Network-Based Detection of Wash Trading*

Allen Sirolly† Hongyao Ma‡ Yash Kanoria§ Rajiv Sethi¶

November 6, 2025

Abstract

Wash trading refers to the practice of buying and selling securities without taking a net position, for the purpose of artificially inflating recorded volume. It is prohibited by law in the United States, but evidence suggests that it is widespread on some exchanges, especially those involving cryptocurrencies where trader identities can be shielded. The reliable detection of wash trading is challenging because it can be implemented using a variety of different approaches, some of which resemble authentic and lawful strategies such as automated market making. We propose an iterative network-based procedure for detection based on the idea that wash traders form approximately closed clusters of colluding counterparties, seldom transacting with other market participants. Applying this method to the *Polymarket* exchange, we estimate that transaction patterns indicative of wash trading began to trend upward in July 2024, peaking at nearly 60 percent of volume in December 2024. This activity persisted through late April 2025 before subsiding substantially, and once again increased to about 20 percent of volume in early October 2025.

*The authors thank Eric Budish, Jake Marcinek, Michael Ostrovsky, David Parkes, Nicola Rosaia, Xintong Wang, and Jens Witkowski, as well as participants in the Columbia DRO Brown Bag Seminar, for helpful comments and discussion.

†Decision, Risk, and Operations Division, Columbia Business School, as6383@columbia.edu

‡Decision, Risk, and Operations Division, Columbia Business School, hongyao.ma@columbia.edu

§Decision, Risk, and Operations Division, Columbia Business School, ykanoria@columbia.edu

¶Economics Department, Barnard College, Columbia University and the Santa Fe Institute. rs328@columbia.edu

1 Introduction

A wash trade is a financial market transaction that is fictitious, meaning it doesn't involve a bona fide change in market position. The clearest examples involve trades between accounts that have common beneficial ownership, or repeated buying and selling by colluding parties at a price within the bid-ask spread. Such transactions increase trading volume without changing the market position of any individual or entity in the economy, and are prohibited by law in the United States.[1] Since volume is a key measure of market participation and conviction, wash trading distorts the interpretation of market signals and has real economic effects.

The detection of wash trading based on common ownership is challenging even on regulated exchanges that must abide by Know-Your-Client (KYC) requirements, since off-market collusion between account holders is not easy to identify. On unregulated exchanges that use stablecoins or other cryptocurrencies as means of payment, this problem becomes especially acute, as the real-world identities of individual wallet owners are not observed, despite the fact that all transactions between wallets are on-chain. Existing detection algorithms in the literature often rely on identifying simple trading patterns such as two wallets trading the same asset back and forth at the same price, or a closed cycle of transactions leaving all accounts with no change in their aggregate net position. Such methods are insufficient since wash traders can easily adopt complex trading patterns across many wallets to evade such simple detection. Furthermore, the intended trades may be intercepted by an authentic market participant, breaking the circuit. Another approach is to focus on the buying and selling of securities at high frequency, with limited market exposure at any point in time. This by itself is also of limited effectiveness, as such behavior is a characteristic feature of automated market making strategies, which are common and lawful.

In this paper, we propose an iterative network-based method for the *unsupervised* detection of wash trades. We begin by identifying instances in which wallets repeatedly close their open positions. This alone does not indicate a wash trade; a wallet may close a position to lock in gains or losses, or otherwise as part of a legitimate trad-

[1] According to the Chicago Mercantile Exchange, a wash trade is "a form of fictitious trade in which a transaction or a series of transactions give the appearance that authentic purchases and sales have been made, but where the trades have been entered without the intent to take a bona fide market position or without the intent to execute bona fide transactions subject to market risk or price competition" (CME Group, 2025). Along similar lines, the Commodity Futures Trading Commission (which regulates the trading of derivatives, including prediction market contracts) defines wash trading as "entering into, or purporting to enter into, transactions to give the appearance that purchases and sales have been made, without incurring market risk or changing the trader's market position" (CFTC, 2025). Wash trading is prohibited by the Commodity Exchange Act of 1936, codified in 7 U.S.C. §6c(a)(2)(A)(i) (Legal Information Institute, 2025).

ing strategy which aims to make a profit. What distinguishes wash traders, however, is that they trade primarily with other wash traders, in particular with counterparties who also repeatedly open and close their positions. This is the core insight behind our detection algorithm, which has key advantages of being simple and interpretable. We assign an initial score to each wallet based on its position-closing propensity, then iteratively update the score of each wallet as a weighted average of this initial score and the scores of the wallet's trading counterparties. The scores (provably) converge under iteration, at which point we classify trades between wallets with sufficiently high scores as likely wash trades.

We apply this method to estimate the prevalence of wash trading on *Polymarket*. This is a prediction market on which participants can buy and sell binary options—contracts that pay a fixed amount if a referenced outcome occurs and nothing otherwise. The medium of exchange is the USDC stablecoin, which we shall refer to for simplicity as a dollar. Users based in the United States are formally prohibited from trading on Polymarket but this may be circumvented using a VPN or other measures.[2] Since November 2022, Polymarket has operated using a central limit order book (CLOB) for each contract. By mid October 2025, 29.0 billion contracts valued at 15.5 billion dollars had been traded.

Our analysis leverages the complete history of Polymarket trades from November 2022, which include wallet addresses, trade amounts, and share prices. We apply our algorithm to the exchange's *trade graph* across all contracts, where nodes correspond to wallets and edges represent trades between counterparties. By design, wallets with scores close to 1 are likely wash traders, since they almost always close their open positions, and trade almost exclusively with other wallets exhibiting similar behavior. These wallets often trade large volumes on a short time scale relative to a market's duration (see Figure 25). The scale and complexity of their activities is remarkable, and is evident in the trade graph examples we present in Appendix D. In contrast to several existing approaches in the literature (discussed in Section 2), our approach does not require that wash trades occur in a closed cycle. This allows us to detect sequences of likely wash trades that are interrupted by a non-colluding wallet (Example 1), as well as a wide variety of complex patterns—such as irregular chains, clusters, or other seemingly chaotic structures—that might begin or end with trades with authentic market participants (Examples 2 through 6). It also allows us to flag likely wash trades that are disguised by commingling with occasional authentic orders or by occasionally holding positions until market resolution (Example 7).

[2]In July 2025, it was reported that Polymarket was staged to reenter the US market, following the closing of investigations by the Department of Justice and the CFTC, and Polymarket's acquisition of the derivatives exchange QCX (Bloomberg, 2025).

Overall, our algorithm classifies nearly 25% of Polymarket's historical volume as activity consistent with wash trading, and flags 14% of the 1.26 million wallets which have ever traded on the platform.[3] The estimated wash volume peaked at nearly 60% of total weekly exchange volume in December 2024, subsiding to less than 5% by May 2025 before resurging sharply to about 20% by early October 2025, the end of our sample period (see Figure 7). The prevalence of this activity also varies significantly by market category: 45% of all-time volume in Sports markets is classified by our algorithm as likely wash trading, compared to 17% in Election markets, 12% in Politics markets, and 3% in Crypto markets. At their peaks, our estimates reached as high as 95% in Election markets during the week of March 24, 2025, and 90% in Sports markets for the week of October 21, 2024 (see Figure 30). We emphasize that these results are estimates, as there is no definitive "ground truth" proving whether a transaction is a wash trade. However, by examining auxiliary data, such as direct transfers between wallets and common display names, we further identify clusters of wallet addresses whose trading volume is almost entirely self-contained. Our algorithm effectively identifies these large clusters, some of which number in the tens of thousands (see Table 10).

There are several institutional features that together enable and potentially provide an economic incentive for large scale wash trading. First, Polymarket does not implement Know-Your-Customer (KYC) verification, making it straightforward for a user to generate and trade via multiple wallet addresses anonymously. Second, as of this writing, Polymarket does not charge transaction fees, which makes wash trading more feasible than on exchanges which do. Third, the anticipation of a potential token launch—a new cryptocurrency distributed to users—incentivizes so-called *airdrop farming*.[4] Airdrops are a common strategy to scale markets with substantial network effects, retroactively rewarding users with free tokens based on their activities prior to the token launch. This, in turn, incentivizes users to "artificially inflate their trading volume in the hopes of scooping a larger airdrop reward" (Gladwin, 2024).

The ability to detect wash trading is important for the long-term health and growth of the market. First, when a wash trader places executable orders within the current prevailing bid-ask spread, this contributes neither liquidity nor information to the prediction market. Any potential future airdrop should therefore be designed to exclude such activities. Second, investors and new users often rely on trading volume as a key indicator of platform health and adoption. Robust detection, therefore,

[3] Wash trading has been suspected on Polymarket for some time. For instance, an October 2024 article in *Fortune*, citing non-public reports by crypto research firms Chaos Labs and Inca Digital, reported significant wash trading in the 2024 U.S. Presidential Election market (Schwartz, 2024).

[4] In October 2025, Polymarket confirmed a future airdrop (CoinDesk, 2025).

builds confidence that the platform's metrics accurately reflect its growth and user engagement. Third, and crucially for a prediction market, high volume is often interpreted as evidence that the prediction implied by a contract's price aggregates the wisdom of a larger crowd. Hence, trustworthy volume metrics allow participants to properly interpret the market consensus. Furthermore, markets on the home page of the exchange are by default ranked in decreasing order of 24-hour volume. Effective wash trade detection thereby also helps prevent easy manipulation of these rankings to draw attention to a particular set of contracts.

The rest of the paper is organized as follows: Section 2 discusses related work. Section 3 describes in detail the mechanics of Polymarket and the collected data. Section 4 presents examples of activity consistent with wash trading uncovered in the trade data, and Section 5 presents our algorithms for estimating the extent of wash trading. Section 6 presents our empirical measures of wash trading, additional evidence of common ownership based on the network of direct USDC transfers on Polygon, and a comparison of our results with those obtained using algorithms in related work.

2 Related Work

Cao et al. (2016) considers the problem of detecting wash trading attempts from a stream of limit orders, such that they may be caught or blocked before execution. Given the most recent order, their detection algorithm first searches for sets of potential matching limit orders via a "volume-matching" procedure (a knapsack problem).[5] It then searches among the volume-matched sets for closed cycles, i.e., sets of orders in which the aggregate positions of the traders would remain unchanged if executed. (In this setting, wash traders may strategically set limit orders with some level of price mismatch but which nevertheless become matched trades; this is not a complication for us since we observe executed trades.) The authors test the algorithm against real order data for seven stocks in which fake wash trades following prespecified patterns are injected.

There are a number of papers which attempt to estimate the prevalence of wash trades on cryptocurrency exchanges. Cong et al. (2023) checks for statistical discrepancies in trade sizes between regulated and unregulated exchanges—in the distribution of the first digit (Benford's law); in the pattern of trade size clustering; and in the tail exponent of the trade size distribution. They estimate that wash trading

[5] That is, the author does not assume the common "price-time priority" protocol for order matching.

averages "more than 70% of the reported volume" on unregulated exchanges, perhaps representing attempts to game exchange rankings.

A number of papers overcome the limitations of statistical detection methods by accessing trade-level data containing trader IDs. Victor and Weintraud (2021) examine wash trading on IDEX and EtherDelta, two Ethereum-based token exchanges. Their detection algorithm first generates candidate sets of wallets based on frequently-occurring trades among strongly-connected components (SCCs) of the trade graph for each token. Adapting the method from Cao et al. (2016), it then performs a volume-matching step that searches for subsets of trades within each SCC "that lead to no change in the individual position of each participating trader," up to a small margin. They conservatively find that "at least 10% of traded tokens have a wash trading share of at least 20%" on both IDEX and EtherDelta, with the respective exchange-wide wash shares peaking at 90% and 60% of weekly volume.

Aloosh and Li (2024) use leaked transaction data from the defunct Mt. Gox exchange and use this as ground truth to evaluate indirect statistical detection methods. They find that trades with identical buyer and seller addresses ("self-self" trades) comprise 32.5% of trading volume over their 2011–2013 sample period. To detect wash trades involving multiple wallet addresses, the authors implement the volume-matching algorithm of Victor and Weintraud (2021), from which they find that cyclic wash volume is negligible compared to self-self volume. Note that Polymarket has blocked self-self trading on its exchange since April 27, 2023.[6]

von Wachter et al. (2022) examine wash trading on the OpenSea marketplace for non-fungible tokens (NFTs; typically digital artworks), where sellers may choose between fixed-price and auction listings. In such a setting, a malicious user can manipulate volume and prices through wash trades, inflating a target NFT's perceived value to unsuspecting potential buyers. The authors detect closed cycles and rapid "path-like patterns" with minimal price movement in the trade graphs for 52 large NFT collections, and "identify 2.04% as the lower bound of suspicious sale transactions that closely follow the general definition of wash trading." Note that, in contrast to NFTs sold on OpenSea which are indivisible, shares on Polymarket may be bought or sold in any quantity; as a consequence, wash trading schemes may be more complex and harder to identify in our setting. (There are other platforms which support trade in fractionalized NFTs.)

In addition to the academic studies discussed above, several crypto industry firms which monitor exchanges have published details on methods for wash trade detection.

[6] https://discord.com/channels/710897173927297116/775506448041115669/1101210578346835968, accessed October 14, 2025.

CoinMetrics' *Trusted Exchange Framework* grades exchanges on a collection of factors including "trusted volume". The tests for this category include the statistical tests similar to those used in Cong et al. (2023), as well as a quantification of "round-trip trades", defined as "two trades with nearly identical prices and amounts executed on the opposite side (buy or sell) in close proximity in time and sequence order".[7] Chaos Labs, whose findings on Polymarket wash trading were reported by *Fortune*, describes its "wash trading detection module" for use in a chain launch incentive program by the dYdX platform. The module comprises trade graph analysis for "abnormal ownership changes in relation to trade volume among clustered accounts", followed by manual screening to reduce the incidence of false positives.[8]

Separate from the problem of detection, our findings give empirical support to theories of market microstructure. Glosten and Milgrom (1985) posit an information-based theory of market microstructure, in which liquidity takers trade on more up-to-date information than market makers; this informational edge leads to higher returns for takers on average. We present supporting evidence for this theory in Section 6, where we find that clusters of wash-trading wallets—which are by definition not trading on information, and whose orders are occasionally intercepted by non-colluding outsiders—typically lose money in aggregate.

Finally, our work is relevant for studies which use trade-level data from Polymarket. Examples in this nascent research area—which use trades to infer latent trader beliefs—are Eichengreen et al. (2025) and Chen et al. (2024).

3 Data

In this section we provide an overview of the mechanics of the Polymarket exchange, describe the structure and scope of the trade-level data, and document several anomalies that suggest the presence of artificial trading.

In the language of Ethereum, Polymarket is a *decentralized application* (dApp) which runs on Polygon, a "Layer-2" extension of the Ethereum blockchain which allows for higher transaction throughput with lower gas fees. We use the API of the blockchain explorer Polygonscan to retrieve all historical Polymarket transactions on Polygon.[9] Each transaction involves a transfer of USDC collateral or shares (condi-

[7]https://5264302.fs1.hubspotusercontent-na1.net/hubfs/5264302/special-insights/coinmetrics-research_trusted-exchange-framework-2-3.pdf and https://coinmetrics.substack.com/p/state-of-the-network-issue-323, accessed September 5, 2025.

[8]https://chaoslabs.xyz/posts/dydx-chain-launch-incentives-program-wash-trading-detection, accessed July 14, 2025.

[9]https://polygonscan.com/. As of August 15, 2025, the Polygonscan API has been subsumed

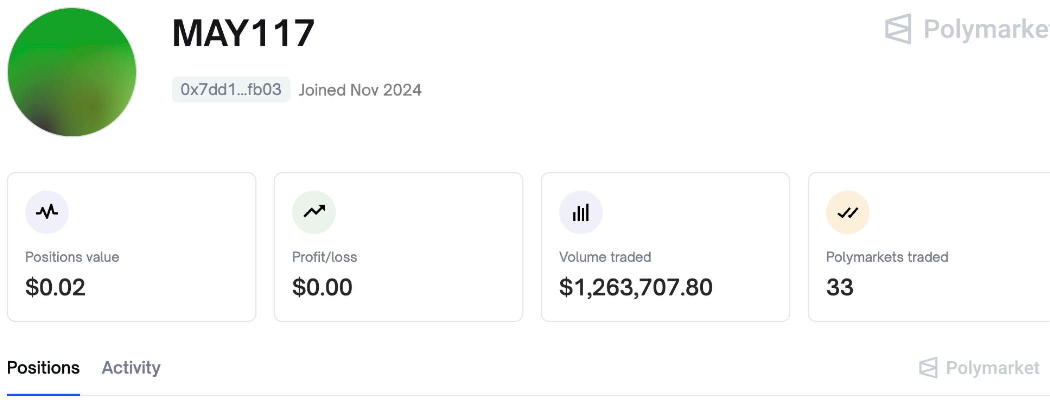


Figure 1: Example of a Polymarket wallet page with display name (URL: https://polymarket.com/profile/0x7dd15526dd14c21b8ff82fd7e0756eee9d71fb03). There are 200 wallets with display names starting with "MAY" that trade almost exclusively with each other, achieving a total volume of over 116 million shares and aggregate profit of merely –$57.86 (see Table 10). Note that (despite the presence of the dollar sign) volume refers to the number of *shares* that the wallet traded, not the dollar value of those shares.

tional tokens) between a wallet address—a 42-character hexadecimal string starting with "0x"—and a Polymarket *module*; see Appendix A. We are also able to retrieve direct transfers between wallet addresses, which provide strong evidence of common ownership; see Section 6. Note that in Polymarket's central limit order book (CLOB), orders are matched off-chain, meaning that the state of the CLOB and the order matching rules are not discernible from on-chain data.

We additionally use Polymarket's Gamma Markets API to retrieve market-level information, including the names of markets, their start and end times, and IDs for the conditional tokens corresponding to the possible outcomes. Furthermore, we obtain data from Polymarket's publicly accessible user profiles, including display names and wallet `created_at` timestamps, which may provide further evidence of association.[10]

For a dramatic example of a wallet that is very likely to be engaged in wash trading, see Figure 1. This trader registered volume well in excess of a million shares across 33 separate markets, but managed to do so with a profit of precisely zero. There are many such wallets in the data, as described in the figure caption.

by the Etherscan V2 API; see https://docs.etherscan.io/etherscan-v2, accessed August 24, 2025.

[10] https://docs.polymarket.com/developers/gamma-markets-api/overview, accessed July 14, 2025. Note that about 96,000 wallets (7% of wallets) do not have public user attributes. These attributes, however, are only used for auxiliary analysis and descriptive statistics—our proposed algorithm relies exclusively on the transaction-level data, which is publicly available for all wallets. Also note that users may change their display names; we collected snapshots on several dates in August, September and October 2025.

Markets, Events, and Trades Each *market* is a question with a binary outcome, e.g., "Will the U.S. have a recession in 2025?" These outcomes are typically represented by "Yes" and "No".[11] In contrast to traditional exchanges in which a stock or token is traded perpetually, each market has a concrete expiration time (sometimes within hours of when trading opens), which means that traders may rely on the payout mechanism to recoup their collateral, rather than selling shares. Payouts depend on the realized outcome, which is determined by a decentralized voting mechanism known as the UMA protocol.[12]

An *event* comprises one or more markets, depending on whether there is a binary outcome or multiple possible outcomes. For example, the event "What products will Apple launch on September 9?" is comprised of the markets "Will Apple launch an Apple Watch on September 9?", "Will Apple launch an iPhone SE on September 9?", and others. An important category of event is called *negative risk* (or NegRisk), which describes collections of markets whose outcomes are mutually exclusive and collectively exhaustive. For example, the NegRisk event "Who will win the election?" may be comprised of the markets "Will A win the election?", "Will B win the election?", and "Will somebody else win the election?" The order books for such markets associated with a NegRisk event are not automatically linked, despite the logical dependency among the outcomes. This occasionally creates opportunities for arbitrage—for instance if the sum of all "No" contract prices is less than a dollar—but such opportunities tend to be quickly exploited and extinguished.

