

IvyDB TradeFlow: A Framework for Options Volume Classification and Microstructure Analytics

Garrett DeSimone, PhD Quant Research Dept.
Abhi Gupta, Quant Research Dept.

Executive Overview

IvyDB TradeFlow is an options volume classification tool that identifies buyer- and seller-initiated trades, aggregated at the symbol level across 5-minute, 30-minute, and end-of-day intervals. Its core advantage is classification accuracy: by incorporating trade condition codes and exchange venue data alongside VWAP, IvyDB TradeFlow assigns directional volume more precisely than naive classification algorithms and resolves midpoint trades that those methods leave unsigned.

Beyond the headline buy/sell split, IvyDB TradeFlow provides analytical depth through distinct tables covering venue, VWAP, trade condition codes, and trade location relative to the quote—together with proprietary indicators that separate retail from institutional order flow.

The following sections detail and quantify the accuracy improvement delivered by the new algorithm and illustrate the applications enabled by the added tables. Throughout, the illustrations are drawn from a single trading session—SPY options on 17 June 2026, the day of the June FOMC decision, with the statement released at 14:00 ET — so that every table can be read against the same event.

Trade Classification Methodology

This methodology covers all trades executed on U.S. options exchanges. The algorithm classifies each trade as buyer- or seller-initiated through a sequential, multi-stage process, evaluating each trade against the stages in order until it resolves.

Trades are first screened for characteristics indicating execution against a previously posted passive order and classified accordingly. Remaining trades are screened for small price improvements obtained through exchange auction mechanisms. Trades still unresolved are compared to the prevailing quote; those executing nearer the offer are treated as buys; those nearer the bid as sells.

Midpoint trades, where price alone does not indicate direction, are resolved using supplementary market information available at the time of the trade. Any trade not resolved by the preceding stages is classified by the direction of price movement relative to the prior trade.

IvyDB TradeFlow's advantages over naive assignment (classification by distance from the bid/ask) are twofold:

- 1) **Higher accuracy on commonly classified trades**, reflected in IvyDB TradeFlow's distinct cross-sectional heterogeneity.
- 2) **Greater coverage**, through resolution of midpoint trades that naive methods leave unsigned.

