.33
BOX	64	0	0	0.206	49.80	351.56	733.74	1038.35	2128.53	5011.63	5011.63	944.57
CBOE	2133	0	0.001	0.199	2.61	9.88	22.33	28.35	57.73	163.39	336.39	33.26
ISE	3347	0	0	0.015	1.10	9.63	42.48	36.48	95.70	534.03	3642.90	131.49
ARCA	574	0	0	0.032	2.52	29.28	82.97	94.57	222.33	711.34	2139.66	173.91
C2	418	0	0	0.134	6.08	50.31	112.69	153.79	319.83	780.44	1259.99	168.61
PHLX	1016	0	0	0.117	2.73	19.91	110.01	61.62	172.79	1340.39	10933.96	539.88

Table B10

Different Time Threshold Parameters May Lead to Different Groupings

This table illustrate how outcomes may vary and depend on the choice of different time threshold parameters. Panel A provides an example consisting of 5 trades over a four millisecond interval. Panel B shows the grouping result if we set *Threshold1* to be 4 milliseconds and *Threshold2* to be 3 milliseconds. Panel C shows the grouping result if we set both *Threshold1* and *Threshold2* to be 1 millisecond instead.

Panel A

ID	Exchange	Class	Trade Date	Expiration	Put/Call	Strike	Price	Size	Timestamp	Condition
1	C	SPY	20161018	20161118	C	211	5.74	1	9:30:02.678	L
2	C	SPY	20161018	20161118	C	214	3.66	2	9:30:02.678	L
3	C	SPY	20161018	20161118	C	217	1.98	1	9:30:02.679	L
4	C	SPY	20161018	20161021	C	207	7.3	4	9:30:02.681	L
5	C	SPY	20161018	20161021	C	210	4.45	4	9:30:02.682	L

Panel B

ID	Exchange	Class	Trade Date	Expiration	Put/Call	Strike	Price	Size	Timestamp	Condition	Base Trade Time	Pre Trade Time	Base Diff	Delta Diff	Group
1	C	SPY	20161018	20161118	C	211	5.74	1	9:30:02.678	L	9:30:02.678				1
2	C	SPY	20161018	20161118	C	214	3.66	2	9:30:02.678	L	9:30:02.678	9:30:02.678	0	0	1
3	C	SPY	20161018	20161118	C	217	1.98	1	9:30:02.679	L	9:30:02.678	9:30:02.678	0.001	0.001	1
4	C	SPY	20161018	20161021	C	207	7.3	4	9:30:02.681	L	9:30:02.678	9:30:02.679	0.003	0.001	1
5	C	SPY	20161018	20161021	C	210	4.45	4	9:30:02.682	L	9:30:02.678	9:30:02.681	0.004	0.001	1

Panel C

ID	Exchange	Class	Trade Date	Expiration	Put/Call	Strike	Price	Size	Timestamp	Condition	Base Trade Time	Pre Trade Time	Base Diff	Delta Diff	Group
1	C	SPY	20161018	20161118	C	211	5.74	1	9:30:02.678	L	9:30:02.678				1
2	C	SPY	20161018	20161118	C	214	3.66	2	9:30:02.678	L	9:30:02.678	9:30:02.678	0	0	1
3	C	SPY	20161018	20161118	C	217	1.98	1	9:30:02.679	L	9:30:02.678	9:30:02.678	0.001	0.001	1
4	C	SPY	20161018	20161021	C	207	7.3	4	9:30:02.681	L	9:30:02.681				2
5	C	SPY	20161018	20161021	C	210	4.45	4	9:30:02.682	L	9:30:02.681	9:30:02.681	0	0.001	2

Table B11

Examples Illustrating the Grouping and Matching of Complex Trades in Step 2.

In Panel A, we show how to regroup and match the complex trades which are previously identified as an unknown 5-leg complex instrument in the first step of the algorithm by imposing finer time threshold parameters. In Panel B, we illustrate the pecking order approach to match trades (greater than 4) with identical or almost identical execution timestamps.

Panel A

ID	Exchange	Class	Trade Date	Expiration	Put/Call	Strike	Price	Size	Timestamp	Condition	Spread	Group (Step1)	Group (step2-1)
1	C	SPY	20161018	20161021	C	214	1.12	1	11:31:49.928	L	Butterfly (Other)	1	1
2	C	SPY	20161018	20161021	C	214.5	0.84	2	11:31:49.928	L	Butterfly (Other)	1	1
3	C	SPY	20161018	20161021	C	217	0.11	1	11:31:49.928	L	Butterfly (Other)	1	1
4	C	SPY	20161018	20161021	C	221	0.01	20	11:31:49.931	L	Diagonal (Call)	1	2
5	C	SPY	20161018	20161118	C	220	0.73	20	11:31:49.932	L	Diagonal (Call)	1	2

Panel B

ID	Exchange	Class	Trade Date	Expiration	Put/Call	Strike	Price	Size	Timestamp	Condition	Spread	Step	Spread Number
1	A	SPY	20161018	20161118	C	220	0.8	1	9:30:02.066	L	Strangle/Combo	2.2.5	4
2	A	SPY	20161018	20161118	P	205	1.2	1	9:30:02.066	L	Strangle/Combo	2.2.5	4
3	A	SPY	20161018	20161111	C	222.5	0.18	1	9:30:02.067	L	Iron Condor	2.2.1	1
4	A	SPY	20161018	20161111	C	227	0.01	1	9:30:02.067	L	Iron Condor	2.2.1	1
5	A	SPY	20161018	20161118	C	219	1.11	10	9:30:02.067	L	Iron Condor	2.2.1	2
6	A	SPY	20161018	20161118	C	222	0.36	10	9:30:02.067	L	Iron Condor	2.2.1	2
7	A	SPY	20161018	20161111	P	197	0.3	1	9:30:02.067	L	Iron Condor	2.2.1	1
8	A	SPY	20161018	20161111	P	202.5	0.67	1	9:30:02.067	L	Iron Condor	2.2.1	1
9	A	SPY	20161018	20170120	C	220	2.61	3	9:30:02.068	L	Vertical (Call)	2.2.5	3
10	A	SPY	20161018	20170120	C	221	2.24	3	9:30:02.068	L	Vertical (Call)	2.2.5	3
11	A	SPY	20161018	20161118	P	200	0.67	10	9:30:02.068	L	Iron Condor	2.2.1	2
12	A	SPY	20161018	20161118	P	203	0.92	10	9:30:02.068	L	Iron Condor	2.2.1	2

Table B12

Comparison of the Algorithmic Approach to the Results of the Manual Exercise

This table compares the results of the algorithmic approach with the results of the manual exercise on a sample of SPY options multi-leg trades executed on 2016/10/18. The table provides the number of trades and the total contract volume at each step of the algorithmic approach. The "Trades Matched" column and the "Volume Matched" column show the results of the comparison.

Step	# of Trades	Fraction of Trades	Trades matched	Volume	Fraction of Volume	Volume matched
1	12456	93.68%	100.00%	457732	92.28%	100.00%
2-1	186	1.40%	82.26%	2416	0.49%	93.09%
2-2	456	3.43%	88.82%	11538	2.33%	93.52%
3	198	1.49%	100.00%	24366	4.91%	100.00%
All	13296	100.00%	99.37%	496052	100.00%	99.82%