In each market, wallets can perform the following actions using the Polymarket web interface or API:

- Buy: A wallet may place an order to buy shares at price (at most) p. This may be marketable, in which case a trade is consummated against a seller willing to accept p or a buyer of the complementary contract willing to pay $1 - p$. Or it may be non-marketable, in which case the order remains in the order book until it trades against a future order or is canceled.[13]
- Sell: Similarly, a wallet may place an order to sell.
- Split: A wallet may convert one unit of (USDC) collateral into one "Yes" and one "No" share.

[11] For markets whose conditional tokens are not labeled "Yes" and "No"—e.g., sports matches which label outcomes "Team A" and "Team B"—we map their outcomes to "Yes" and "No" arbitrarily.

[12] https://uma.xyz/#how-it-works, accessed July 14, 2025.

[13] Orders may also be partially marketable, with some of the demand filled and the unfilled remainder entering the order book.

- Merge: A wallet may convert one "Yes" and one "No" share into one unit of collateral.
- Redeem: Once a market has resolved, wallets must manually redeem their shares to receive their payout value in collateral.

In addition, NegRisk events permit the following:

- Convert: A wallet may convert one or more "No" shares into a payout-equivalent amount of "Yes" shares (in the complementary set of markets nested under the event) and collateral.[14]

We focus our analysis on buy and sell transactions only, as the other transaction types do not involve a counterparty and do not contribute to CLOB volume.

In every *trade*, a wallet takes either a *long* position (by buying "Yes" or selling "No") or a *short* position (by buying "No" or selling "Yes"). Each trade involves a liquidity taker and one or more liquidity providers or makers. The taker is the party submitting a marketable order, which crosses with an order (or multiple orders) that are currently resting in the book. See Table 1 for an example. Each trade occupies a position within a block on the Polygon blockchain, indicated by the block number and the transaction index; the timestamp is at the block level, not the trade level. After matching shares line-by-line as in the bottom of Table 1, one may categorize each trade as one of three types: (i) a "buy/buy" trade, in which the long and short wallets buy "Yes" and "No" shares respectively, i.e., taking opposite sides of a bet; (ii) a "buy/sell" trade, in which the wallet on one side sells shares to a buyer on the other side; and (iii) a "sell/sell" trade in which the long and short wallets sell "No" and "Yes" shares, analogous to (i). Since April 20, 2023, prices have been available in some markets in increments of \$0.001 (one-tenth of a cent); as of this writing, the increment switches from \$0.01 to \$0.001 (and remains at \$0.001) when a market's last traded price reaches $p < \$0.04$ or $p > \$0.96$.[15]

Trading Volume Throughout the remainder of the paper, we refer to two commonly-used measures of volume, namely *share volume* and *dollar volume*. Share volume measures the number of shares traded, while dollar volume measures the value of those

[14] Using the preceding election example, one "No" share in "Will A win?" plus one "No" share in "Will B win?" has the same contingent payout as one "Yes" share in "Will somebody else win?" plus \$1 (\$1 if A or B wins, \$2 if somebody else wins). See https://github.com/Polymarket/neg-risk-ctf-adapter#polymarket-multi-outcome-markets and https://discord.com/channels/710897173927297116/775506448041115669/1187619141683781663, accessed July 14, 2025.

[15] https://discord.com/channels/710897173927297116/775506448041115669/1098745036503535696 and https://docs.polymarket.com/developers/CLOB/websocket/market-channel#tick-size-change-message, accessed October 14, 2025.

			long				short			
block	index	timestamp	wallet_id	type	price	shares	shares	price	type	wallet_id

10000000	1	2025-01-02 03:04:05	`0x123...abc`	buy	0.955	1000	500
							200
							300
								0.045
								0.045
								0.955
									buy
									buy
									sell
										`0x234...bcd`
										`0x345...cde`
										`0x456...def`

↓

			long				short		
block	index	timestamp	wallet_id	type	price	shares	price	type	wallet_id
10000000	1	2025-01-02 03:04:05	`0x123...abc`	buy	0.955	500	0.045	buy	`0x234...bcd`
10000000	1	2025-01-02 03:04:05	`0x123...abc`	buy	0.955	200	0.045	buy	`0x345...cde`
10000000	1	2025-01-02 03:04:05	`0x123...abc`	buy	0.955	300	0.955	sell	`0x456...def`

Table 1: Top: A fictional example trade. Bottom: A reformatted view of the same trade, in which the shares on the long side of the trade are matched line-by-line with each of the short counterparties. Note that we calculate the per-share price by dividing the collateral amount by the share quantity. Also note that for buying on the long side and selling on the short side, the price shown is that of a "Yes" share; otherwise it is that of a "No" share.

shares in USDC. (Since USDC is a stablecoin whose value is pegged to the U.S. dollar, we use "$" to denote dollar volume for convenience.) We adopt the convention that in buy/buy and sell/sell trades, N shares each of "Yes" and "No" together contribute N units of share volume (and $\$N$ of dollar volume), as opposed to $2N$ units.[16] In the example trade in Table 1, this means that the total share volume is 1000, while the total dollar volume is $\$(700 + 300 \times 0.955) = \$986.50$. Polymarket's daily share and dollar volume since January 2024 are shown in Figure 13 in Appendix B.

Our complete data set comprises the entire history of Polymarket CLOB trades from November 21, 2022 through October 12, 2025: 67.7 million trades (82.7 million using the bottom representation in Table 1) by 1.33 million wallets in 102,532 markets (45,732 events). Over this period, 29.0 billion shares corresponding to $15.5 billion in dollar volume were traded. Figure 14 in Appendix B presents distributions to characterize market and wallet activity. The most heavily traded event in Polymarket history referenced the outcome of the 2024 U.S. Presidential Election, in which 3.7 billion in share volume ($2.0 billion in dollar volume) was traded.[17] Figure 15 shows the monthly number of new events; the rapid growth in the series since June 2025 is driven by the launch of short-duration cryptocurrency events, e.g., "Bitcoin Price – August 27, 4PM ET." Figures 16 and 17 show the number of daily active (i.e., trading) wallets, and the fraction of wallets which were active after 90–120 days, by

[16]This is a standard procedure to avoid double counting volume. Our measure of share volume is consistent with volume numbers reported by Polymarket (except that Polymarket's volumes are affixed with a dollar sign, which we reserve for dollar volume).

[17]https://polymarket.com/event/presidential-election-winner-2024, accessed October 8, 2025.

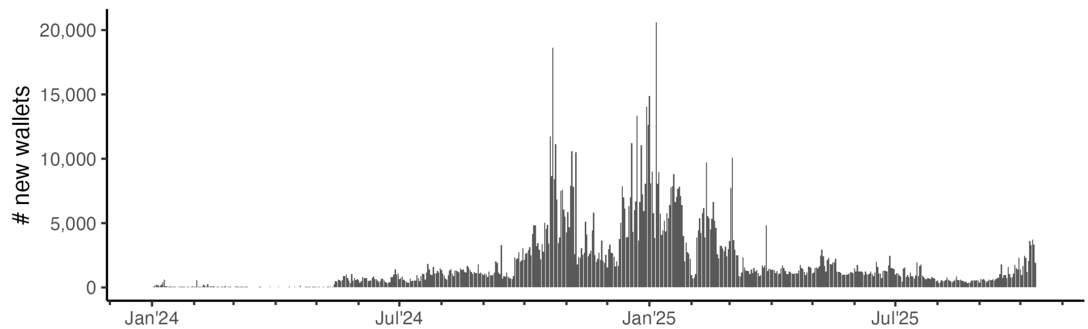


Figure 2: Daily number of newly created wallets (which traded on Polymarket at least once; dates prior to January 1, 2024 are not shown).

their `created_at` date.

Anomalies Before identifying and quantifying wash trades, we present evidence of anomalies using several aggregate measures. First, there is unnatural growth in the number of new wallets. Figure 2 shows a large increase in the rate of new wallet activity starting May 14, 2024. By early January 2025, more than half of the wallets that had ever traded made their first trade after Election Day (November 5, 2024), only two months earlier. Over the same period of time, there was no corresponding increase in open interest (the total value of outstanding shares in markets that had yet to be resolved), as shown in Figure 23 in Appendix B. Moreover, the display names of many of these new wallets conform to common patterns: for example, Figure 18 in Appendix B reveals an unusually high frequency of names with exactly 10 or 17 characters; additionally, tens of thousands of wallets created in quick succession have a name which is a first name followed by several digits or characters, e.g., "Veronica2nge", "Winona9r1g", "Maxwellmk3d". Figures 19 and 20 show that a large fraction of these wallets have a cumulative profit-or-loss of less than \$1, and trade mostly at prices less than \$0.01 or greater than \$0.99. By August 2025, the rate of growth of new wallets returned to a level slightly below that immediately after May 14, 2024; starting in mid-September, however, the growth rate again increased substantially.

Second, beginning in January 2024, there is an increase in the fraction of share volume attributable to buy/sell trades (relative to buy/buy or sell/sell), and a divergence between this metric when measured by dollar volume versus share volume. Namely, the buy/sell fraction of share volume increases significantly compared to the buy/sell fraction of dollar volume; see Figure 3. This is explained by a large increase in the fraction of buy/sell share volume at near-zero prices, as shown in Figure 4.[18] If

[18]The first sharp increase corresponds to the start of trading for the 2024 U.S. Presidential

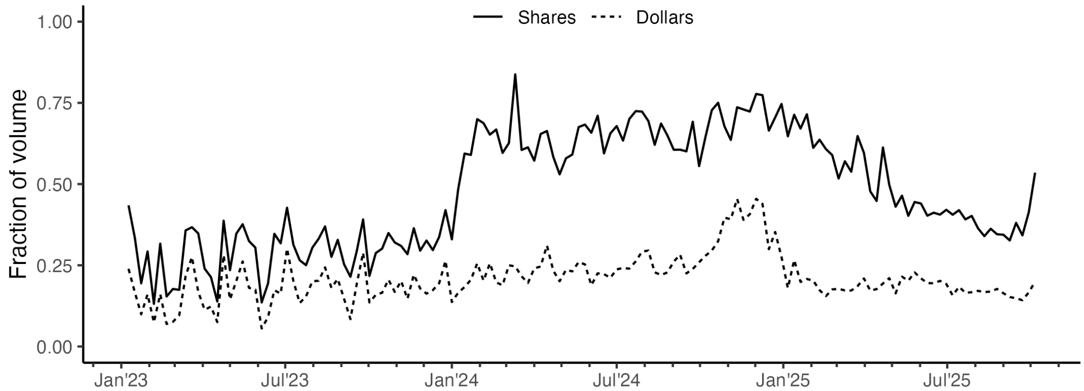


Figure 3: The fraction of weekly volume due to buy/sell trades, versus buy/buy or sell/sell. Note the large and persistent increase in the buy/sell fraction of share volume beginning January 2024, and the divergence of share volume relative to dollar volume. (The pattern is similar when 2024 U.S. Presidential Election trades are excluded.)

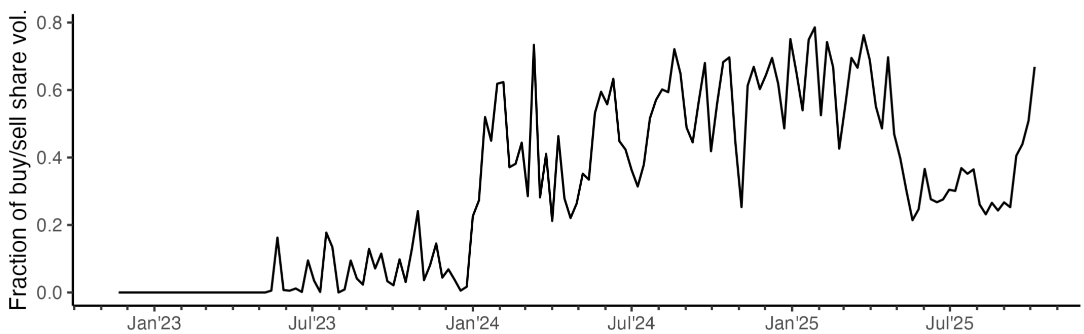


Figure 4: The fraction of weekly buy/sell share volume at price $p <$ \$0.01. (Note that \$0.001 tick sizes were first made available for some markets starting April 20, 2023.) See also Figures 21 and 22 in Appendix B.

wallets are trying to generate large share volume with minimal capital commitment, an effective strategy is to repeatedly buy and sell a large number of very low-priced shares.[19] Figure 21 in Appendix B shows that a sizable fraction of weekly active wallets trade at prices less than \$0.01 or greater than \$0.99.

Third, as shown in Figure 24 in Appendix B, the ratio of platform-wide volume to open interest increased substantially in 2024 and 2025, compared to 2023. This corresponds to the increase in buy/sell trading which contributes to volume but not open interest (in contrast to buy/buy trading which adds to both volume and open interest). Note that this may reflect a rise in legitimate algorithmic trading.

Election markets on January 4, 2024.

[19]This would increase share volume to a much greater degree than dollar volume, but Polymarket has tended to conflate the two measures (Schwartz, 2024).

4 Wash Trades

As the definitions of wash trades by the CTFC and CME Group suggest (see Footnote 1), wash trading involves the buying and selling of securities without incurring market risk or changing one's market position, while creating the appearance of an authentic transaction. In general, traders do not get to choose their counterparties when placing orders. However, it is possible for an individual to trade with a desired counterparty (or an account with the same beneficial ownership) by coordinating to place orders in quick succession via two or more wallets within the prevailing bid-ask spread. This may be done manually or programmatically through Polymarket's trading API.

For example, a trader may commit $\$x$ of capital between two wallets to buy x "Yes" shares and x "No" shares via two wallets *in a single trade*, such that the overall net position (zero) pays $\$x$ regardless of the market outcome.[20] While such a buy/buy trade is a wash trade, it is, by itself, not particularly efficient at generating large volume per unit of capital committed, if the trader has to wait hours, days, or longer to recover the invested capital upon market resolution. To generate additional volume using the shares owned, the trader may subsequently sell shares to additional wallets under the same ownership control, or execute a sell/sell trade between the two wallets with open positions. In either case, the positions of one or both wallets is closed while the trader's overall net position (zero) remains unchanged.

It is also possible that, after having legitimately acquired a non-zero net position in the market, the trader sells shares repeatedly through a sequence of wallets under common ownership and then closes out the position at the prevailing price. This needs to be done quickly—relative to the arrival rate of new market orders which shift prices—to avoid the risk of unfavorable price movements that decrease the value of the position.

In both of the above cases, there is the possibility of an "interception" in the following scenario: A trader who intends to execute a wash trade pings the Polymarket API to get the best bid and ask prices. Before the trader submits orders for two wallets under their control, a third, unaffiliated wallet places a limit order within the bid-ask spread. One of the orders sent by the wash trader may execute against this order, while the other remains unfilled. If the trader initially had a zero net position, this is no longer the case after the interception, resulting in a potential profit or loss.

[20] Recall that Polymarket offers this transaction directly in the form of a "split" (see Section 3). However, split transactions do not involve counterparties (and do not contribute to volume) and therefore do not qualify as wash trades.

4.1 Examples

We now present several examples of wash trading activity in different markets. While not exhaustive, the examples illustrate the diversity of trading patterns that potential wash traders have adopted. These include strategies which appear designed to evade detection. (Readers may skip to Section 5 for a description of our detection algorithm.)

In addition to the fields included in the example trade in Table 1, we calculate the cumulative net position (net_pos) of each wallet following each of its trades in the market (positive if the wallet has a net long position, and negative if it has a net short position). This number may reflect previous trades a wallet made which are not shown in the table.

In discussing the examples, we refer to wallets executing a sequence of wash trades as "colluders", and to non-colluders as "the market". We also say that a wallet *buys with* another if each buys shares on opposing sides (i.e., "Yes" and "No") to consummate a trade, committing a total of 1 dollar per unit volume. Similarly, a wallet *sells with* another if each sells shares on opposing sides to consummate a trade.

The first example illustrates two wallets trading shares back and forth, as well as an interception by a legitimate account.

Example 1 (Will the Republican candidate win Pennsylvania by 1.0%–1.5%?)**.** As shown in Table 2, MAY175 first buys 7,291.07 shares with MAY20. MAY175 then trades its "No" shares with MAY176 repeatedly, alternating as buyer and seller. After 90 such trades—over a 30-minute period during which there are only two non-MAY trades in the market—MAY176's buy order for the "No" shares appears to be intercepted by `0x203...cd1`, which buys "Yes" shares to increase its long position. This leaves each of MAY175 and MAY176 with 7,291.07 "No" shares. MAY175 closes its position by selling with MAY20, and MAY176 later closes by selling to the market. The rapid back-and-forth trading by MAY175 and MAY176 is visible as a sharp increase in the market's overall volume as seen in Figure 25 in Appendix B (left column, first row).

The following example illustrates colluders' opening and closing of positions via buy/buy then sell/sell trades. The trades, however, occur at different prices and involve more than two wallets.

Example 2 (Will Chop Robinson win NFL Defensive Rookie of the Year?)**.** As shown in Table 3, `zhongxin` buys 1,000 shares with `0xb19...ebd`. It then sells with `0xb19...ebd` in two separate trades to close out both positions. About ten minutes

block	index	timestamp	long				shares	short			
			display_name	net_pos	price	type		type	price	net_pos	display_name
64356641	97	2024-11-16 13:50:07	MAY20	7291.07	0.017	buy	7291.07	buy	0.983	-7291.07	MAY175
64356648	81	2024-11-16 13:50:21	MAY175	0	0.983	sell	7291.07	buy	0.983	-7291.07	MAY176
64356654	53	2024-11-16 13:50:33	MAY176	0	0.983	sell	7291.07	buy	0.983	-7291.07	MAY175
64356669	48	2024-11-16 13:51:07	MAY175	0	0.983	sell	7291.07	buy	0.983	-7291.07	MAY176
64356679	15	2024-11-16 13:51:29	MAY176	0	0.983	sell	7291.07	buy	0.983	-7291.07	MAY175
		⋮					⋮				
64357423	34	2024-11-16 14:18:11	MAY175	0	0.983	sell	7291.07	buy	0.983	-7291.07	MAY176
64357433	36	2024-11-16 14:18:33	MAY176	0	0.983	sell	7291.07	buy	0.983	-7291.07	MAY175
64357456	81	2024-11-16 14:19:23	`0x203...cd1`	16386.21	0.018	buy	7291.07	buy	0.982	-7291.07	MAY176
64357503	71	2024-11-16 14:21:01	MAY175	0	0.983	sell	7291.07	sell	0.017	0	MAY20

Table 2: Trades for Example 1, in the market "Will the Republican candidate win Pennsylvania by 1.0%-1.5%?" Also see Figure 5.

later, it buys 1,000 shares with `0xaa3...b5c` in two separate trades, and again sells with `0xaa3...b5c` in two separate trades to close out both positions. In both cases, the sell price differs significantly from the buy price ($0.984 vs. $0.949 and $0.944 vs. $0.984); there are no intervening trades, meaning that the price differential may indicate a wide bid-ask spread in a thin market (there had only been two trades in the market prior to `zhongxin`'s first trade, and `zhongxin`'s 24 trades in the first 3% of the market's duration account for more than 50% of the total market share volume). Moreover, these wallets are affiliated through the network of direct USDC transfers shown in Figure 34 in Appendix B.

block	index	timestamp	long				shares	short			
			wallet_id	net_pos	price	type		type	price	net_pos	wallet_id
65233769	55	2024-12-08 13:43:52	`0xb19...ebd`	1000	0.051	buy	1000	buy	0.949	-1000	zhongxin
65233799	30	2024-12-08 13:45:08	zhongxin	-600	0.984	sell	400	sell	0.016	600	`0xb19...ebd`
65233813	26	2024-12-08 13:45:38	zhongxin	0	0.984	sell	600	sell	0.016	0	`0xb19...ebd`
		⋮					⋮				
65234102	24	2024-12-08 13:56:16	`0xaa3...b5c`	100	0.016	buy	100	buy	0.984	-100	zhongxin
65234113	23	2024-12-08 13:56:40	`0xaa3...b5c`	1000	0.016	buy	900	buy	0.984	-1000	zhongxin
65234144	24	2024-12-08 13:57:46	zhongxin	-950	0.944	sell	50	sell	0.056	950	`0xaa3...b5c`
65234192	43	2024-12-08 13:59:28	zhongxin	0	0.944	sell	950	sell	0.056	0	`0xaa3...b5c`

Table 3: Trades for Example 2, in the market "Will Chop Robinson win NFL Defensive Rookie of the Year?"