Data Structure and Use Cases

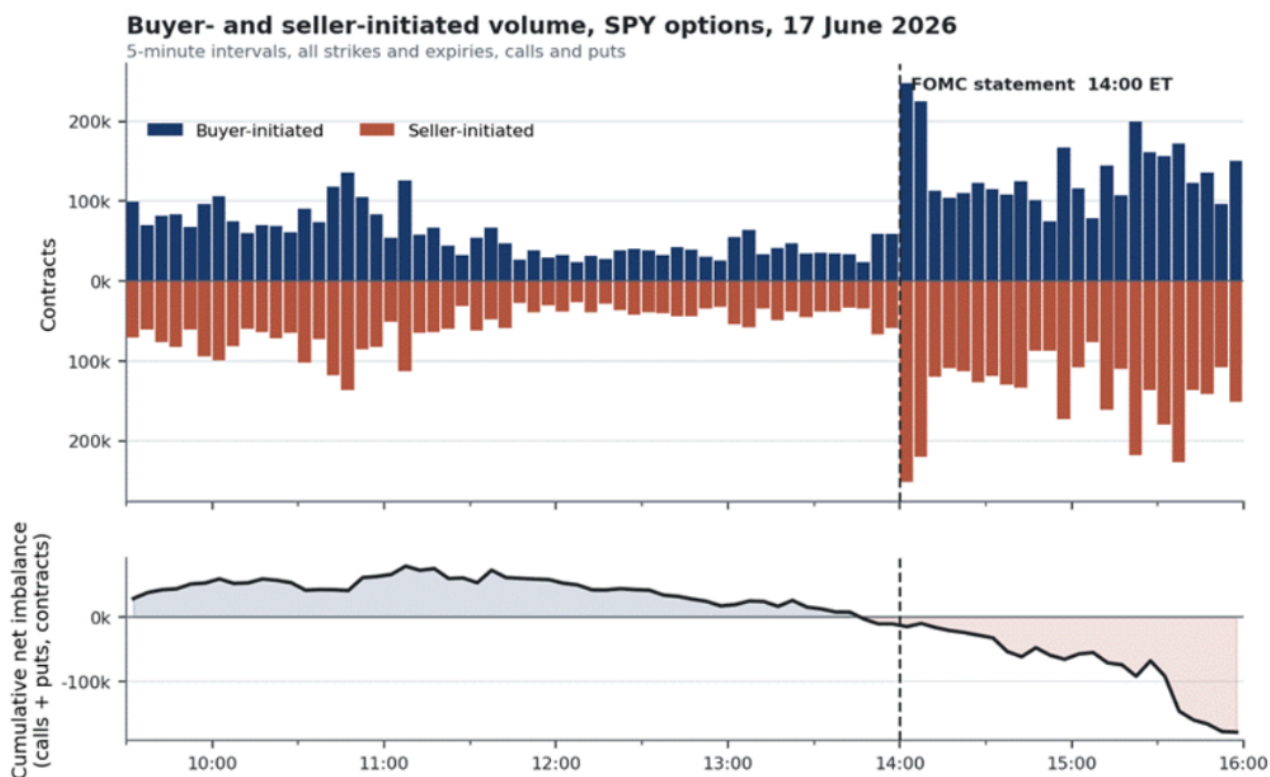
IvyDB TradeFlow is delivered as four relational tables (Signed_Volume, Trade_Conditions, Market_Structure, and Signed_Exchange), each keyed by option symbol, security identifier, and date, with symbols and security identifiers cross-referenced to IvyDB US. Signed_Volume carries the headline classification output; the remaining three tables decompose that output along the dimensions of execution: the condition under which each trade printed, the trade's price location relative to the prevailing quote, and the venue on which it executed.

Signed Volume

Signed_Volume is the core table of the dataset. Each record aggregates all trades in a single option series over a reporting interval, with every trade assigned as buyer- or seller-initiated by the classification algorithm described above; the small remainder that cannot be resolved is retained as unclassified rather than discarded, so directional and unclassified volume sum to total traded volume

Buy/Sell Assignment

The buy/sell decomposition supports the construction of net order imbalance, the workhorse variable in the literature linking option order flow to future underlying returns and shifts in the implied volatility surface. The separation of buy and sell VWAP provides a direct read on the prices at which liquidity demanders transacted, supporting measures of effective execution cost and directional price pressure.



Source: OptionMetrics TradeFlow, Signed_Volume (5-minute). Intervals plotted at bucket start. Net imbalance = buyer-initiated less seller-initiated contracts.

Figure 1. Buyer- and seller-initiated volume and cumulative net imbalance, SPY options, 17 June 2026.

Retail and Institutional Measures

IvyDB TradeFlow reports volume in five proprietary participant buckets in each direction — three retail and two institutional — identified from the combination of trade condition, trade size, price location relative to the prevailing quote, and notional value.

Retail 1 captures single-leg auction (SLAN) prints below ten contracts. SLAN trades route through Cboe's Automated Improvement Mechanism (AIM), a price-improvement auction designed to provide execution advantages for retail order flow, and academic research establishes SLAN as the gold-standard identifier of retail options flow (Bryzgalova, Pavlova, and Sikorskaya, 2023).

Retail 2 captures automatic executions below ten contracts printing at the bid or the offer — a broader measure that includes some institutional algorithmic contamination but retains a retail

signal because retail orders are predominantly marketable, hitting the displayed quote rather than providing liquidity.

Retail 3 captures automatic executions below ten contracts with a notional value under five thousand dollars, isolating price-improved executions inside the NBBO—a signature of wholesaler internalization for retail orders. The three conditions are mutually exclusive, so the buckets may be summed to a single retail measure.

Retail 1 is the strictest, requiring both the auction mechanism through which retail flow is routed and retail-scale size. Retail 1 combined with Retail 3 adds small-premium price-improved executions for a mid-range measure, and all three together give the widest coverage of small marketable flow. Researchers can match the definition to the question, tightening the estimate where precision matters and widening it where coverage does.