Example 3 illustrates what we refer to as a "triangular trade", in which shares are sold to a third wallet in between buy/buy and sell/sell trades.

Example 3 (Will the Ravens win the AFC Championship?)**.** As shown in Table 4, `srxget4` buys 63,000 shares with `nojkfaes`. `nojkfaes` subsequently sells its shares to `gfhdfgtyh5e`. `gfhdfgtyh5e` then sells with `srxget4`, such that all three wallets have closed their positions within two minutes. These are the only trades by the two wallets `srxget4` and `gfhdfgtyh5e` across all markets, while `nojkfaes` repeats this pattern of trades many times with other wallets in different markets.[21] The trading relationships between `nojkfaes` and its counterparties (across all markets) may be visualized as a triangular "hub-and-spoke" graph as shown in Figure 11, which includes additional "hub" wallets other than `nojkfaes`.

			long					short			
block	index	timestamp	display_name	net_pos	price	type	shares	type	price	net_pos	display_name
66297767	78	2025-01-04 14:24:51	srxget4	63000	0.212	buy	63000	buy	0.788	-63000	nojkfaes
66297806	73	2025-01-04 14:26:13	nojkfaes	0	0.788	sell	63000	buy	0.788	-63000	gfhdfgtyh5e
66297806	78	2025-01-04 14:26:47	gfhdfgtyh5e	0	0.788	sell	63000	sell	0.212	0	srxget4

Table 4: Triangular trades in Example 3, in the market "Will the Ravens win the AFC Championship?"

In the following example, shares are traded along a chain of wallets, before eventually being sold back to the market.

Example 4 (Will the US add more than 300k jobs in December 2024?)**.** As shown in Table 5, wallet `0x90b...c91` buys 95 shares with the market (i.e., it buys "No" shares, taking the short side). It sells these shares to `0xa44...b8f`, which sells to `0xb5b...cd7`, and so on in a trading chain of length 985 involving 16 distinct wallets. The trading chain terminates when the final wallet in the chain, `0x186...0bf`, sells the shares back to the market at 2024-12-28 12:50:06 UTC. The trading takes place over a short duration (relative to the market duration), corresponding to the large spike in volume seen in Figure 25 in Appendix B (right column, second row).

			long					short			
block	index	timestamp	wallet_id	net_pos	price	type	shares	type	price	net_pos	wallet_id
65985710	75	2024-12-27 12:48:09	0xf0b...cab	95.01	0.014	buy	95	buy	0.986	-95	0x90b...c91
65985722	49	2024-12-27 12:48:35	0x90b...c91	0	0.985	sell	95	buy	0.985	-95	0xa44...b8f
65985728	51	2024-12-27 12:48:49	0xa44...b8f	0	0.985	sell	95	buy	0.985	-95	0xb5b...cd7
65985733	51	2024-12-27 12:48:59	0xb5b...cd7	0	0.985	sell	95	buy	0.985	-95	0x748...89e
65985739	74	2024-12-27 12:49:11	0x748...89e	0	0.985	sell	95	buy	0.985	-95	0xcce...190
65985745	104	2024-12-27 12:49:25	0xcce...190	0	0.985	sell	95	buy	0.985	-95	0xca8...7c4
		⋮					⋮				
66013305	77	2024-12-28 05:47:22	0x205...a60	0	0.985	sell	95	buy	0.985	-95	0x186...0bf
66024488	58	2024-12-28 12:50:06	0x186...0bf	-89.79	0.981	sell	5.21	buy	0.981	-5.21	0x88b...118
66024488	58	2024-12-28 12:50:06	0x186...0bf	0	0.981	sell	89.79	sell	0.019	0	0x810...1d8

Table 5: Chain-like trades in the market "Will the US add more than 300k jobs in December 2024?"

In Example 5, a wallet repeatedly trades with three counterparties at a time, while always maintaining a net position of at least 3,000 shares.

[21] See https://polymarket.com/profile/0x1c5fb90dddda9668522ffe94214c9808a28a5eb7 for the Polymarket profile page of nojkfaes, accessed July 26, 2025.

Example 5 (Will SPD, FDP, and Greens form the next German Government?)**.** Wallet `0xfd9...fe9` (with display name `monasa`) buys 15,169.29 shares from the market immediately preceding the first trade in Table 7.[22] It then sells 11,840 shares at $0.002 per share to three wallets, and subsequently buys back 4,000 shares at $0.001 per share from each of the same three wallets. It then repeatedly sells and buys back 12,000 shares 280 times over the course of several hours, in each iteration transacting with a new trio of colluders. By the end, `monasa` has traded 6.7M shares while maintaining a minimum position of at least 3,000 shares (note that it sheds 294.1 shares during this sequence due to two interceptions). Finally, it sells its entire position to the market over a period of 9 days. See the top cluster of Figure 40 in Appendix D. This same trading pattern is repeated in the market "Will CDU/CSU and Greens form the next German Government?", the only other market that `monasa` trades. The combined 1,811 counterparties that `monasa` trades with in the above manner all have a `created_at` timestamp between 2024-12-23 06:32:48 UTC and 2024-12-23 06:50:21 UTC.[23]

The next example illustrates complex patterns of staggered buy/sell trading among a large cluster of wallets.

Example 6 (Will the Denver Nuggets win the 2025 NBA Finals?)**.** As shown in Table 6, `0x24c...fd6` buys 15,500.2 shares with/from the market. It immediately sells 14,986.15 of these shares to `0x747...fec` and the remaining 514.05 shares to `0x093...5f4`, closing its position. `0x747...fec` buys additional shares from the market in the same trade in which it bought from `0x24c...fd6`. It similarly sells in a staggered manner, i.e., across consecutive trades, to close its position. A large set of approximately 1,200 colluding wallets continues this staggered buy/sell trading pattern for 5 days, from 2025-01-16 11:58:37 UTC until 2025-01-21 10:38:30 UTC. See the large cluster of wallets on the right side of Figure 42 in Appendix D.

[22] See https://polymarket.com/profile/0xfd9157825cf0319ec610970cf156ec4bb5008fe9 for the Polymarket profile page of monasa, accessed July 26, 2025.

[23] There are four additional wallets with large volumes which trade almost exclusively with the same counterparties as `monasa`, in different markets. They are included in the cluster labeled "monasa" in Table 10. Note that under our preferred parameters, Algorithm 2 (which we describe in Section 5) does not flag monasa's trades as wash trades; we discuss this limitation briefly in Section 7.

			long					short			
block	index	timestamp	wallet_id	net_pos	price	type	shares	type	price	net_pos	wallet_id
66767535	104	2025-01-16 11:58:37	0x24c...fd6	105	0.05	buy	105	buy	0.95	- 14 545.23	0xb7d...789
66767535	104	2025-01-16 11:58:37	0x24c...fd6	11 414	0.05	buy	11 309	sell	0.05	- 17 558.53	0xa2f...d22
66767535	104	2025-01-16 11:58:37	0x24c...fd6	15 493.6	0.05	buy	4079.6	sell	0.05	- 10 973.41	0x210...495
66767535	104	2025-01-16 11:58:37	0x24c...fd6	15 500.2	0.05	buy	6.6	buy	0.95	- 10 980.01	0x210...495
66767560	84	2025-01-16 11:59:31	0x747...fec	14 986.15	0.05	buy	14 986.15	sell	0.05	514.05	0x24c...fd6
66767560	84	2025-01-16 11:59:31	0x747...fec	15 906.55	0.05	buy	920.4	sell	0.05	- 11 900.41	0x210...495
66767560	84	2025-01-16 11:59:31	0x747...fec	15 934.2	0.05	buy	27.65	buy	0.95	- 11 928.06	0x210...495
66767603	108	2025-01-16 12:01:03	0x093...5f4	514.05	0.05	buy	514.05	sell	0.05	0	0x24c...fd6
66767603	108	2025-01-16 12:01:03	0x093...5f4	15 893.6	0.05	buy	15 379.55	sell	0.05	554.65	0x747...fec
66767650	78	2025-01-16 12:02:43	0x868...c68	554.65	0.05	buy	554.65	sell	0.05	0	0x747...fec
66767650	78	2025-01-16 12:02:43	0x868...c68	13 365	0.05	buy	12 810.35	sell	0.05	3083.25	0x093...5f4
66767650	78	2025-01-16 12:02:43	0x868...c68	18 066	0.05	buy	4701	sell	0.05	- 16 629.06	0x210...495
66767650	78	2025-01-16 12:02:43	0x868...c68	18 128	0.05	buy	62	buy	0.95	- 16 691.06	0x210...495
66767695	71	2025-01-16 12:04:17	0x536...c0e	3083.25	0.05	buy	3083.25	sell	0.05	0	0x093...5f4
66767695	71	2025-01-16 12:04:17	0x536...c0e	18 174.2	0.05	buy	15 090.95	sell	0.05	3037.05	0x868...c68
66767729	124	2025-01-16 12:05:31	0x955...877	15 484.55	0.05	buy	15 484.55	sell	0.05	2689.65	0x536...c0e
66767729	124	2025-01-16 12:05:31	0x955...877	18 521.6	0.05	buy	3037.05	sell	0.05	0	0x868...c68
66767775	58	2025-01-16 12:07:07	0xf01...7bf	2689.65	0.05	buy	2689.65	sell	0.05	0	0x536...c0e
		⋮					⋮				

Table 6: Staggered trades for Example 6, in the market “Will the Denver Nuggets win the 2025 NBA Finals?” Also see Figure 42.

			long					short			
block	index	timestamp	wallet_id	net_pos	price	type	shares	type	price	net_pos	wallet_id
70841377	88	2025-04-28 02:42:23	`0xdb1...991`	4000	0.002	buy	4000	sell	0.002	11 169.29	`0xfd9...fe9`
70841377	89	2025-04-28 02:42:23	`0x624...d36`	3840	0.002	buy	3840	sell	0.002	7329.29	`0xfd9...fe9`
70841377	90	2025-04-28 02:42:23	`0xdda...0bb`	4000	0.002	buy	4000	sell	0.002	3329.29	`0xfd9...fe9`
70841384	57	2025-04-28 02:42:37	`0xfd9...fe9`	7329.29	0.001	buy	4000	sell	0.001	0	`0xdb1...991`
70841385	48	2025-04-28 02:42:39	`0xfd9...fe9`	11 329.29	0.001	buy	4000	sell	0.001	0	`0xdda...0bb`
70841386	45	2025-04-28 02:42:41	`0xfd9...fe9`	15 329.29	0.001	buy	4000	sell	0.001	0	`0x624...d36`
		⋮					⋮				
70848306	76	2025-04-28 06:51:35	`0x39a...665`	4000	0.002	buy	4000	sell	0.002	11 035.19	`0xfd9...fe9`
70848306	78	2025-04-28 06:51:35	`0xb05...e89`	4000	0.002	buy	4000	sell	0.002	7035.19	`0xfd9...fe9`
70848306	79	2025-04-28 06:51:35	`0x4f7...5a7`	4000	0.002	buy	4000	sell	0.002	3035.19	`0xfd9...fe9`
70848314	60	2025-04-28 06:51:51	`0xfd9...fe9`	7035.19	0.001	buy	4000	sell	0.001	0	`0x4f7...5a7`
70848314	61	2025-04-28 06:51:51	`0xfd9...fe9`	11 035.19	0.001	buy	4000	sell	0.001	0	`0x39a...665`
70848315	68	2025-04-28 06:51:53	`0xfd9...fe9`	15 035.19	0.001	buy	4000	sell	0.001	0	`0xb05...e89`
		⋮					⋮				
71173307	36	2025-05-06 07:24:08	`0x702...3c5`	1100	0.001	buy	898.73	sell	0.001	0	`0xfd9...fe9`

Table 7: Trades in the market "Will SPD, FDP, and Greens form the next German Government?" Also see Figure 40.

				long					short			
block	index	timestamp	market_id	display_name	net_pos	price	type	shares	type	price	net_pos	display_name
71305824	86	2025-05-09 13:45:25	540415	`0x648...dc5`	97150.46	0.002	buy	28.06	buy	0.998	-28.06	Mazric
71305999	84	2025-05-09 13:51:35	541068	Mazric	77	0.54	buy	77	buy	0.46	-5012	`0x4da...eb5`
71306013	59	2025-05-09 13:52:05	521837	`0xa61...abd`	3649.22	0.008	buy	12.1	buy	0.992	-12.1	Mazric
71575738	58	2025-05-16 05:55:28	541473	Lanze	30128	0.41	buy	30128	buy	0.59	-30128	Mazric
		⋮										
71577703	74	2025-05-16 07:05:06	541473	Mazric	-3	0.58	sell	30125	sell	0.42	3	Lanze
71578287	58	2025-05-16 07:25:46	540818	Felvra	33568	0.48	buy	33568	buy	0.52	-33568	Mazric
		⋮										
71580005	65	2025-05-16 08:26:38	540818	Mazric	-3	0.53	sell	33565	sell	0.47	3	Felvra
		⋮										
71587566	73	2025-05-16 12:54:27	544800	Mazric	33970	0.56	buy	33970	buy	0.44	-33970	Therzia

Table 8: Trades for the wallet with display name "Mazric" across multiple markets. Mazric, Lanze, Felvra, and Therzia belong to the cluster labeled "Fantasy" in Table 10.

In the final example, we observe a pair of wallets which hold large complementary positions until market resolution; one wallet realizes a large profit, while the other realizes a large offsetting loss.

Example 7 (Mazric)**.** Table 8 shows a selection of trades for Mazric across multiple markets.[24] The first three trades appear to be legitimate trades with non-colluding wallets. Mazric then buys 30,128 shares with Lanze, a colluder with which Mazric then sells 30,125 shares, resulting in a small residual open position which is held to market resolution. Mazric performs several similar trades in different markets with other colluding wallets such as Felvra. In market 544800 ("Mets vs. Yankees"), Mazric buys 33,970 shares with Therzia. Mazric and Therzia both hold their positions until market resolution (less than 24 hours following the trade). While they have assumed no collective risk, Mazric realizes a large profit and Therzia realizes a large offsetting loss. The named wallets in this example are part of a large wash trading cluster whose constituent wallets all trade in a similar manner (see the "Fantasy" cluster in Table 10).

5 Detection Algorithms

As we can see from the above examples, many wash trading strategies generate significant volume without tying up capital for extended periods of time. This motivates us to consider instances in which wallets close their open positions, i.e., return to a close-to-zero net position and recoup their USDC collateral. Merely opening and closing a position is insufficient to identify a wash trade, as such actions are common in legitimate, profit-seeking strategies.[25] The fundamental distinction lies in a wash trader's intent to trade exclusively with wallets under their own control. This difference in counterparty selection becomes starkly visible when analyzing the trading network induced by all wallets' activities. Within this network, legitimate trading connects a diverse set of market participants, whereas wash trading carves out dense and mostly isolated subgraphs of self-dealing.

This is the main motivation behind our network-based wash trade detection algorithm. We assign an initial score to each wallet based on its position-closing propensity, and then iteratively update the score to a weighted average of the wallet's initial

[24]See `https://polymarket.com/profile/0xef665608a4f6d4cedc83feddd7aa25843cacb05c` for the Polymarket profile page of Mazric, accessed July 14, 2025.

[25]For example, an analysis of transaction-level data from Intrade referencing the winner of the 2012 US presidential election found that two of the largest traders by volume closed more than half of their open positions within the same second (Rothschild and Sethi, 2016). These traders were engaged in arbitrage, not wash trading.

score and its trading counterparties' scores. This iterative process provably converges to a unique set of final scores. We then classify trades between wallets with scores above a threshold as likely wash trades, based on a market-specific threshold chosen to minimize the fraction of "spillover" trades between estimated wash traders and estimated non-wash traders.

5.1 A General Detection Algorithm

Algorithms 1 and 2 specify our approach for estimating the extent of wash trading. Algorithm 1 consists of two parts—score initialization and iterative network-based score estimation—and outputs wallet scores as well as trades flagged using a fixed threshold θ. Algorithm 2 additionally performs market-specific threshold selection—which we consider a third part of our overall approach—and outputs trades flagged using these thresholds. (As we discuss at the end of this section, this modular structure allows one to modify or replace each part independently, as one sees fit.) We describe and discuss each part in turn, providing required definitions along the way.

Part I: Score Initialization. The first part of Algorithm 1 defines an initial score for each wallet which quantifies its propensity to close its positions.

We first give definition to the notion of "almost closing" a position, which will admit robustness against wash trading behavior like that in Example 7. To keep notation concise we fix the market m. Each wallet i in the data for this market is associated with a finite sequence of trades, and each of these trades either increases or decreases its net long position. Some transactions can lead to a reversal of position (long to short or vice versa) and for ease of analysis we split these into two trades, one of which takes the position to zero while the other opens up a position in the opposite direction. Under this convention, all pairs of consecutive post-trade positions are either non-negative or non-positive. We refer to the t-th trade in the chronological sequence of trades as simply trade t.

With this in mind, let $P_{i,t}$ represent wallet i's net long position after trade t, and let $X_{i,t} := |P_{i,t}|$. We say that trade t is an *expansion* if $X_{i,t} > X_{i,t-1}$, and a *contraction* if $X_{i,t} < X_{i,t-1}$. We say that trade t is a *terminal contraction* if it is a contraction and is either immediately followed by an expansion, or is the last trade in this market for wallet i. A terminal contraction is a *closure* if it is at least a $100(1-c)\%$ drawdown from the maximum position obtained following the previous closure (or the beginning of i's trading).

Closures can be identified iteratively starting with the first. To make this precise, set $\tau_{i,0} = 0$ and let $\tau_{i,\ell}$ denote the trade corresponding to the ℓth closure. Then $\tau_{i,\ell}$

must be a terminal contraction that satisfies

$$X_{i,\tau_{i,\ell}} \leq cM_{i,\ell}$$

where

$$M_{i,\ell} = \max\{X_{i,t} \mid \tau_{i,\ell-1} < t < \tau_{i,\ell}\}.$$

The first closure is then the first terminal contraction that satisfies the above conditions, and successive closures can be identified iteratively. For our implementation below, we choose $c = 0.005$.

Define $Q_i^m := \sum_\ell \mathbb{1}\{\tau_{i,\ell}^m < \infty\}$, the number of times that i closed a position in market m according to our above criterion, and let $\boldsymbol{x}^{(0)}$ be the vector with

$$x_i^{(0)} = \sum_m s_i^m \mathbb{1}\{Q_i^m > 0\},$$

where s_i^m is the share of i's total volume traded that was traded in market m. As discussed above, the vector $\boldsymbol{x}^{(0)}$ serves as an initialization capturing wallets' tendency to close their positions. (We weight by volume traded to be robust against strategies—as in Example 7—in which a wallet attempts to disguise wash trading by making small, legitimate trades in multiple markets.)

Part II: Iterative Network-Based Score Estimation. Part II of Algorithm 1 is our iterative procedure for updating wallets' scores. Let $\boldsymbol{B}$ be the volume-weighted adjacency matrix, where element b_{ij} is the fraction of wallet i's share volume *across all markets* where wallet j is the counterparty, among i's counterparties $N(i)$.[26] The score update for iteration k is

$$x_i^{(k)} \leftarrow \frac{1}{2}\left\{x_i^{(0)} + \sum_{j \in N(i)} b_{ij}\, x_j^{(k-1)}\right\}.$$

We iterate until the score vector $\boldsymbol{x}^{(k)}$ converges, with a relative ℓ_2 tolerance $tol = 10^{-5}$. For a fixed threshold $\theta \in [0, 1]$, we would flag trades between two wallets i and j as likely wash trades if both wallets' final scores exceed θ. However, our approach also calculates market-specific thresholds in Algorithm 2. Before describing this calculation in detail, we first discuss several properties of the score estimation step.

[26]One can view $\boldsymbol{B}$ as a simple attention mechanism in the context of graph neural networks; see the section on soft edge selection in Qiao et al. (2025).

Algorithm 1: Wash Trade Detection (Fixed Threshold)

Input: Wallets $\mathcal{W}$; Adjacency matrix $\boldsymbol{B}$; Tolerance tol; Threshold θ.
Output: Converged wallet scores; Set of flagged trades.

```
/* --- Part I: Score Initialization ---                                          */
```

foreach $i \in \mathcal{W}$ **do**
    $x_i^{(0)} \leftarrow \sum_m s_i^m \mathbb{1}\{Q_i^m > 0\}$;
end

```
/* --- Part II: Iterative Network-Based Score Estimation ---                     */
```

$k \leftarrow 0$;
repeat
    $k \leftarrow k+1$;
    foreach $i \in \mathcal{W}$ **do**
        $x_i^{(k)} \leftarrow \frac{1}{2}\left\{x_i^{(0)} + \sum_{j \in N(i)} b_{ij}\, x_j^{(k-1)}\right\}$;
    end
until $\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k-1)}\|_2 / \|\boldsymbol{x}^{(k-1)}\|_2 < \texttt{tol}$;
$\boldsymbol{x} \leftarrow \boldsymbol{x}^{(k)}$;
//
return $\boldsymbol{x}$, Trades between wallets i and j with $\min\{x_i, x_j\} \geq \theta$.

Algorithm 2: Wash Trade Detection (Market-Specific Threshold)

Input: Wallet scores $\boldsymbol{x}$; Markets $\mathcal{M}$; Market data $\{\mathcal{W}(m), N_m(i), \mathcal{P}(m), v_{ij,m}\}_{m \in \mathcal{M}}$; Threshold parameters $\underline{\theta}$, $\overline{\theta}$, $\overline{Y}$, ζ.
Output: Set of flagged trades.

```
/* --- Part III: Market-Specific Threshold Selection ---                         */
```

foreach $m \in \mathcal{M}$ **do**
    foreach $i \in \mathcal{W}(m)$ **do**
        $r_{im} \leftarrow \min\left\{x_i,\ \max_{j \in N_m(i)} x_j\right\}$;
    end
    Define $Y_m(\theta) := 1 - \frac{\sum_{(i,j) \in \mathcal{P}(m)} v_{ij,m} \mathbb{1}\{\min\{r_{im}, r_{jm}\} \geq \theta\}}{\sum_{(i,j) \in \mathcal{P}(m)} v_{ij,m} \mathbb{1}\{\max\{r_{im}, r_{jm}\} \geq \theta\}}$;
    Define $\Theta_m(\underline{\theta}, \overline{\theta}, \overline{Y}) := \left\{\theta \in \{r_{im}\}_{i \in \mathcal{W}(m)} \mid \theta \in [\underline{\theta}, \overline{\theta}] \text{ and } Y_m(\theta) \leq \overline{Y}\right\}$;
    if $\Theta_m(\underline{\theta}, \overline{\theta}, \overline{Y})$ *is not empty* **then**
        $\theta_m^* \leftarrow \min\left(\operatorname{argmin}_{\theta \in \Theta_m(\underline{\theta}, \overline{\theta}, \overline{Y})} \max\{\zeta,\ Y_m(\theta)\}\right)$;
    end
    else
        $\theta_m^* \leftarrow 1$;
    end
end
return For all $m \in \mathcal{M}$, trades between wallets i and j with $\min\{x_i, x_j\} \geq \theta_m^*$.

As we show in Appendix C (see Proposition 1), the sequence of score vectors $\{\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \dots\}$ converges to the $\boldsymbol{x}$ which satisfies the stationary equation

$$\boldsymbol{x} = \frac{1}{2}(\boldsymbol{x}^{(0)} + \boldsymbol{B}\boldsymbol{x}) \tag{1}$$

or

$$(\boldsymbol{I} - \frac{1}{2}\boldsymbol{B})\boldsymbol{x} = \frac{1}{2}\boldsymbol{x}^{(0)}. \tag{2}$$

That is, $\boldsymbol{x}$ gives one-half weight to $\boldsymbol{x}^{(0)}$ and one-half weight to a volume-weighted average of each wallet's counterparties' scores.[27] Since $\boldsymbol{B}$ is a row-stochastic matrix with diagonal elements equal to zero, all its eigenvalues are contained within the unit disc, such that $\boldsymbol{I} - \frac{1}{2}\boldsymbol{B}$ is invertible and a unique solution vector $\boldsymbol{x}$ exists.[28]

We further show in Appendix C (see Proposition 2) that the volume-weighted score $\boldsymbol{v}^\top \boldsymbol{x}^{(k)}$—where element v_i of $\boldsymbol{v}$ is wallet i's total trading volume—is conserved across the iterations of Algorithm 1. (The overall volume-weighted average score in the data is $\sum_i v_i x_i / \sum_i v_i = 0.747$.) Thus, one may view the detection algorithm as *redistributing* the initial scores $\boldsymbol{x}^{(0)}$ to clusters of wallets which collectively have a high propensity to close their positions.

Finally, we note that our formulation has an analogue in the literature on non-Bayesian social learning, namely in Friedkin and Johnsen (1999)'s extension of the classic DeGroot (1974) model for opinion dynamics. In the Friedkin-Johnsen model, agents have varying susceptibility to influence by neighbors (or, conversely, stubbornness with respect to their initial beliefs), parametrized by a diagonal matrix $\boldsymbol{A}$; here, we have $\boldsymbol{A} = \frac{1}{2}\boldsymbol{I}$.[29]

Part III: Market-Specific Threshold Selection. In Algorithm 2, we compute market-specific thresholds for more robust estimation of likely wash volume. Let x_i be wallet i's converged score (an output of Algorithm 1), and let $\mathcal{W}(m)$ be the set of wallets in market m. Define r_{im} to be the largest threshold at which wallet $i \in \mathcal{W}(m)$ would be flagged as having a wash trade in market m, i.e.,

$$r_{im} := \max\left\{\theta \mid \min\{x_i, x_j\} \geq \theta \text{ for some } j \in N_m(i)\right\}, \tag{3}$$

[27] More generally, one could assign weight $\rho \in (0,1)$ to $\boldsymbol{x}^{(0)}$ and weight $1-\rho$ to the average counterparty score. We tentatively choose $\rho = \frac{1}{2}$ as a default value, though it remains of interest to select this parameter in a more systematic way.

[28] That $|\lambda| \leq 1$ for all eigenvalues λ of $\boldsymbol{B}$ is a consequence of the Gershgorin Circle Theorem.

[29] An analysis of the Friedkin-Johnsen model is given by Parsegov et al. (2017).

where $N_m(i)$ is the set of i's counterparties in market m. It is easy to see that

$$r_{im} = \min\left\{x_i, \max_{j \in N_m(i)} x_j\right\}. \tag{4}$$

For threshold $\theta \le \max_i r_{im}$, we define the *relative spillover* from estimated wash traders (wallets with $r_{im} \ge \theta$) to non-wash traders in market m as

$$Y_m(\theta) := 1 - \frac{\sum_{(i,j)\in\mathcal{P}(m)} v_{ij,m} \mathbb{1}\{\min\{r_{im}, r_{jm}\} \ge \theta\}}{\sum_{(i,j)\in\mathcal{P}(m)} v_{ij,m} \mathbb{1}\{\max\{r_{im}, r_{jm}\} \ge \theta\}}, \tag{5}$$

where $\mathcal{P}(m)$ is the set of trading wallet pairs in market m, and $v_{ij,m}$ is the total share volume traded between wallets i and j in market m. Motivated by the observation that wash traders tend to trade among themselves, $Y_m(\theta)$ is a measure of how well θ separates estimated wash traders from non-wash traders, with smaller $Y_m(\theta)$ indicating better separation. We tailor the threshold to the market by choosing θ^*_m which minimizes $Y_m(\theta)$ within a range $[\underline{\theta}, \overline{\theta}]$, provided that some $Y_m(\theta)$ is smaller than a maximum acceptable value $\overline{Y}$.[30] We denote this feasible set by

$$\Theta_m(\underline{\theta}, \overline{\theta}, \overline{Y}) := \left\{\theta \in \{r_{im}\}_{i \in \mathcal{W}(m)} \mid \theta \in [\underline{\theta}, \overline{\theta}] \text{ and } Y_m(\theta) \le \overline{Y}\right\}. \tag{6}$$

We include a slack parameter $\zeta \ge 0$ in the objective to accommodate spillovers resulting from interceptions or intentional trades with "the market", i.e., non-colluders. If there are multiple θ which minimize the objective, we select the smallest, corresponding to the largest estimated wash volume; on the other hand, if $\Theta_m(\underline{\theta}, \overline{\theta}, \overline{Y})$ is empty, we set $\theta^*_m = 1$, corresponding to no detected wash volume (though one could equally well choose any $\theta_m > \max_{i \in \mathcal{W}(m)} r_{im}$):

$$\theta^*_m := \begin{cases} \min \underset{\theta \in \Theta_m(\underline{\theta}, \overline{\theta}, \overline{Y})}{\operatorname{argmin}} \max\{\zeta, Y_m(\theta)\} & \text{if } \Theta_m(\underline{\theta}, \overline{\theta}, \overline{Y}) \text{ is not empty} \\ 1 & \text{otherwise.} \end{cases} \tag{7}$$

For our implementation, we set $\zeta = 0.001$, $\underline{\theta} = 0.8$, $\overline{\theta} = 0.99$, and $\overline{Y} = 0.1$. Trades between two wallets i and j in market m are flagged as likely wash trades if both wallets' final scores exceed θ^*_m.

Example 1 (continued)**.** While wallets' score updates make use of information from *all* markets, we can visualize the progression of Algorithm 1 in individual markets;

[30] We specify a lower bound $\underline{\theta}$ since, for a small enough threshold θ', all trades are flagged as wash trades, and $Y_m(\theta') = 0$.

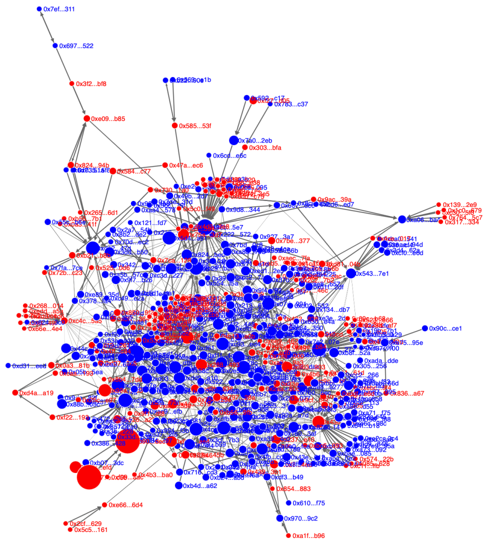
an example is shown in Figure 5. From the collection of trades for the market "Will the Republican candidate win Pennsylvania by 1.0%–1.5%", we first construct the market's *trade graph*, where nodes represent wallets and directed edges represent trades between wallets.[31] At initialization, 32.5% of 453 wallets have a score $x_i \geq 0.9$; by iteration $k = 3$, only 13.7% of wallets do, and only five of those wallets (shown in red) also have a counterparty with score $x_j^{(3)} \geq 0.9$. The coloring of the graph in panel (d) remains unchanged in subsequent iterations, and is the same as that under the market-specific threshold $\theta_m^* = 0.9152$ selected by Algorithm 2 (with associated spillover $Y_m(\theta_m^*) = 0.0217$). The red nodes in the bottom left are precisely the wallets MAY20, MAY175, and MAY176 whose trades are shown in Table 2. These trades account for 56.8% of the total share volume in the market (73.4% of the total dollar volume).

The strength of our approach is evident in the additional examples shown in Appendix D, in which our detected wash trades often correspond to anomalous spatial patterns in the trade graphs (which can be seen independently from our algorithm; see footnote 31).[32] Our score update rule based on a wallet's activity in *all markets* reduces the incidence of false positives (i.e., classifying authentic trades as wash trades), compared to a rule which updates scores based only on a wallet's market-specific activity (and which would assign each wallet a market-specific score). This is because it is possible, especially in thin markets with few traders, that two non-affiliated wallets could incidentally open and close positions against one another. But it is unlikely that this could happen in all markets that a wallet trades in without explicit coordination.

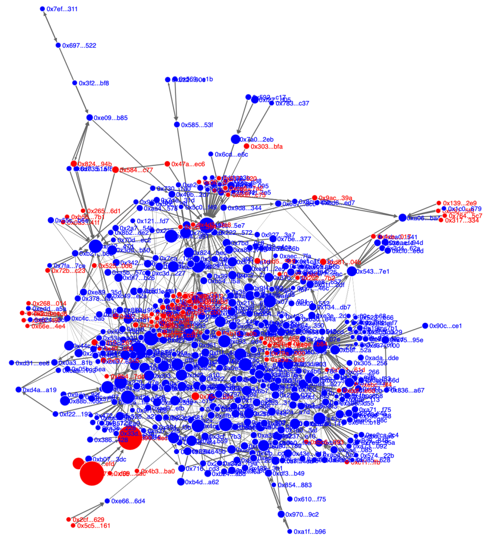
We note that the three parts of Algorithms 1 and 2 are modular, and can be modified or replaced independently. As one example, in place of our current Part II one

[31] We plot these graphs using the R package networkD3, an API for the Javascript visualization library D3.js. We use the *d3-force* module, which produces graph layouts by simulating physical forces among nodes and edges; in many cases, this results in segregated layouts that make wash-trading activity discernible to the eye. See `https://github.com/christophergandrud/networkD3` and `https://d3js.org/d3-force`, accessed July 14, 2025.

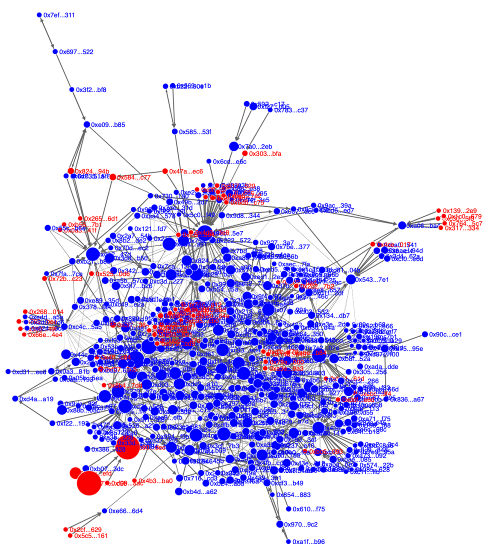
[32] In some markets, we observe trades between many—hundreds, sometimes thousands—of "sybil" wallets and a small number of wallets which appear to be legitimate market makers. This pattern is visually prominent in Figures 45 and 48, and several other example graphs. Such trades typically involve the sybil wallets each buying a small number of shares from makers (in markets with a price increment of \$0.001), and selling back to the makers at a slightly lower price to realize a small per-share loss. In Figure 45, there are 3,866 sybils which buy at \$0.996 and sell at \$0.995, counterparty to 12 market makers which buy at \$0.004 and sell at \$0.005; these trades account for 4.6% of the total market share volume. While such trades may be regarded as inauthentic, they do not necessarily require explicit coordination between sybils and market makers. (There are other cases, including Example 5, in which the trades do appear to be coordinated among a cluster of largely self-trading wallets.) Under our chosen parameters, Algorithm 2 generally does not classify such trades as wash trades, though the prevalence of these trades is suggested by Figure 4 and Figure 21.



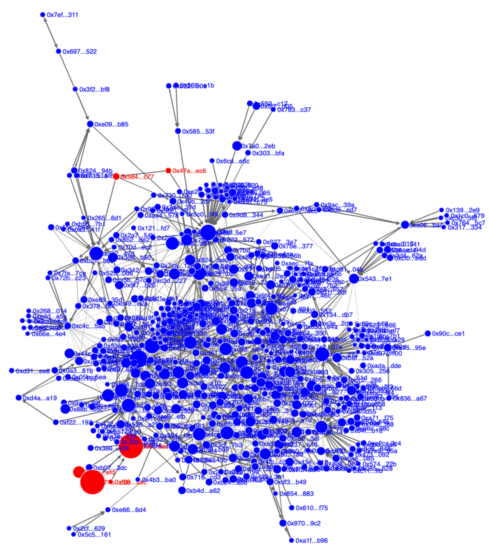
(a) $k = 0$ (initialization).



(b) $k = 1$.



(c) $k = 2$.



(d) $k = 3$. (*)

Figure 5: The progression of the iterative score updates in Algorithm 1 for the market "Will the Republican candidate win Pennsylvania by 1.0%–1.5%?", discussed in Example 1. Each node i is a wallet, with circular area proportional to $\sqrt{\text{share volume}_i}$, and an edge with weight $w_{ij} \in (0, 1]$ indicates that a fraction w_{ij} of i's share volume traded was traded with counterparty j in this market. Edge thickness is proportional to $\exp\{w_{ij}^2\}$. In panels (a)–(c), wallets with a kth iteration score $x_i^{(k)} \geq 0.9$ are colored red; all other wallets are colored blue. In panel (d), only neighboring wallets i and j with $\min\{x_i^{(3)}, x_j^{(3)}\} \geq 0.9$ are colored red. Trades by the red wallets in the bottom left of (d)—MAY20, MAY175, and MAY176—are shown in Table 2.

could conceivably design and use a "leave-one-out" message-passing algorithm, where the message from node i to j at iteration k does not depend on the messages received by i from j thus far (i.e., up to iteration $k-1$). After all, the past messages sent by node j (whose job it is to estimate whether trader j is a wash trader) to node i should not influence what j is learning from i. Leave-one-out algorithms—like max-product and sum-product belief propagation in graphical models—have well-documented theoretical and practical advantages in many settings (e.g., see Kschischang et al., 2002; Lu et al., 2008; Donoho et al., 2010; Karger et al., 2011; Murphy et al., 2013).