Institutional 1 captures intermarket sweep orders (ISOs) executing at or through the prevailing quote. ISOs are large orders that sweep displayed liquidity across multiple exchanges simultaneously, waiving standard price protection in exchange for immediate execution—a signature of sophisticated participants prioritizing speed over price.

Institutional 2 extends this measure by including automatic executions at or through the quote that exceed both ten contracts and five thousand dollars of notional value, capturing aggressive institutional flow that shares the ISO urgency profile without using the ISO mechanism. Institutional 1 is contained in Institutional 2; the two are not additive, and Institutional 2 is the appropriate measure to use.

Both measures isolate aggressive execution by design: every qualifying trade prints at or through the prevailing quote, indicating the initiator crossed the spread for immediate execution. Passive execution, midpoint prints, patient position-building, and market-making activity fall outside these measures. These fields, therefore, capture institutional liquidity demand urgency specifically, rather than total institutional participation.

Net aggressive imbalance, retail- and institutional-classified flow

SPY options, 17 June 2026, 5-minute intervals

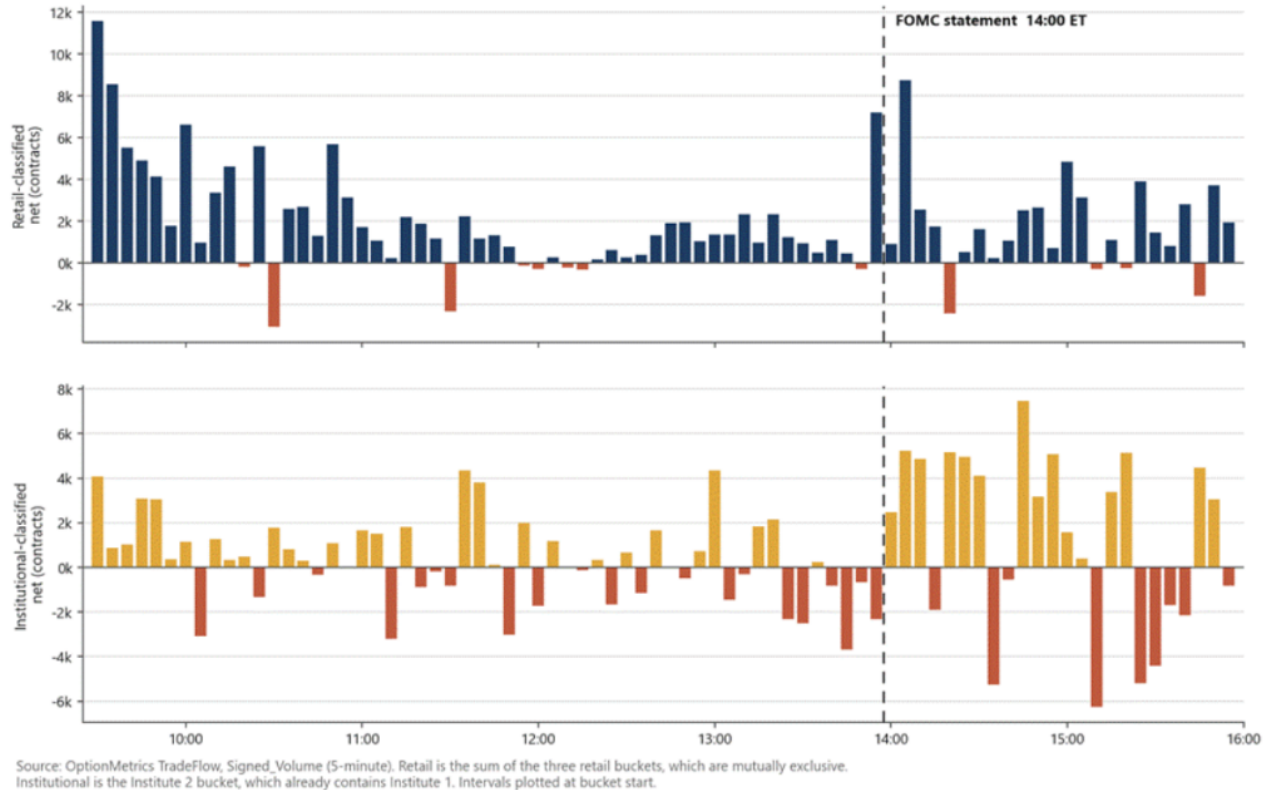


Figure 2. Net aggressive imbalance in retail- and institutional-classified flow, SPY options, 17 June 2026.

VWAP

Alongside the directional split, each record reports quote and trade counts; volume-weighted average (VWAP) prices computed across all trades and separately for buyer- and seller-initiated volume; the distribution of buy and sell trade counts across five lot-size buckets ranging from single-contract trades to blocks of one hundred contracts or more; and proprietary indicators of aggressive institutional and retail activity in each direction.

Lot Size Distribution

The lot-size distribution characterizes participant composition, as single-contract and small-lot trades are strongly associated with retail activity while large blocks reflect institutional execution. Because records are produced at intraday intervals, these measures can be tracked within the

trading day, enabling event studies around earnings announcements, macroeconomic releases, and other scheduled information events.

Trade Conditions

The Trade_Conditions table decomposes each series' classified trade counts by the exchange-reported condition under which each trade printed, such as single-leg auction executions, multi-leg executions, and standard automatic executions. Each row represents one condition code for a given symbol and interval, carrying buyer-initiated, seller-initiated, and unclassified trade counts across the same five lot-size buckets used in Signed_Volume.

Condition Codes

Each row is keyed by the exchange-reported condition code. Seventeen codes appear in SPY options over a single session, covering standard automatic executions, single-leg auctions with and without intermarket sweep protection, the multi-leg family, intermarket sweep orders, and a tail of floor, cross and stock-tied prints. Buyer-initiated, seller-initiated and unclassified counts are carried across the same five lot-size buckets used in Signed_Volume, so each mechanism can be read at trade size.

Multi-Leg Executions

Multi-leg codes identify the legs of spread and combination trades. Each leg is signed individually and on the same basis as an outright trade, so isolating or excluding multi-leg flow is a filter on the data rather than an assumption in the model. Multi-leg prints were 8.2% of SPY option trades on 17 June 2026, 7.6% before the FOMC statement and 8.7% after it.

Auction Flow as a Retail Marker

Auction executions, in which orders receive small price improvements through exchange mechanisms, are widely used in the academic literature as a marker of marketable retail orders routed through wholesalers, providing an independent lens on retail participation.

Figure 3 tracks the mechanism mix across the June 2026 FOMC decision. Standard automatic executions fall from 64.2% of SPY option trades before the 14:00 statement to 55.5% after it, while single-leg auctions rise from 23.3% to 29.5%. The lower panel isolates small-lot single-leg auction prints: the 14:00-14:05 interval carried 36,747 against a pre-statement median of 4,270 per interval, and the median over the remainder of the session was 14,425.

Execution mechanism mix, SPY options, 17 June 2026

5-minute intervals. Auction prints are the literature's marker for marketable retail orders routed through wholesalers.