5.2 Specialized Detection Algorithms

Our general detection algorithm does not distinguish between different classes of wash trading strategies. Here we describe several specialized methods, each of which is designed to detect a specific type of wash trade such as those identified in the examples in Section 4.1. The types of trades flagged by these methods are mutually exclusive but non-exhaustive: there is significant volume detected by Algorithm 2 that is not captured by this classification. In Section 6, we examine the prevalence of these different types of wash trades over time, and the overlap between these trades and the trades identified by our general detection algorithm.

To place the magnitudes below in context, we note that the total share volume that represents positions which are opened and drawn down to any extent is 21.5B shares, or 74.4% of total exchange volume. Positions which are opened and subsequently closed account for 19.4B shares, or 66.8% of total exchange volume.[33]

Dyadic Wash Trades We consider instances in which wallets i and j buy together and subsequently sell together, or in which i buys from j and j buys from i, in the same market.[34] As in Example 2, wallets may split volume across multiple trades, so it does not suffice to restrict to trades with identical sizes. Instead, similar to our calculation of the net long position of each wallet, we calculate the net long position of a wallet *vis-à-vis each of its counterparties* in market m.[35] Let $P_{ij,t}$ represent wallet

[33] Recall that we refer to a position as closed if there is at least a $100(1-c)\%$ drawdown from the maximum position size obtained following the previous closure.

[34] If i and j are known to be wallets with the same owner, then any trade between i and j is a wash trade. In establishing our upper bound, we ignore this possibility as it would lead to the vacuous conclusion that all trades could be wash trades, if one cannot rule out that all wallets are mutually affiliated. In other words, our definition of dyadic trading (for the purpose of computing this upper bound) deliberately excludes cases in which, e.g., i and j buy together and hold shares to market resolution.

[35] For example, if i buys 500 shares (long) with j (short), then i increases its net long position against j by 500.

i's net long position against wallet j after trade t, and let $X_{ij,t} := |P_{ij,t}|$. An upper bound on the amount of dyadic wash trading in market m is given by

$$\sum_{(i,j)\in\mathcal{P}(m)} \sum_{t\in T(i)\cap T(j)} 2\cdot \mathbb{1}\{X_{ij,t} < X_{ij,t-1}\}|X_{ij,t} - X_{ij,t-1}| \tag{8}$$

where $\mathcal{P}(m)$ is the set of trading wallet pairs in market m. That is, there is a potential wash trade each time i reduces its absolute net position vis-à-vis j, with multiplication by two to account for both the opening and closing of the position.[36] Summing over all markets gives 5.6B shares (19.3% of total share volume) as the overall upper bound on dyadic wash volume.

We can narrow our estimate by further considering the timing of trades. Likely dyadic wash trades involve pairs of wallets which open and (almost) close positions against each other in short time intervals.[37] This criterion based on pairwise position closures accounts for 4.23B shares traded, or 14.6% of total share volume ($2.26B, or 14.5% of dollar volume), when no time limit is imposed and 2.16B shares traded, or 7.5% of total share volume ($936M, or 6.0% of dollar volume), with a time limit of 180 seconds. Note that the estimate with no time limit is smaller than the upper bound, since (8) includes cases in which i only partially closes its position against j, or vice versa.

Triangular Wash Trades We consider trades like the one depicted in Example 3, in which i buys with j; j sells to k; and i sells with k in the same market. To construct an upper bound on the total triangular trade volume in each market, we first formalize this trading pattern by introducing the following notation for *directed trade relations*:

- $u \upuparrows v$: u (long) buys with v (short);
- $u \rightarrow v$: u (long) sells to v (short);
- $u \leftarrow v$: v (short) sells to u (long);
- $u \downdownarrows v$: u (long) sells with v (short).

Next, we define a *triangular motif* as a tuple of trade relations (α, β, γ) of the form $(i \upuparrows j,\ j \rightarrow k,\ k \downdownarrows i)$ or $(i \upuparrows j,\ k \leftarrow i,\ j \downdownarrows k)$. Letting T_α denote the set of trade indices (ordered by time) for trades described by relation α, we define such a motif

[36] Recall that we split trades which lead to position reversals into separate trades, such that $P_{ij,t}P_{ij,t-1} \geq 0$ for all pairs of consecutive trades between wallets i and j.

[37] Note that this will exclude cases in which i does not fully close its position vis-à-vis j, including instances in which i does so strategically to evade detection, or when i's order is partially intercepted by a third wallet k.

to be "weakly time-compatible" if

$$\min_{t \in T_\alpha} t < \min \left\{ \max_{t \in T_\beta} t,\ \max_{t \in T_\gamma} t \right\}$$

and

$$\min_{t \in T_\beta} t < \max_{t \in T_\gamma} t.$$

We use this condition to rule out cases in which, e.g., all "$j \to k$" trades occur before all "$i \upuparrows j$" trades, which precludes the existence of triangular trades with an $(i \upuparrows j,\ j \to k,\ k \downdownarrows i)$ motif. (The condition is otherwise very relaxed with respect to the timing of trades, and is defined as it is for bounding purposes only.) Finally, letting Δ be the set of all weakly time-compatible triangular motifs present in market m; x_s, the number of shares exchanged in $s \in \Delta$; and $V_\alpha := \sum_{t \in T_\alpha} v_t$, the total share volume associated with relation α, we solve the following linear program:

$$\begin{aligned}
\text{maximize} \quad & 3 \sum_{s \in \Delta} x_s \\
\text{subject to} \quad & \sum_{s \in \Delta : u \upuparrows v} x_s \leq V_{u \upuparrows v} \quad \forall (u, v) \\
& \sum_{s \in \Delta : u \to v} x_s \leq V_{u \to v} \quad \forall (u, v) \\
& \sum_{s \in \Delta : u \leftarrow v} x_s \leq V_{u \leftarrow v} \quad \forall (u, v) \\
& \sum_{s \in \Delta : u \downdownarrows v} x_s \leq V_{u \downdownarrows v} \quad \forall (u, v).
\end{aligned} \tag{9}$$

Summing the objective value over all markets gives 649M shares (2.2% of total share volume) as the overall upper bound for triangular wash volume.

Similar to the case of dyadic trades, we narrow this estimate by considering the timing of trades. Without aggregating pairwise volume as we did to construct the upper bound, we restrict our attention to buy/buy, buy/sell and sell/sell trades that constitute a triangular trade and are within 180 seconds of each other. This criterion accounts for 100M shares, or 0.34% of total share volume ($86.5M, or 0.56% of total dollar volume).[38]

[38]For computational tractability, we temporally partition several large markets—"Will Donald Trump (Kamala Harris) win the 2024 US Presidential Election?" and "Will Donald Trump be inaugurated?"—and calculate wash volume separately for each part. Specifically, we split each market into 20 parts such that each part has trades, in sequence, which account for 5% of the market's total share volume. We also do this for the calculation of chain and cluster volume described below.

Chain Wash Trades We consider simple trading chains like the one depicted in Example 4, in which i sells to j, j sells to k, k sells to l, and so on. To identify these chains in each market m, we restrict attention to the subgraph induced by wallets which have exactly one counterparty in each direction—i.e., one upstream seller and one downstream buyer—with the two counterparties distinct. We then find connected components comprising at least three wallets where the coefficient of variation of trade sizes (in shares) is no more than 0.1. Summing over all markets, we estimate that chain wash trades represent 306M shares, or 1.1% of total share volume ($31.6M, or 0.20% of total dollar volume).

Cluster Wash Trades We consider wash trades of the variety in Example 6, in which wallets buy and sell against multiple counterparties, perhaps deliberately to avoid closing positions pairwise. Similar to our approach for chains above, we consider buy/sell trades and restrict attention to the subgraph induced by wallets which have more than one counterparty in at least one direction in market m. We find connected components of at least four wallets for which the coefficient of variation of trade sizes (in shares) is no more than 0.1. Summing over all markets, we estimate that cluster wash trades represent 1.37B shares, or 4.74% of total share volume ($67M, or 0.43% of total dollar volume).

6 Results

We now provide estimates of the amount of wash trading on Polymarket resulting from the application of Algorithms 1 and 2 to the complete history of trades from November 21, 2022 to October 12, 2025. At tolerance level $tol = 10^{-5}$ with parameter $c = 0.005$ (for determining position closures), Part I of our algorithm (wallet score estimation) converged in 12 iterations.

Figure 6 shows the fraction of Polymarket's overall historical share volume that is classified as wash trading by Algorithm 1, as a function of a *fixed* threshold parameter θ. At a very conservative threshold of $\theta = 0.99$, 18.0% of all historical share volume is generated by trades between wallets whose scores are both at least θ. Intuitively, this high threshold flags wallets that have a strong tendency to close their positions, and which trade heavily with counterparties that also have a strong tendency to close their positions. Such associations should be regarded as highly irregular, especially when traders should (in theory) be agnostic about the identity of their counterparties.

In comparison, Algorithm 2—which selects market-specific thresholds to minimize our spillover metric in (5)—flags 24.2% of historical share volume as likely wash trad-

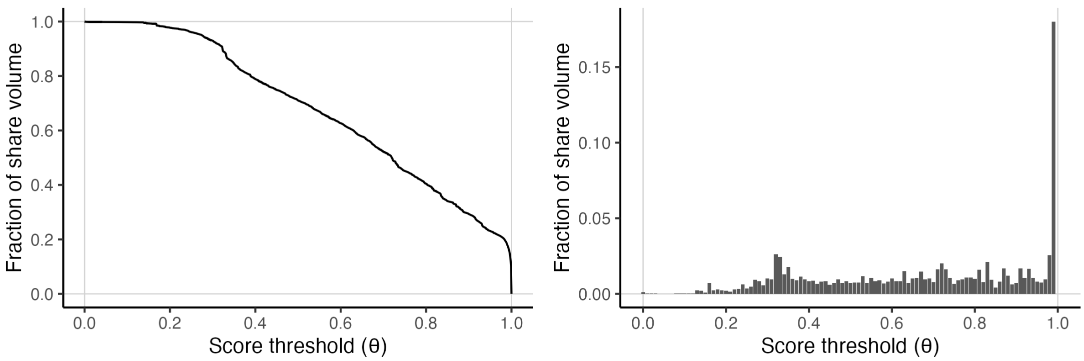


Figure 6: Left: The overall fraction of Polymarket's 29.0B share volume (through October 12, 2025) classified as wash trading by Algorithm 1 using a *fixed* threshold, as a function of the threshold θ. Right: The incremental fraction of overall share volume classified as wash (for increasing θ in increments of 0.01).

ing. (The same aggregate fraction is flagged by Algorithm 1 under a fixed threshold $\theta = 0.939$.) Analogous to Figure 6, Figure 28 in Appendix B shows the sensitivity of this aggregate estimate to the threshold parameters $(\underline{\theta}, \overline{\theta}, \overline{Y}, \zeta)$. We observe that our estimate is insensitive to the choice of ζ and $\underline{\theta}$, and moderately sensitive to the choice of $\overline{\theta}$ and $\overline{Y}$. We choose the latter two parameters to give reasonable performance on a set of example markets and the high-volume markets in Table 13 (in terms of flagging trades corresponding to anomalous clusters of wallets in these markets' trade graph visualizations; see footnote 31).

Figure 7 shows the fraction of weekly exchange volume that is classified as wash trading from January 2024 to October 2025. The temporal pattern in the detected volume offers guidance on the interpretation of different fixed thresholds (for Algorithm 1): in particular, at $\theta \geq 0.8$ there is essentially no detected wash trading prior to June 2024, and at $\theta \geq 0.9$ there is a period from May–September 2025 where there is relatively little detected wash trading. Such a pattern would be unlikely if Algorithm 1 were detecting authentic trading volume, hence $\theta = 0.9$ may be regarded as a strict, or conservative, threshold. Algorithm 2 generally produces lower estimates of wash volume, and therefore may be regarded as even more conservative than Algorithm 1 with $\theta = 0.9$. Using Algorithm 2, we find that detected wash trading peaked in December 2024 at approximately 60% of overall weekly exchange volume. From June until late September 2025, detected wash trading accounted for less than 5% of weekly volume (this may be because Polymarket made efforts to curb wash trading, or because wash-trading wallets no longer close their open positions or trade exclusively with each other; we discuss this possibility further in Section 7). In the first week of October 2025, the detected wash fraction again increased sharply to about 20% of weekly volume.

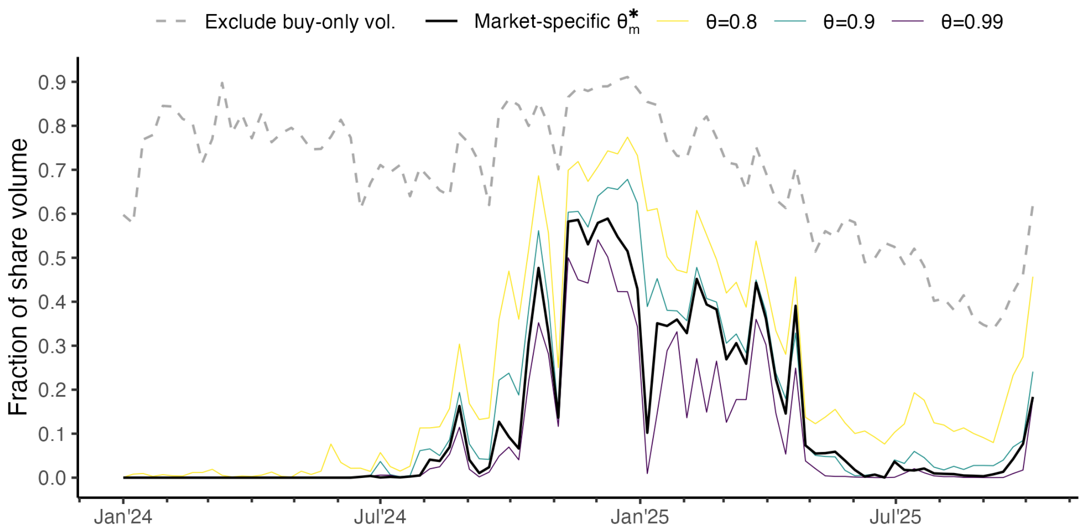


Figure 7: The weekly fraction of Polymarket share volume classified as wash volume by Algorithm 2 (black) and, for reference, by Algorithm 1 under different fixed thresholds θ. The dashed gray line is the fraction of total share volume which excludes volume from wallets which *only buy* shares in a given market (using the complete set of trades for the market).

Figure 7 also shows the fraction of weekly share volume which excludes volume generated by wallets which *only buy* (i.e., never sell) in a given market. While the set of trades by such "buy-only" wallets is not necessarily disjoint with all wash trades—in Example 7, colluding wallets buy shares together and hold them to market resolution—a wash trading strategy which involves only buying would generally be a capital-inefficient way to generate high volume.[39] (Figure 27 in Appendix B shows that no more than about 2% of the weekly buy-only volume is classified as wash trading by Algorithm 2.) The decrease in the detected wash fraction is accompanied by a decrease in the "exclude buy-only" fraction of volume, and the latter decreases to a level lower than that which prevailed before the surge in detected wash trading in the second half of 2024, before increasing significantly starting in mid-September 2025.[40]

Figure 8 shows the cumulative distribution function of the fraction of share volume classified as wash trading, by market and by wallet. Using Algorithm 2 we find that, across the entire sample period, fewer than 10% of markets have any amount of wash

[39] It is possible, however, to devise an effective buy-only wash trading strategy: one could buy shares with a colluder and hold to resolution in markets which are expected to resolve shortly. Alternatively, one could use wallets A and B to buy "Yes" and "No" shares, transfer B's "No" shares to A, and then merge A's "Yes" and "No" shares to recover collateral.

[40] Note that the "exclude buy-only" fraction may be underestimated in recent months due to the possibility of censoring in active markets, as some wallets which have so far been observed to have only bought shares may sell them in the future. However, we find that the magnitude of this censoring bias has been small in the past.

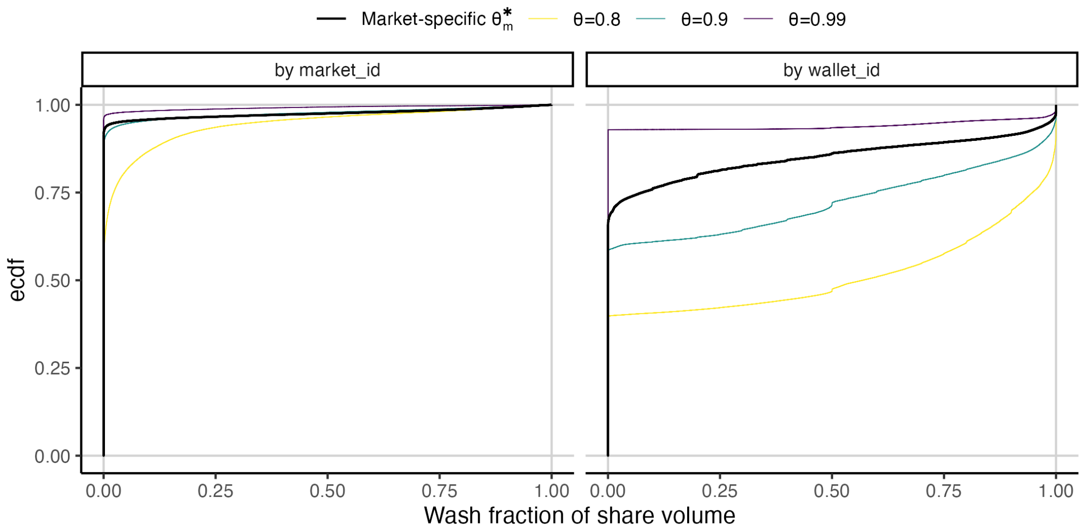


Figure 8: CDF of the fraction of share volume classified as wash trading by Algorithm 2 (black)—and, for reference, by Algorithm 1 under different fixed thresholds θ—by market (left) and by wallet (right).

trading, while about 30% of wallets do (14% of wallets have more than half of their volume classified as wash).

Figure 30 in Appendix B shows the estimated wash fraction of share volume by event category. (The largest categories by cumulative volume to date are Crypto, Elections, Politics, and Sports, which have each traded more than \$2.0B in dollar volume.) Wash trading is highly prevalent across all event categories at some point in the sample period. Detected wash trading peaked at 21% of the weekly share volume in Crypto markets (week of October 6, 2025); 95% in Election markets (week of March 24, 2025); 60% in Politics markets (week of December 23, 2024); and 90% in Sports markets (week of October 21, 2024).

Table 13 shows the estimated wash fraction of share volume for the 50 largest markets by share volume. Most of these markets have either a high fraction (≥ 0.8) or a low fraction (≤ 0.2) of detected wash volume. Notably, Algorithm 2 does not detect wash trades in the three largest markets, "Will Donald Trump (Kamala Harris) win the 2024 US Presidential Election?" and "Will Donald Trump be inaugurated?", as none of these markets can be assigned a threshold $\theta_m \in [\underline{\theta}, \overline{\theta}]$ which satisfies our spillover criterion $Y_m(\theta) \leq \overline{Y}$.[41] On the other hand, 98.5% of volume in "Will Nicolae Ciucă win the 2024 Romanian Presidential election?"—which traded only \$2.6M in

[41] Recall that our chosen threshold parameters for Algorithm 2 are $(\underline{\theta}, \overline{\theta}, \overline{Y}) = (0.8, 0.99, 0.1)$. We also note that our estimates contrast with those of crypto research firm Chaos Labs, which "concluded that around one-third of trading volume—and overall users—on the presidential market alone was likely wash trading, along with across all markets", as reported in Schwartz (2024).