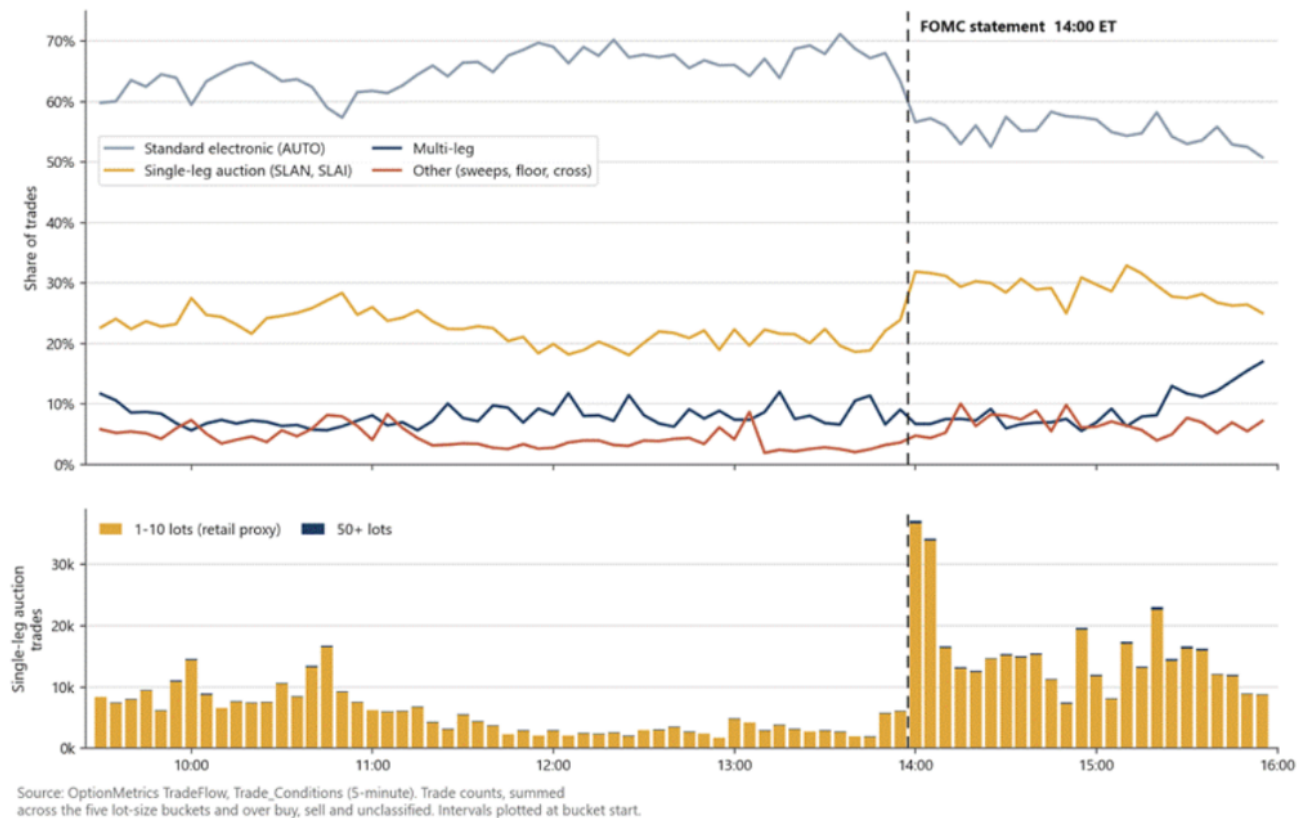


Figure 3. Execution mechanism mix and small-lot auction activity, SPY options, 17 June 2026.

More broadly, the table supports microstructure research on the growth of auction volume, the composition of flow across execution channels, and how the information content of order flow varies by mechanism.

Market Structure

The Market_Structure table records where each trade executed relative to the National Best Bid and Offer prevailing at the time of the trade. Each row represents one price-location attribute (at the NBBO, inside the NBBO, at the midpoint, or outside the NBBO) for a given symbol and interval, with buyer-initiated, seller-initiated, and unclassified volume reported for each.

Price Location Relative to the NBBO

Each row carries buyer-initiated, seller-initiated and unclassified volume for one price-location attribute, symbol and interval. The four attributes are not a flat partition: midpoint volume is delivered as a subset of inside-NBBO volume, so total traded volume is the sum of the at-NBBO, inside-NBBO and outside-NBBO attributes. On that basis the table reconciles exactly to TradeVolume in Signed_Volume for the same symbol and interval.

Midpoint Resolution and Classification Coverage

A trade at the midpoint is equidistant from the bid and the offer, so a classifier that assigns direction by distance from the quote leaves it unsigned. IvyDB TradeFlow resolves these trades using supplementary market information available at the time of the trade and reports them as a labelled category.