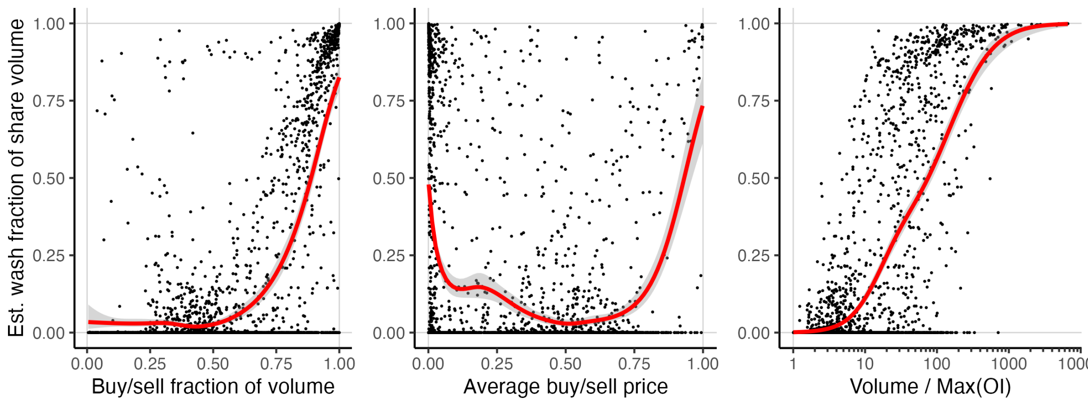


Figure 9: Univariate predictors of the fraction of wash volume detected by Algorithm 2 in each market, for markets with total volume at least 10^6 shares traded (2.8% of all markets): buy/sell fraction of share volume (left); the share-weighted average buy/sell price (middle); and the ratio of share volume to maximum open interest, or "speculative ratio" (right). Curves are fit using a generalized additive model (GAM) smoother with logit link.

dollar volume but is the fifth largest market by share volume—is classified as wash trading.[42]

In the Nicolae Ciucă market, nearly 60% of shares traded were traded in buy/sell trades (as opposed to buy/buy or sell/sell), with a share-weighted average buy/sell trade price of $0.00147. The left and middle panels of Figure 9 illustrate that markets with an unusually high fraction of buy/sell volume, and a high or low average buy/sell price, tend to have high amounts of detected wash trading. As we note in Section 3, if the goal of a wash trader is to generate large share volume (as opposed to dollar volume), then the most efficient trading strategy—in terms of requiring the least capital—is to buy and trade cheap shares at a fraction of a penny.[43] Reflecting this, Figure 29 in Appendix B shows that detected wash trading is considerably more prevalent among buy/sell and sell/sell trades than buy/buy trades, as collateral is only recycled (to be used in future wash trades) when shares are sold. The right panel of Figure 9 also shows that a high ratio of volume to open interest is strongly associated with wash trading. This reflects the relationship between the buy/sell frequency and wash trading, since buy/sell trades generate volume but have no impact on open interest (the total number of outstanding shares in active markets).

We further find that our algorithm identifies wallets which close their positions quickly without substantial changes in price, even though timing and price do not

[42] `https://polymarket.com/event/romania-presidential-election`, accessed September 8, 2025.

[43] The counterparties of the initial "buy" trade may be non-colluding traders who prefer selling their positions to holding them to market resolution, or who buy the complementary outcome at a high absolute price with the hope of accumulating a small return over a short time horizon.

enter as inputs to the algorithm. Figure 32 in Appendix B shows the distribution of the median duration and price change between the opening and subsequent closing of positions, for wallets with high scores ($x_i \geq 0.9$) versus those with comparatively low scores ($x_i < 0.8$). More than 83% of high-score wallets close positions in at least half of the markets they trade in, while only 13% of low-score wallets do. Conditional on closing a position at least once, 44% of high-score wallets have a median open-to-close time less than one minute, while only 6% of low-score wallets do. Similarly, more than 96% of high-score wallets have a median open-to-close price change less than \$0.01, while only 35% of low-score wallets do.

A challenge in comprehensively evaluating the performance of our detection algorithm is the lack of preexisting ground-truth labels which identify wash trades. However, we take several steps toward partial evaluation by comparing our results with those of alternative, but more limited, methods for detection. In the remainder of this section, we compare wash trades detected by Algorithm 2 and by our specialized methods; demonstrate the use of direct USDC transfers on Polygon to show common ownership between multiple wallets, with the implication that any trades between these wallets are wash trades (our algorithm flags the vast majority of these trades as such); and present identified clusters of wallets engaged in wash trading, for which we can calculate the recall of our algorithm. (We discuss several examples which also give insight into the scale and nature of individual wash-trading strategies.) In Section 6.1 we briefly compare our method with the volume-matching detection algorithm of Victor and Weintraud (2021).

Comparison with Specialized Wash Detection Figure 10 shows the fraction of weekly share volume classified as wash trading by each of the specialized detection methods of Section 5. In general, the total amount of detected wash volume is smaller than that detected by Algorithm 2. After August 2024, there is a rise in the detected wash volume, similar to the pattern observed in Figure 7. Dyadic trades appear to account for the majority of this specialized wash volume prior to January 2025, at which point cluster trades become the dominant form, until detected wash trading mostly subsides in late April 2025, until again increasing in late September 2025. Triangular trade volumes are tiny throughout.

Table 9 shows the fraction of trades (weighted by share volume) detected by each of our specialized methods which are also flagged as wash trades by Algorithm 2. Algorithm 2 detects 64% of dyadic volume (with 180-second time limit), 12% of triangular volume (with 180-second time limit), 85% of chain volume, and 95% of cluster volume, and 16% of the remaining unclassified volume. The reason that a

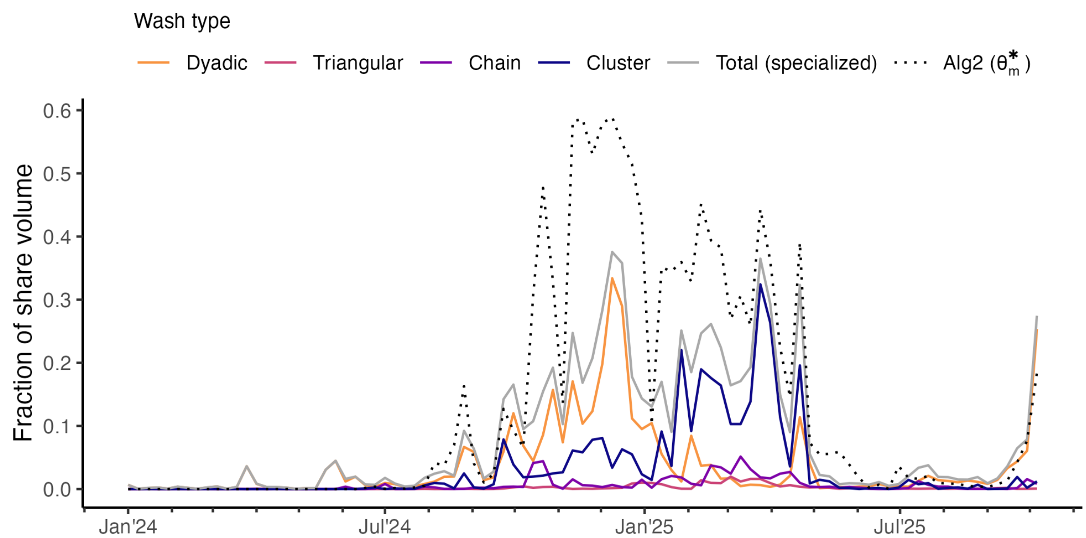


Figure 10: The weekly fraction of Polymarket share volume classified as wash by each of the specialized detection algorithms described in Section 5 (solid), versus wash volume detected by Algorithm 2 (dotted).

comparably small fraction of triangular volume is detected is subtle: Most of the triangular wash trades are perpetrated by a cluster of 890 wallets which we identify as "TwoTrade" in Table 10; as seen in Figure 11, these are organized into a number of "hub-and-spoke" like sub-clusters, in which a central wallet is party to all triangular trades, while the non-central wallets each participate in only one such trade (see Example 3). In a number of instances, the triangle is *not closed* (and the open positions held to resolution, such that some wallets realize large gains or losses), hence the initial scores $x_i^{(0)}$ of some of the peripheral wallets are zero. This is common enough that many of the final scores of the "TwoTrade" wallets fall between 0.8 and 0.9, with their wash trades undetected at $\theta = 0.9$ but detected at $\theta = 0.8$ (using the fixed thresholds of Algorithm 1). Perhaps not incidentally, the fact that these wallets traded in many markets means that the cluster's overall profit-or-loss is limited by

		Pct. of volume detected			
Specialized wash type	Volume (shares)	$\theta = 0.8$	$\theta = 0.9$	$\theta = 0.99$	Alg2 (θ^*_m)
Dyadic	2 242M	90.3%	73.6%	53.2%	63.8%
Triangular	100M	86.7%	34.7%	1.5%	12.1%
Chain	306M	99.1%	97.5%	55.6%	85.3%
Cluster	1 373M	99.2%	98.8%	85.8%	95.1%
Unclassified	24 949M	31.7%	20.5%	10.7%	16.1%

Table 9: The fraction of wash volume identified (at the trade level) by the specialized detection algorithms that is also identified by Algorithm 2 (with market-specific thresholds θ^*_m) and, for reference, by Algorithm 1 under different fixed thresholds θ.

diversification of risk; collectively, these "TwoTrade" wallets realized a $1,469.38 gain on 81.9M of share volume ($78.6M of dollar volume) before ceasing their trading activity in May 2025.

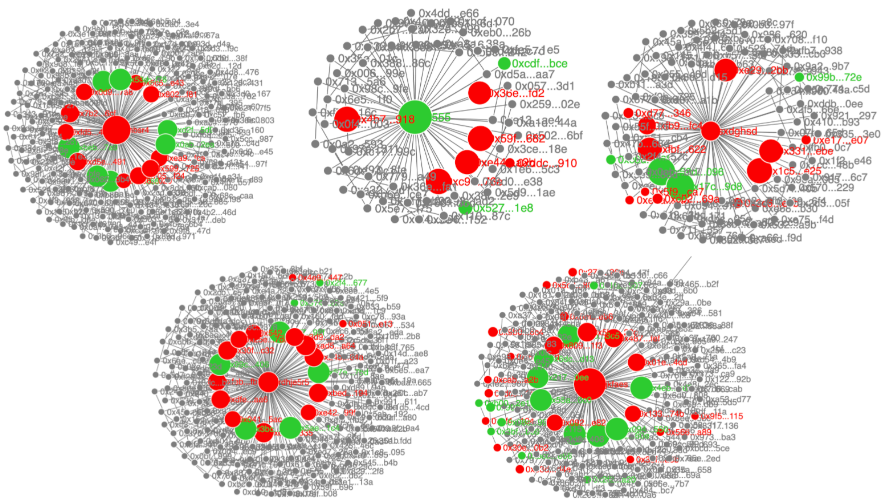


Figure 11: Wallets which execute triangular wash trades as illustrated in Example 3. In some instances, wallets are left exposed when an order is intercepted by an outside trader or when the triangle is otherwise not closed, leading to a gain or loss. Each wallet i is sized proportional to $\sqrt{|profit_i|}$ and is colored green (red) if it has a gain (loss) in excess of $10. Each peripheral wallet (belonging to a completed triangle) trades exactly twice in a single market, once as a buyer and once as a seller. Collectively, the wallets shown realized a $1,469.38 gain on 81.9M of share volume ($78.6M of dollar volume); see the "TwoTrade" cluster in Table 10 for more details.

Further Evidence of Common Ownership In a wash trading scheme, wallets which trade for a limited time eventually transfer their capital to another wallet (which may not necessarily trade on Polymarket). For example, wallet B may buy shares from wallet A, in which case wallet A receives USDC (from a Polymarket module acting as an intermediary to the trade); wallet A may then directly transfer the USDC to wallet C, so that wallet C may then buy the shares held by wallet B without committing new capital. As we mentioned in Section 3, we are sometimes able to observe these direct transfers between wallet addresses on Polygon, which provides strong evidence of common ownership, thereby allowing us to flag trades between these wallets as true wash trades. We start by observing sets of wallets with similar volumes and first-trade dates. We then use the Polygonscan API to retrieve their transfers to external wallet addresses (i.e., excluding the Polymarket modules listed in Table 11), from which we construct networks of transfers.[44] We discuss several examples.

[44] There are often transfers to a Uniswap protocol address, e.g., `0xd36ec33c8bed5a9f7b6630855f1533455b98a418`, in which case the counterparty is observ-

In one instance, wash trades are executed between wallets in 81 parallel chains between October 12, 2024 and November 4, 2024. In each chain, a wallet trades for a short time (usually less than 24 hours) with the simultaneously active wallets in the other chains, before transferring its capital through a series of non-trading wallets to the next trading wallet in the chain. In total, this formation comprises 12,036 wallets, including 1,823 trading wallets which collectively lose $666.00 on 2.9M of share volume ($2.6M of dollar volume); 85.0% of their volume is within-cluster, and 56.0% of this volume is classified as wash trading by Algorithm 2. See the cluster called "81chain" in Table 10.

In another instance, we discover a large network of 1,028 trading wallets which collectively traded 792M of share volume ($407M of dollar volume) almost exclusively in sports markets, starting October 23, 2024 and with a cumulative loss of only $511.31. Algorithm 2 flags virtually all (99.9%) of their within-cluster volume as wash trades. These wallets are responsible for 27% (21%) of the total dollar (share) volume in sports markets between October 23, 2024 and November 25, 2024, including 70% of the dollar volume during the week of October 28, 2024. The graph of direct USDC transfers among them is shown in Figure 34 in Appendix B. Their capitalization can be traced to the wallet with display name "fengchu", which transfers approximately 5,000 USDC to each of six children—named "fdetdddw", "duichong", "DuiChong1", "duic", "miya", and "DuiDui"—between 2024-10-24 23:35:57 UTC and 2024-10-25 03:13:13 UTC.[45] The trade graph for this cluster is shown in Figure 35 in Appendix B; we classify the wallets into the trading clusters called "MAY", "miya", "duic", "duichong", and "zhongxin" which are presented in Table 10.

Wash-Trading Clusters In addition to associating wallets by their direct USDC transfers on Polygon, we identify clusters of wallets which trade among themselves by similarities in their display names and trading statistics, e.g., total volume, capital commitment, and number of markets traded. Starting from a small set of wallets which we suspect are wash traders, we discover other colluding wallets by crawling the trade graph, sequentially adding new wallets which are frequent counterparties to existing wallets in the set. Table 10 presents these clusters, along with statistics such as aggregate volume and profits, the fraction of traded volume that is within-cluster, and the fraction of volume classified as wash trading by our algorithm. We discuss

able from the corresponding transaction. It is also common that transfers are sent through an exchange such as Binance or OKX, in which case the counterparty is not traceable.

[45] See Table 14 in the Appendix. fengchu is itself funded on October 22, 2024 by untraceable transfers via Bybit and OKX; see https://polygonscan.com/tokentxns?a=0xc469af64E33E8e1caBE2CB761aD1C3552F29dd61&p=5, accessed July 14, 2025.

several notable examples.

We identify a cluster of 514 wallets—many of which are actively trading as of our October 12, 2025 data cutoff—which we call "Fantasy" for the evocative display names of the constituent wallets, e.g., "Myrkos", "Uvenlor", and "Qorveth". The cluster includes the wallets named in Table 8—"Mazric", "Lanze", "Felvra" and "Therzia"—which are representative in their attempts to conceal their wash trades by (i) mixing in legitimate trades; (ii) not fully closing their positions; and (iii) sometimes holding shares to market resolution in short-duration markets, leading to a large realized profit or loss. These wallets have so far collectively traded 109M in share volume ($109M in dollar volume), with 99.8% of trade volume within-cluster and 94.4% of this volume classified as wash trading by Algorithm 2. In aggregate, the cluster has so far realized a $1,346.89 gain, while the mean absolute gain or loss of individual wallets is $17,134.

Another cluster is comprised of more than 43,000 wallets with 10-character display names, e.g., "vOWdcRhNQl" and "GyIutxAtzc", which we accordingly name "TenChar" (see Figure 18 in Appendix B). These wallets have collectively traded 188M in share volume at tenths of a penny, such that they account for only $0.9M in dollar volume; 93.4% of their share volume is traded within-cluster, and 90.9% of that volume is classified as wash trading by Algorithm 2. We find that the cluster has so far realized a $9,872.93 aggregate loss; the mean absolute gain or loss of individual wallets is $0.75, and 99.8% of wallets have an absolute gain or loss less than $4.

Cluster label	# wallets	Agg. profit	Volume traded		Pct. vol. in cluster	In-cluster wash pct.		First trade	Last trade	Markets traded	Events traded
			Shares	Dollars		Alg2	VW21				
zhongxin*	492	-165.7	517.1M	193.4M	99.9%	99.8%	97.0%	2024-10-23	2025-09-10	638	106
duichong*	105	-38.08	30.0M	18.8M	99.9%	100.0%	87.0%	2024-10-25	2024-12-25	36	3
MAY*	200	-57.86	116.5M	113.1M	99.8%	100.0%	89.4%	2024-10-24	2024-11-19	119	11
Fantasy	514	1346.89	109.2M	109.1M	99.8%	94.4%	5.7%	2025-03-29	2025-09-11	4,715	2,060
TwoTrade	890	1469.38	81.9M	78.6M	99.8%	8.7%	72.2%	2024-12-18	2025-05-14	202	133
miya*	171	-175.57	94.8M	49.3M	99.7%	100.0%	96.5%	2024-10-25	2024-12-26	147	13
duic*	59	-74.1	33.7M	32.6M	99.6%	100.0%	90.2%	2024-10-25	2024-12-27	138	9
Carti	205	-7618.43	139.0M	138.7M	99.4%	86.1%	18.6%	2025-01-11	2025-09-11	4,654	2,390
Name	40	-155.04	12.2M	12.1M	99.0%	100.0%	0.4%	2024-12-14	2024-12-29	21	17
New4	300	-68.25	4.1M	4.1M	99.0%	99.3%	9.3%	2024-12-27	2025-03-17	31	25
fourchar	100	-146.23	5.1M	5.1M	97.0%	97.0%	6.3%	2024-11-14	2024-12-04	17	10
Anon	5,238	-1967.31	55.7M	55.0M	96.1%	96.3%	19.2%	2024-12-15	2024-12-31	488	195
monasa	1,816	-2418.86	21.1M	0.1M	95.1%	0.0%	30.0%	2024-12-29	2025-05-06	17	16
Anon2	200	-559.65	10.2M	10.1M	95.0%	100.0%	6.6%	2024-12-02	2024-12-13	49	36
Lander	109,142	-64160.26	932.7M	79.9M	94.1%	93.3%	0.0%	2024-10-13	2025-09-11	10,714	5,709
TenChar	43,011	-9872.93	188.0M	1.0M	93.4%	90.9%	0.1%	2025-01-13	2025-09-11	1,824	622
Anon6	690	-502.86	16.6M	16.5M	93.3%	96.9%	16.1%	2024-12-04	2025-04-15	241	85
81chain	1,823	-666.38	2.9M	2.6M	85.0%	56.0%	18.8%	2024-10-12	2024-11-04	27	15

Table 10: Summary table for a selected set of identified wash-trading clusters. The columns include the aggregate profit and volume traded (i.e., the volume of trades in which either counterparty is a member of the cluster); the fraction of share volume traded within-cluster, i.e., with another member of the cluster as counterparty; the fraction of the within-cluster volume classified as wash trading by Algorithm 2 and by the volume-matching algorithm of Victor and Weintraud (2021) (see Section 6.1); the dates of the first and last trades by any wallet in the cluster (note that the data has a October 12, 2025 cutoff); and the number of unique markets and events traded. Clusters marked with an asterisk (*) are part of the large graph shown in Figure 35 in Appendix B. The clusters were discovered using numerous techniques, including crawling the network of direct USDC transfers on Polygon; sequentially adding wallets which trade almost exclusively with existing wallets assigned to the cluster; and observing common display names and statistical similarities in trading activity.