On 17 June 2026, 1,321,280 of SPY's 12,932,090 traded contracts — 10.2% — printed at the midpoint, and IvyDB TradeFlow assigned a direction to 99.9% of that volume; across all price locations 0.10% of the day's volume is left unclassified. The lower panel of Figure 4 shows the midpoint share by interval, ranging from 6.5% to 16.4% around a session average of 10.2%.

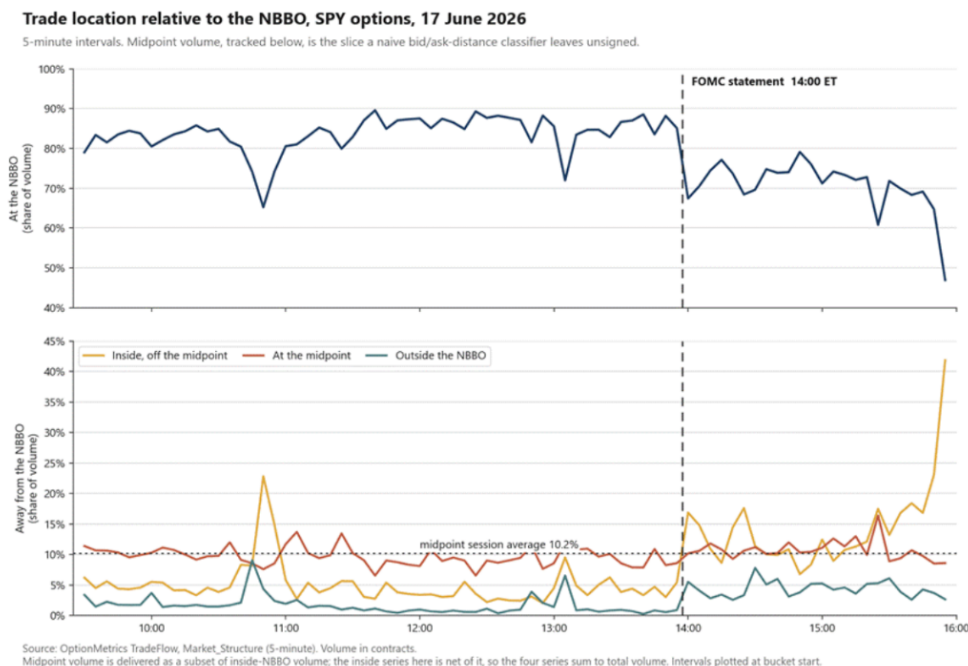


Figure 4. Trade location relative to the NBBO, and the midpoint share of volume, SPY options, 17 June 2026.

Effective Spreads and Price Improvement

Price location is also a direct measure of trade aggressiveness and realized execution cost. Volume executing at the quote reflects liquidity-demanding flow paying the full half-spread; volume inside the NBBO captures price improvement obtained through auctions and wholesaler internalization; and outside-NBBO volume flags sweeps and large negotiated trades printed away from the displayed market.

Figure 4 shows the composition over the session: 76.2% of SPY option volume printed at the NBBO, 20.5% inside it and 3.3% outside. The upper panel tracks the at-the-quote share, which falls from 82.6% before the 14:00 statement to 70.1% after it. The lower panel carries the rest: inside-NBBO volume off the midpoint rises from 5.5% to 14.8%, and outside-NBBO volume from 2.1% to 4.4%.

Together these categories support analysis of effective spreads, price improvement rates, and liquidity conditions by series and time of day, and they complement the Trade_Conditions table by describing the economic outcome of each execution rather than its mechanism.

Signed Exchange

The Signed_Exchange table allocates each series' signed volume across the exchange venues on which its trades were printed. Each row represents one venue for a given symbol and interval, reporting total, buyer-initiated, seller-initiated, and unclassified volume.

Venue Allocation

Each row reports total, buyer-initiated, seller-initiated and unclassified volume for one venue, symbol and interval, and the venue rows sum exactly to the symbol's total in Signed_Volume. SPY options traded on eighteen venues on 17 June 2026, the twelve largest carrying 85.5% of volume and no single venue above 11.8%.