6.1 Comparison with Alternative Detection Methods

We now compare our results with those obtained using the volume-matching algorithm of Victor and Weintraud (2021).[46] The authors write:

> Our aim is... to identify sets of trades between collusive trading accounts that lead to no change in the individual position of each participating trader. In other words, for each account within a set of trading accounts, the total amount of purchased assets equals the total amount of sold assets, such that the involved traders essentially hold the same position they had initially.

The authors' analysis is based on the trade graphs of two Ethereum-based token exchanges. Their algorithm has two steps:

1. First, it detects strongly connected components (SCCs) in the trade graph, along with their multiplicities, or *occurrences*, from an iterative counting procedure.[47] It then filters for SCCs which occur unusually often, for example those above the 99th percentile in the distribution of occurrences.

2. The second step is a volume-matching procedure: for each SCC, restricting trades to a fixed time window, the algorithm searches for subsets of trades for which the traders have almost no change in their net positions, up to an allowable deviation of 1% of the average trade size.

There are several features of the Polymarket data which make VW21's algorithm (unmodified) unsuitable for detection. First, many wash-trading wallets which trade together in a cycle do so only once—either in a given market, or across all markets—such that their SCCs have a low occurrence and will not be flagged as suspicious. Second, neither step is robust to cases where (i) a sequence of wash trades begins or ends with the market, i.e., with traders outside the cluster, as in Example 4 and Example 6; or (ii) wash orders are intercepted by non-collusive wallets.

We apply VW21's algorithm to each market independently. To alleviate the problem of low occurrences, we ignore the step that filters occurrences and instead run the volume-matching step on all SCCs which occur at least once (which could potentially lead to false positives in detection). We use one hour for the time window in step two, which means that the algorithm may fail to detect wash trades when the wash

[46] `https://github.com/friedhelmvictor/lob-dex-wash-trading-paper`, accessed July 14, 2025.

[47] Starting from a directed trade graph with edge weights w_{ij} designating the number of times that i sold a token to j, each counting iteration detects SCCs—incrementing a counter for SCCs already seen—then decrements all positive edge weights in the graph by 1, eliminating edges with zero weight. The iterations continue until no edges remain.

Figure 12: The weekly fraction of Polymarket share volume classified as wash by Victor and Weintraud (2021)'s volume-matching algorithm, using a 1-hour window size and allowable deviation 1% of the average trade size (solid), versus wash volume detected by Algorithm 2 (dotted).

sequence crosses an hourly break before forming a closed cycle. Figure 12 shows the weekly fraction of volume classified as wash trading by VW21's algorithm, and by our Algorithm 2. Overall, we detect 80.5% of the volume labeled as wash by VW21, while VW21 detects only 26.5% of the volume flagged by Algorithm 2.

As an additional point of comparison, we include a column in Table 10 showing, for our self-identified set of wallet clusters, the fraction of the within-cluster volume flagged as wash trading by VW21. Our method consistently flags a high fraction of trading within these clusters as wash (notable exceptions being the "TwoTrade" and "monasa" clusters), while VW21's volume-matching method often does not.

7 Discussion

In this work, we develop an iterative network-based approach for the unsupervised detection of wash trades. The algorithm analyzes the trade graph of all historical transactions, and identifies traders who almost always close their open positions, and trade primarily with other traders exhibiting similar behavior. This approach is flexible, avoiding unnecessary restrictions on specific trading patterns—for example, closed cycles or rapid back-and-forth trades—which may constitute wash trading. In the case of Polymarket, we find convincing evidence that there exist patterns of wash trading which contribute significant volume but are not detected by existing methods in the literature.

Our algorithm has a modular structure, with components which may be independently modified or replaced, meaning that our approach is more general than our particular implementation of it. For example, our choice of the initial vector of scores $\boldsymbol{x}^{(0)}$ (in Part I of Algorithm 1) captures behavior that is plausibly associated with wash trading, namely a strong tendency to close positions; other initializations based on alternative behaviors, or combination of behaviors, are possible. For example, one might consider an initialization based on "dollar days per share volume", i.e., the average capital invested and its duration, relative to the total share volume generated. A low value for this metric reflects a high turnover rate without requiring that positions be fully closed. Another alternative could be the share-weighted average difference in wallet creation times between a wallet and its counterparties—wallets belonging to large trading clusters are often created in quick succession (see Figure 26 in Appendix B).

While our empirical implementation has been limited to Polymarket, the approach is more general and a promising next step would be to apply the algorithms to transaction-level data from other financial markets. The potential for large-scale wash trading means that volume may be unreliable as a metric of authentic platform activity, especially in cryptocurrency-based exchanges which may not have proper safeguards such as Know-Your-Client (KYC) verification, and which do not charge transaction fees. This is especially true when volume generation is encouraged through financial incentives.[48] Until such time as the authenticity of trades can be quickly and reliably established, it may be better to rely on less manipulable measures of platform activity such as open interest, which cannot be inflated without limit by recycling capital across multiple trades.

Finally, we note that the use of our detection algorithm by an exchange would incentivize those who wish to engage in wash trading to adapt in ways that circumvent it. For example, wallets could maintain a buffer position while trading and hold these shares to market resolution to avoid closing positions (as seen in Example 5). This is a familiar phenomenon in financial markets—for instance the detection and exploitation of asset pricing anomalies leads to their disappearance over time. The modularity of our approach, however, may aid in this challenge; some of the alternative score initializations discussed earlier, for example, may be more robust to such

[48]Kalshi, a fiat-based prediction market exchange, launched a volume incentive program in September 2025 (the incentive is limited to trades at prices between \$0.03 and \$0.97; moreover, Kalshi does KYC verification). Apart from a possible token airdrop, Polymarket launched a small volume-based incentive for the creators of parlays, also in September 2025. See https://help.kalshi.com/incentive-programs/volume-incentive-program and https://discord.com/channels/710897173927297116/775506448041115669/1412065932095787058, accessed November 3, 2025.

strategic manipulation. The general question of designing an approach to detection that survives adaptation as part of a game theoretic equilibrium is beyond the scope of this paper but remains an interesting direction for future research.

References

A. Aloosh and J. Li. Direct evidence of Bitcoin wash trading. Management Science, 70(12):8875–8921, 2024. URL https://doi.org/10.1287/mnsc.2021.01448.

Bloomberg. Polymarket Set for US Return After Deal to Buy Tiny Exchange. https://www.bloomberg.com/news/articles/2025-07-21/crypto-betting-site-polymarket-set-for-us-return-after-deal-to-buy-qcx, 2025. Accessed: 2025-09-11.

Y. Cao, Y. Li, S. Coleman, A. Belatreche, and T. M. McGinnity. Detecting wash trade in financial market using digraphs and dynamic programming. IEEE Transactions on Neural Networks and Learning Systems, 27(11):2351–2363, 2016. URL https://doi.org/10.1109/TNNLS.2015.2480959.

CFTC. Wash Trading. https://www.cftc.gov/LearnAndProtect/AdvisoriesAndArticles/CFTCGlossary/index.htm#washtrading, 2025. Accessed: 2025-09-11.

H. Chen, X. Duan, A. E. Saddik, and W. Cai. Political leanings in Web3 betting: Decoding the interplay of political and profitable motives, 2024. URL https://arxiv.org/abs/2407.14844.

CME Group. Wash Trades. https://www.cmegroup.com/education/courses/market-regulation/wash-trades/definition-of-a-wash-trade.html, 2025. Accessed: 2025-09-11.

CoinDesk. Polymarket Will Launch Token and Airdrop After U.S. Relaunch, CMO Says. https://www.coindesk.com/markets/2025/10/24/polymarket-will-launch-token-and-airdrop-after-u-s-relaunch-cmo-says, 2025. Accessed: 2025-11-06.

L. W. Cong, X. Li, K. Tang, and Y. Yang. Crypto wash trading. Management Science, 69(11):6427–6454, 2023. URL https://doi.org/10.1287/mnsc.2021.02709.

M. H. DeGroot. Reaching a consensus. Journal of the American Statistical Association, 69(345):118–121, 1974. URL http://www.jstor.org/stable/2285509.

D. L. Donoho, A. Maleki, and A. Montanari. Message passing algorithms for compressed sensing: I. Motivation and construction. In 2010 IEEE Information Theory

Workshop on Information Theory (ITW 2010, Cairo), pages 1–5. IEEE, 2010. URL https://doi.org/10.1109/ITWKSPS.2010.5503193.

B. Eichengreen, G. Viswanath-Natraj, J. Wang, and Z. Wang. Under pressure? Central bank independence meets blockchain prediction markets, July 2025. URL https://ssrn.com/abstract=5366862.

N. E. Friedkin and E. C. Johnsen. Social influence networks and opinion change. Advances in Group Processes, 16:1–29, 1999. URL https://www2.cs.siu.edu/~hexmoor/classes/CS539-F10/Friedkin.pdf.

R. S. Gladwin. Will Polymarket launch a token? Why some traders are betting on an airdrop. Decrypt, 2024. URL https://decrypt.co/248178/will-polymarket-launch-token-ethereum-airdrop.

L. R. Glosten and P. R. Milgrom. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. Journal of Financial Economics, 14(1):71–100, 1985. ISSN 0304-405X. URL https://doi.org/10.1016/0304-405X(85)90044-3.

D. Karger, S. Oh, and D. Shah. Iterative learning for reliable crowdsourcing systems. Advances in Neural Information Processing Systems, 24, 2011. URL https://proceedings.neurips.cc/paper_files/paper/2011/file/c667d53acd899a97a85de0c201ba99be-Paper.pdf.

F. R. Kschischang, B. J. Frey, and H.-A. Loeliger. Factor graphs and the sum-product algorithm. IEEE Transactions on Information Theory, 47(2):498–519, 2002. URL https://doi.org/10.1109/18.910572.

Legal Information Institute. 7 U.S. Code § 6c - Prohibited transactions. https://www.law.cornell.edu/uscode/text/7/6c, 2025. Accessed: 2025-09-11.

Y. Lu, A. Montanari, B. Prabhakar, S. Dharmapurikar, and A. Kabbani. Counter braids: A novel counter architecture for per-flow measurement. ACM SIGMETRICS Performance Evaluation Review, 36(1):121–132, 2008. URL https://web.stanford.edu/~balaji/papers/08counterbraidsCONF.pdf.

K. Murphy, Y. Weiss, and M. I. Jordan. Loopy belief propagation for approximate inference: An empirical study. arXiv preprint arXiv:1301.6725, 2013. URL https://doi.org/10.48550/arXiv.1301.6725.

S. E. Parsegov, A. V. Proskurnikov, R. Tempo, and N. E. Friedkin. Novel multidimensional models of opinion dynamics in social networks. IEEE Transactions on Automatic Control, 62(5):2270–2285, 2017. URL https://doi.org/10.1109/TAC.2016.2613905.

H. Qiao, H. Tong, B. An, I. King, C. Aggarwal, and G. Pang. Deep graph anomaly detection: A survey and new perspectives. IEEE Transactions on Knowledge and Data Engineering, 37(9):5106–5126, 2025. URL https://doi.org/10.1109/TKDE.2025.3581578.

D. Rothschild and R. Sethi. Trading strategies and market microstructure: Evidence from a prediction market. Journal of Prediction Markets, 10(1), 09 2016. URL https://doi.org/10.5750/jpm.v10i1.1179.

L. Schwartz. Exclusive: Election betting site Polymarket gives Trump a 67% chance of winning but is rife with fake 'wash' trading, researchers say. Fortune, 2024. URL https://fortune.com/crypto/2024/10/30/polymarket-trump-election-crypto-wash-trading-researchers/.

F. Victor and A. M. Weintraud. Detecting and quantifying wash trading on decentralized cryptocurrency exchanges. In Proceedings of the Web Conference 2021, WWW'21, page 23–32, New York, NY, USA, 2021. Association for Computing Machinery. URL https://doi.org/10.1145/3442381.3449824.

V. von Wachter, J. R. Jensen, F. Regner, and O. Ross. NFT wash trading: Quantifying suspicious behaviour in NFT markets, 2022. URL https://arxiv.org/abs/2202.03866.

Appendices

Appendix A Data Collection

Here we provide further details on data collection. Table 11 has the addresses of Polymarket modules which facilitate trades and the distribution of rewards, as well as contract addresses which represent the underlying assets being exchanged (conditional tokens and USDC.e). Table 12 shows the flow of conditional tokens and collateral for each type of user-initiated transaction in CTF and NegRisk markets. The transfer events are obtained using Polygonscan's `accounts/` API endpoint for each module address.

label	address / contract_address
CTF_EXCHANGE	0x4bfb41d5b3570defd03c39a9a4d8de6bd8b8982e
NEG_RISK_CTF_EXCHANGE	0xc5d563a36ae78145c45a50134d48a1215220f80a
NEG_RISK_ADAPTER	0xd91e80cf2e7be2e162c6513ced06f1dd0da35296
NEG_RISK_WRAPPED_COLLATERAL	0x3a3bd7bb9528e159577f7c2e685cc81a765002e2
NEG_RISK_BURN	0xa5ef39c3d3e10d0b270233af41cac69796b12966
LIQUIDITY_REWARDS[49]	0xc288480574783bd7615170660d71753378159c47
HOLDING_REWARDS	0xc536633ff12ee52e280b2af2594031060c5aaf41
USDC.e	0x2791bca1f2de4661ed88a30c99a7a9449aa84174
CONDITIONAL_TOKEN	0x4d97dcd97ec945f40cf65f87097ace5ea0476045
NULL	0x0000000000000000000000000000000000000000

Table 11: Addresses of Polymarket modules and contracts relevant for data collection.

We obtain market-level information from Polymarket's Gamma Markets API.[50] This includes the token_ids which represent the "Yes" and "No" outcomes in each market. Some of the information on token identifiers, resolution prices, and market close timestamps appears to be incorrect, so we retrieve this information directly from ConditionResolution event logs on Polygonscan, accessible via the `logs/` API endpoint.[51] The raw event logs must be parsed (decoded) using the ConditionalToken application binary interface (ABI), and the token_ids obtained by calling a Keccak-

[49]Rewards were distributed directly to users' wallet addresses starting November 24, 2023; see https://discord.com/channels/710897173927297116/775506448041115669/1180273215088627712, accessed July 14, 2025.

[50]https://docs.polymarket.com/developers/gamma-markets-api/overview, accessed July 14, 2025.

[51]*Events* are emitted when certain functions are executed within the *smart contracts* which implement the Conditional Tokens framework. In this case, the signature corresponding to ConditionResolution events is `0xb44d84d3289691f71497564b85d4233648d9dbae8cbdbb4329f301c3a0185894`.

	ERC-20 (Collateral)		ERC-1155 (ConditionalTokens)	
	from	to	from	to
Buy	<wallet>	CTF_EXCHANGE	CTF_EXCHANGE	<wallet>
Sell	CTF_EXCHANGE	<wallet>	<wallet>	CTF_EXCHANGE
Split	<wallet>	CONDITIONAL_TOKEN	NULL	<wallet>
Merge	CONDITIONAL_TOKEN	<wallet>	<wallet>	NULL
Redeem	CONDITIONAL_TOKEN	<wallet>	<wallet>	NULL

	ERC-20 (Collateral)		ERC-1155 (ConditionalTokens)	
	from	to	from	to
Buy	<wallet>	NR_CTF_EXCHANGE	NR_CTF_EXCHANGE	<wallet>
Sell	NR_CTF_EXCHANGE	<wallet>	<wallet>	NR_CTF_EXCHANGE
Split	<wallet>	NR_ADAPTER	NR_ADAPTER	<wallet>
Merge	NR_WRAPPED_COLLATERAL	<wallet>	<wallet>	NR_ADAPTER
Convert	NR_WRAPPED_COLLATERAL	<wallet>	<wallet> NR_ADAPTER	NR_BURN <wallet>
Redeem	NR_WRAPPED_COLLATERAL	<wallet>	<wallet>	NR_ADAPTER

Table 12: The flow of collateral (USDC.e) and Conditional Tokens between wallets and Polymarket modules for each transaction type in CTF markets (top) and NegRisk markets (bottom).

256 hashing function on the decoded outputs. For this latter step, we connect to a Polygon Mainnet node using the Web3 RPC provider Chainstack.[52]

Appendix B Additional Tables and Figures

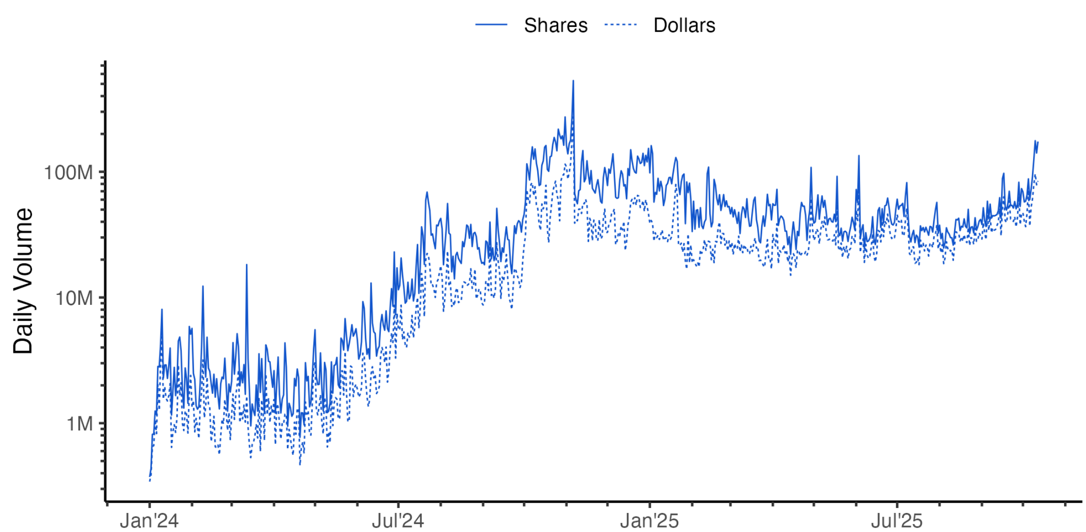


Figure 13: The time series of daily aggregate volume (shares and dollars) for Polymarket (log scale).

[52]https://chainstack.com/. For more detailed instructions on accessing and processing the on-chain data, see https://yzc.me/x01Crypto/decoding-polymarket, accessed September 17, 2025.