Venue Segmentation of Directional Flow

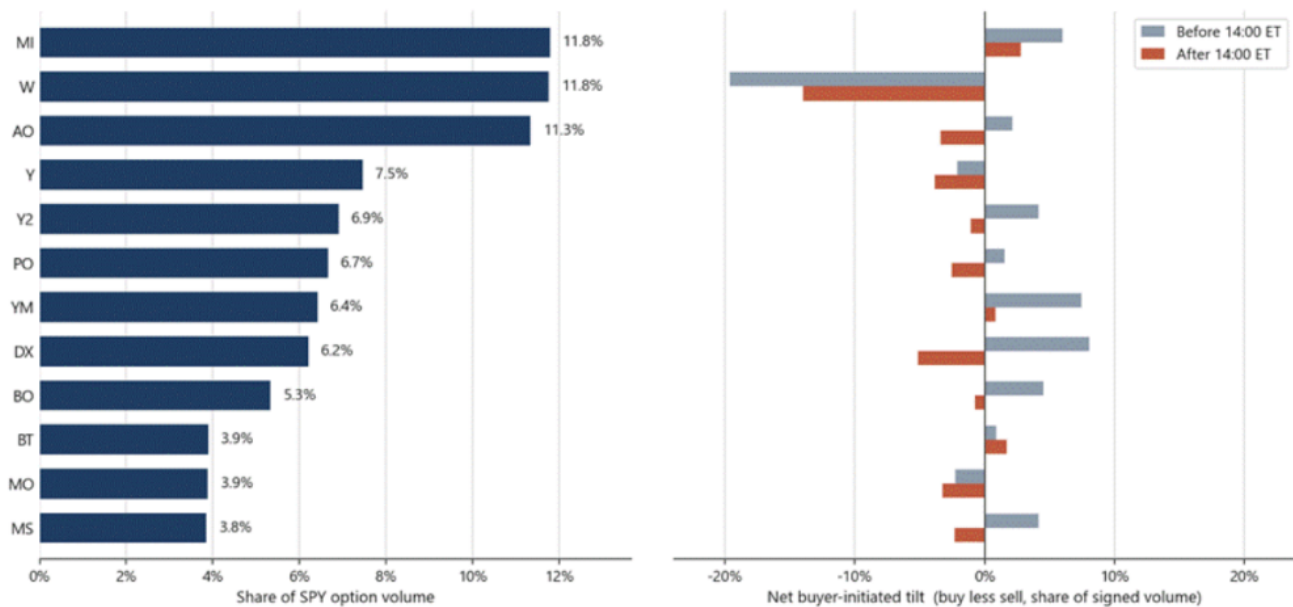
U.S. options exchanges differ substantially in fee structure, auction design, and participant mix, and these differences segment order flow in economically meaningful ways. Venue-level signed volume reveals where directional flow concentrates, separating venues favored by wholesalers executing retail flow from those attracting professional liquidity-taking, and supports research on how venue economics shape the information content of trades.

Figure 5 shows both dimensions for the June 2026 FOMC decision. The three largest venues each carried close to 11.8% of SPY volume, but their net tilts differ: one ran a net seller-initiated tilt of

19.6% of its signed volume before the statement and 13.9% after it, while another ran a net buyer-initiated tilt in both halves. Ten of the twelve largest venues shifted toward net selling after 14:00, one of them by 13.2 percentage points.

Venue concentration and directional tilt, SPY options, 17 June 2026

Twelve largest venues, ranked by volume. Right panel: the same venues' net buyer-initiated tilt, before and after the FOMC statement.



Source: OptionMetrics TradeFlow, Signed_Exchange (5-minute, aggregated to the session).
The 12 venues shown are 85% of the day's volume; 6 smaller venues are omitted. Venues labelled by TradeExchange code.

Figure 5. Venue concentration and net directional tilt by venue, SPY options, 17 June 2026.

The table also enables analysis of exchange market share dynamics, flow migration around fee and mechanism changes, and venue selection as a dimension of execution quality, at the level of individual option series rather than aggregate exchange statistics.