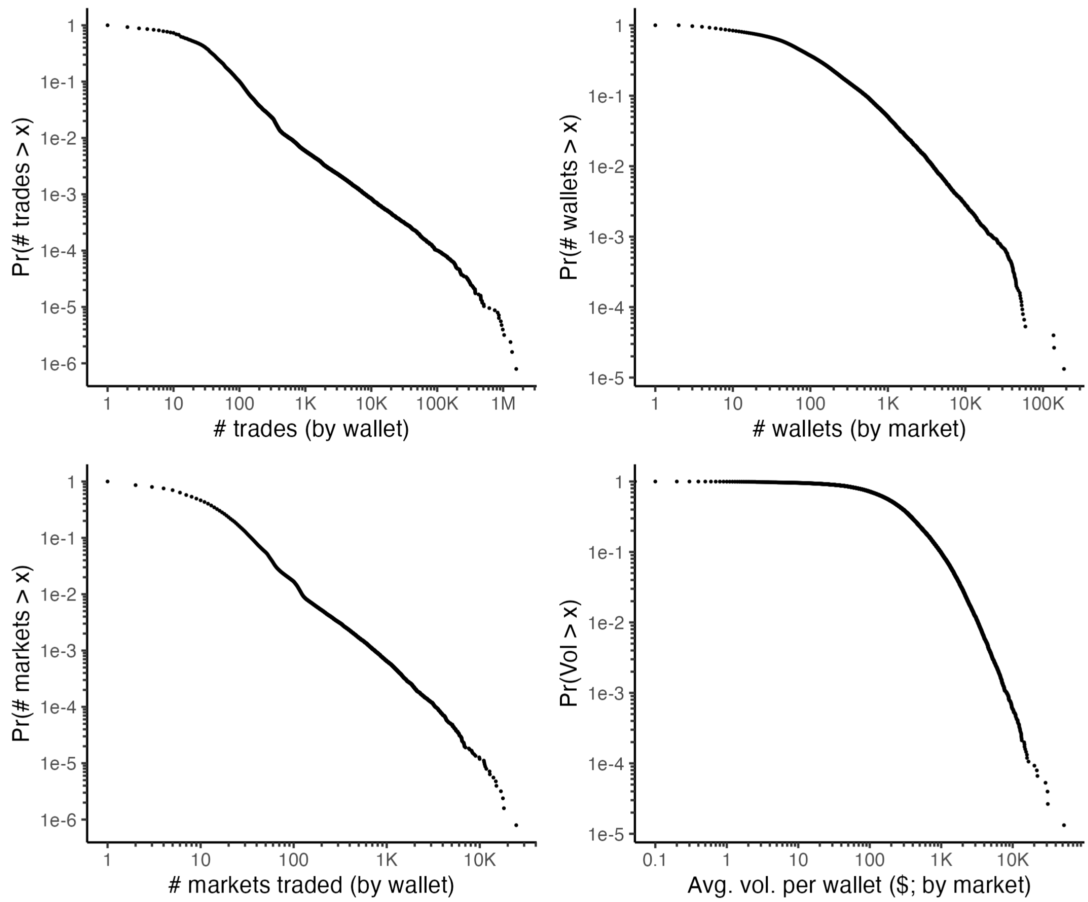


Figure 14: Complementary CDFs for the number of trading wallets per market (top-left); the number of markets traded by wallet (top-right); the number of trades across all markets by wallet (bottom-left); and the average dollar volume traded per wallet-market (bottom-right).

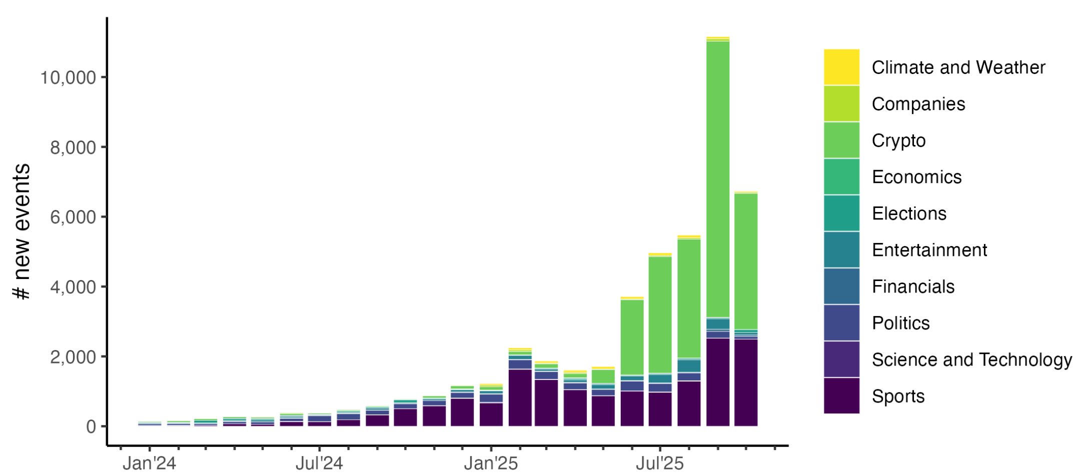


Figure 15: The number of new Polymarket events by start month and category, starting from January 1, 2024, through the October 12, 2025 data cutoff. (Event categories are labeled using OpenAI's ChatGPT 4o-mini due to inconsistent coverage of labels in the Polymarket API.) Note that an event may comprise multiple binary markets.

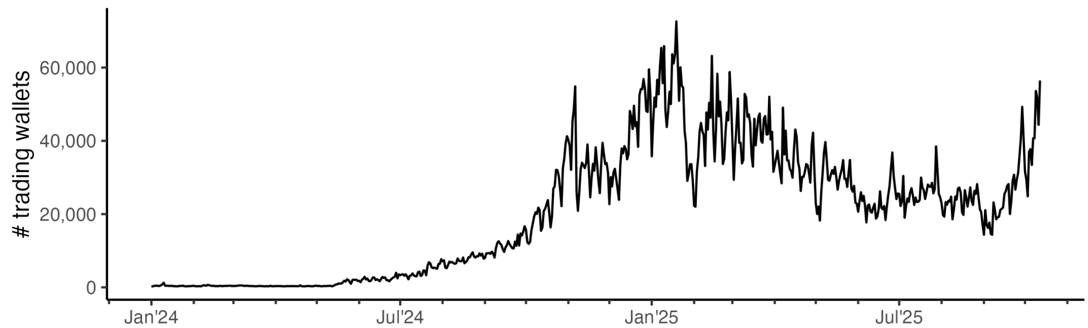


Figure 16: The number of daily active (trading) wallets on the Polymarket CLOB since January 1, 2024.

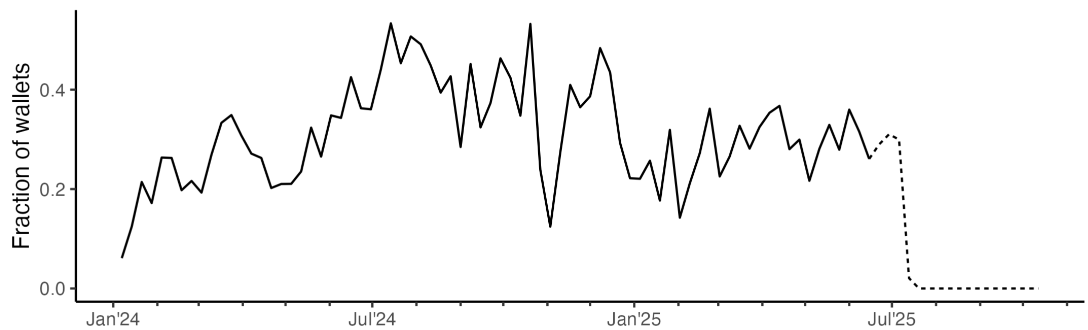


Figure 17: The fraction of wallets by `created_at` week that traded in the 30-day period 90–120 days after their `created_at` date (censored part of the series is dashed).

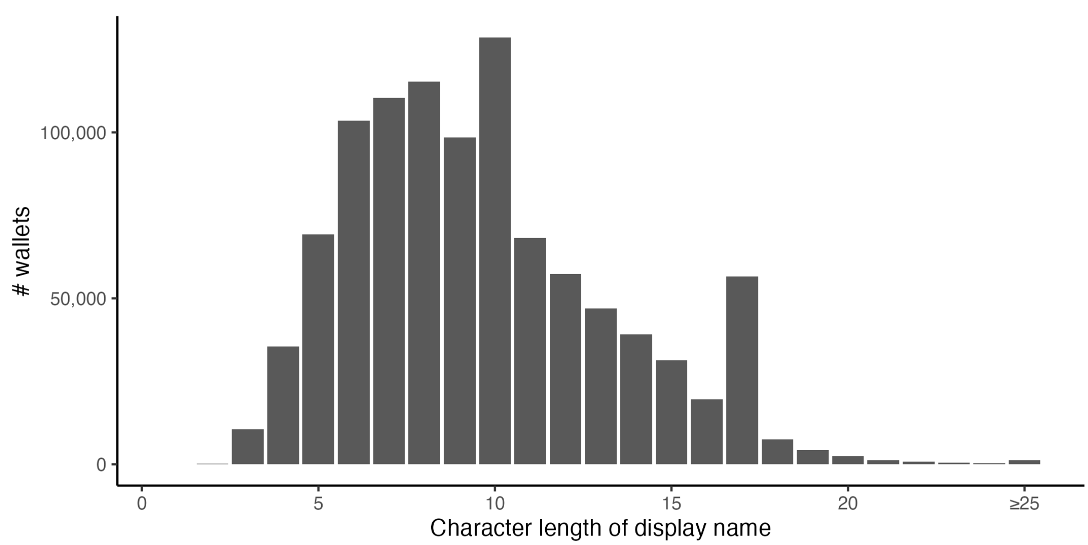


Figure 18: Histogram of the character length of wallet display names. Approximately 250,000 wallets with 42- and 56-character names—which correspond to hexadecimal strings representing Ethereum wallet addresses, or such an address appended with a hyphen and 13 digits (the epoch milliseconds corresponding to the wallet's `created_at` timestamp)—are excluded.

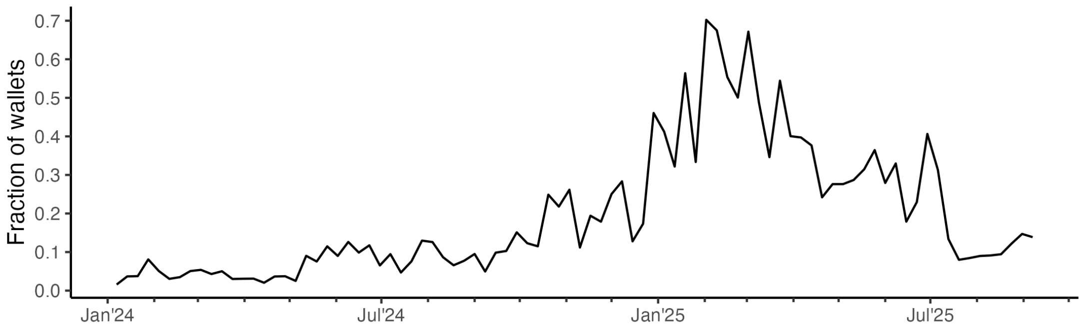


Figure 19: The fraction of wallets by `created_at` week with a cumulative absolute profit-or-loss less than $1 (through the October 12, 2025 data cutoff). A wallet's profit is calculated as its net receipt of USDC plus the current market value of any shares owned.

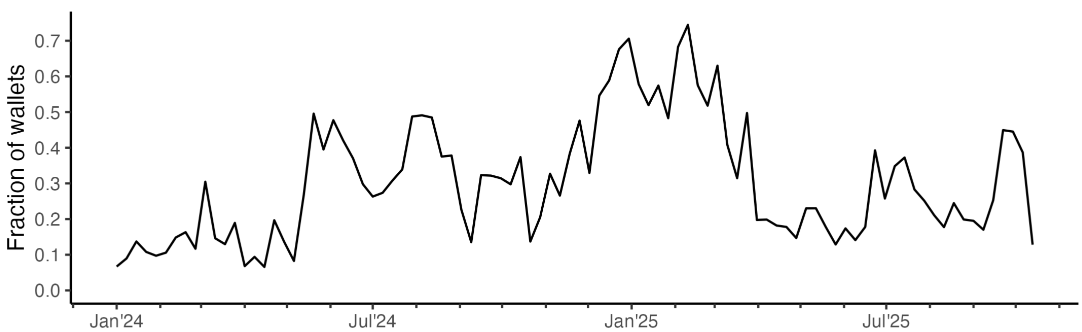


Figure 20: The fraction of wallets by `created_at` week with at least 90% of their cumulative share volume traded at price $p < \$0.01$ or $p > \$0.99$.

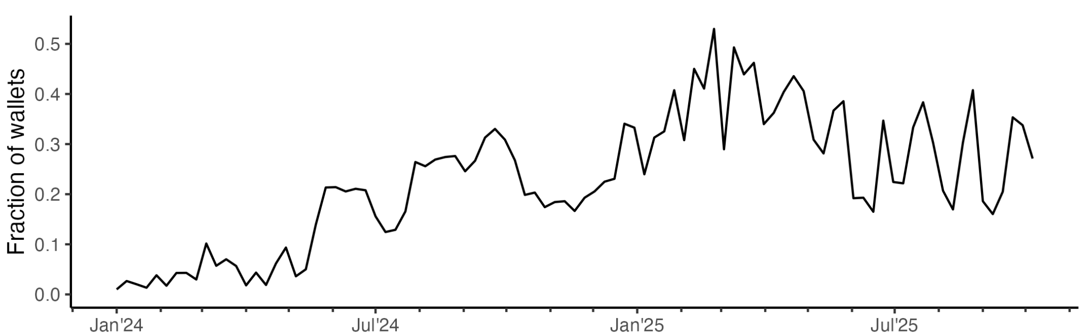


Figure 21: The fraction of weekly active wallets with at least 90% of their weekly share volume traded at price $p < \$0.01$ or $p > \$0.99$.

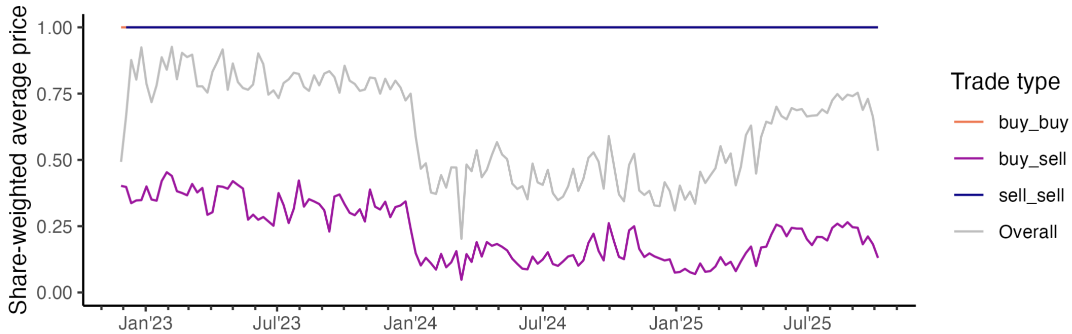


Figure 22: The weekly share-weighted average trade price by trade type. We follow the convention that in buy/buy and sell/sell trades, N shares each of "Yes" and "No" together contribute N units of share volume at average price $1.

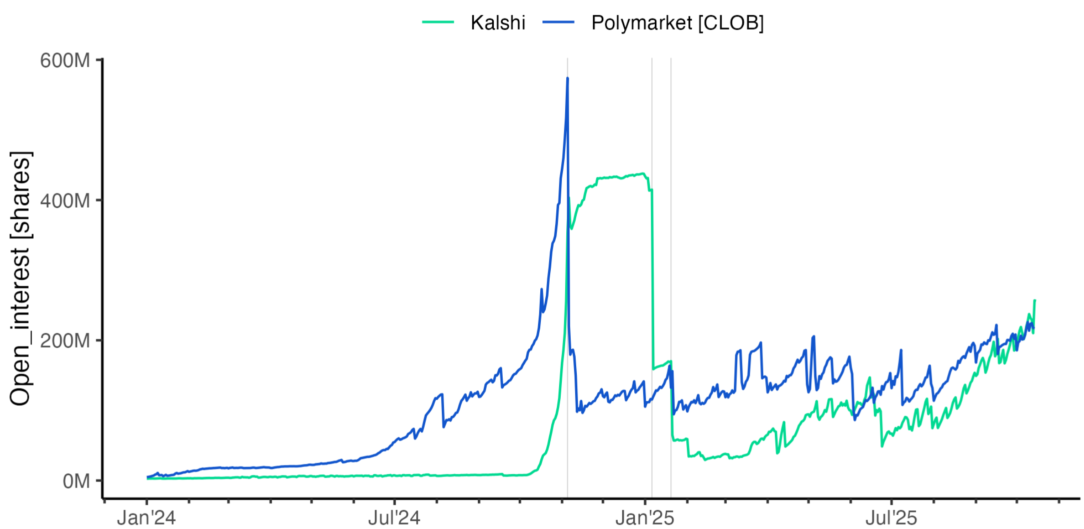


Figure 23: The time series of daily open interest, i.e., the end-of-day total number of outstanding shares (contracts) in active markets, for Polymarket and Kalshi. Polymarket series includes the effect of split and merge transactions (which add to and subtract from open interest, respectively). Note that 2024 U.S. Presidential Election and related markets closed on November 5, 2024 on Polymarket, but not until January 6 or January 20, 2025 on Kalshi. Hence, a large number of contracts were "tied up" on the latter, contributing to the extended period of high open interest. Kalshi data obtained using https://trading-api.readme.io/reference/getmarketcandlesticks-1.

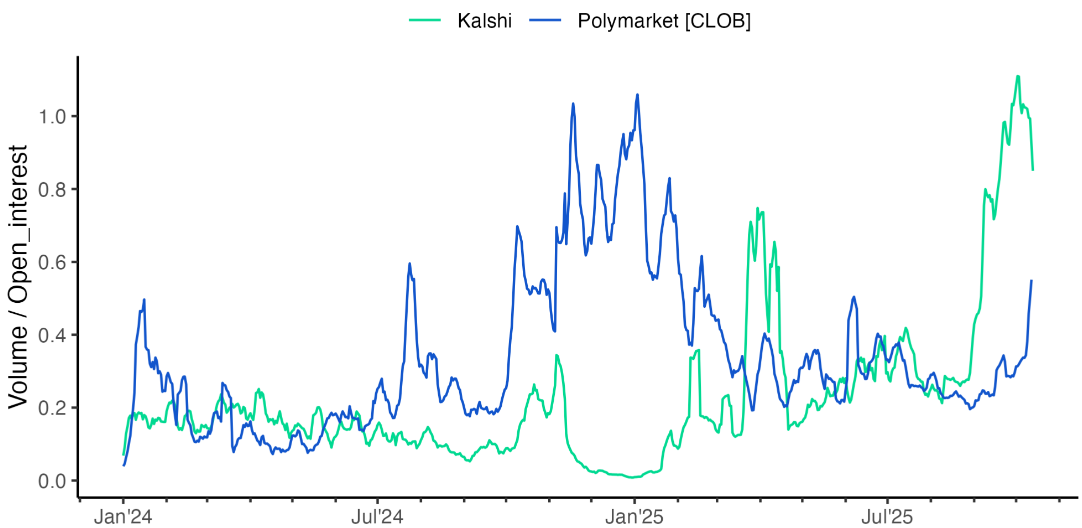


Figure 24: The time series of the daily aggregate volume-to-open interest ratio (7-day moving average) for Polymarket and Kalshi. Polymarket series includes the effect of split and merge transactions on open interest (but these are not counted as volume). Note that 2024 U.S. Presidential Election and related markets closed on November 5, 2024 on Polymarket, but not until January 6 or January 20, 2025 on Kalshi, contributing to the large gap in vol-to-OI during the intervening months. The spikes in Kalshi's vol-to-OI in 2025 are largely attributable to sports betting (e.g., during March Madness). Kalshi data obtained using https://trading-api.readme.io/reference/getmarketcandlesticks-1.

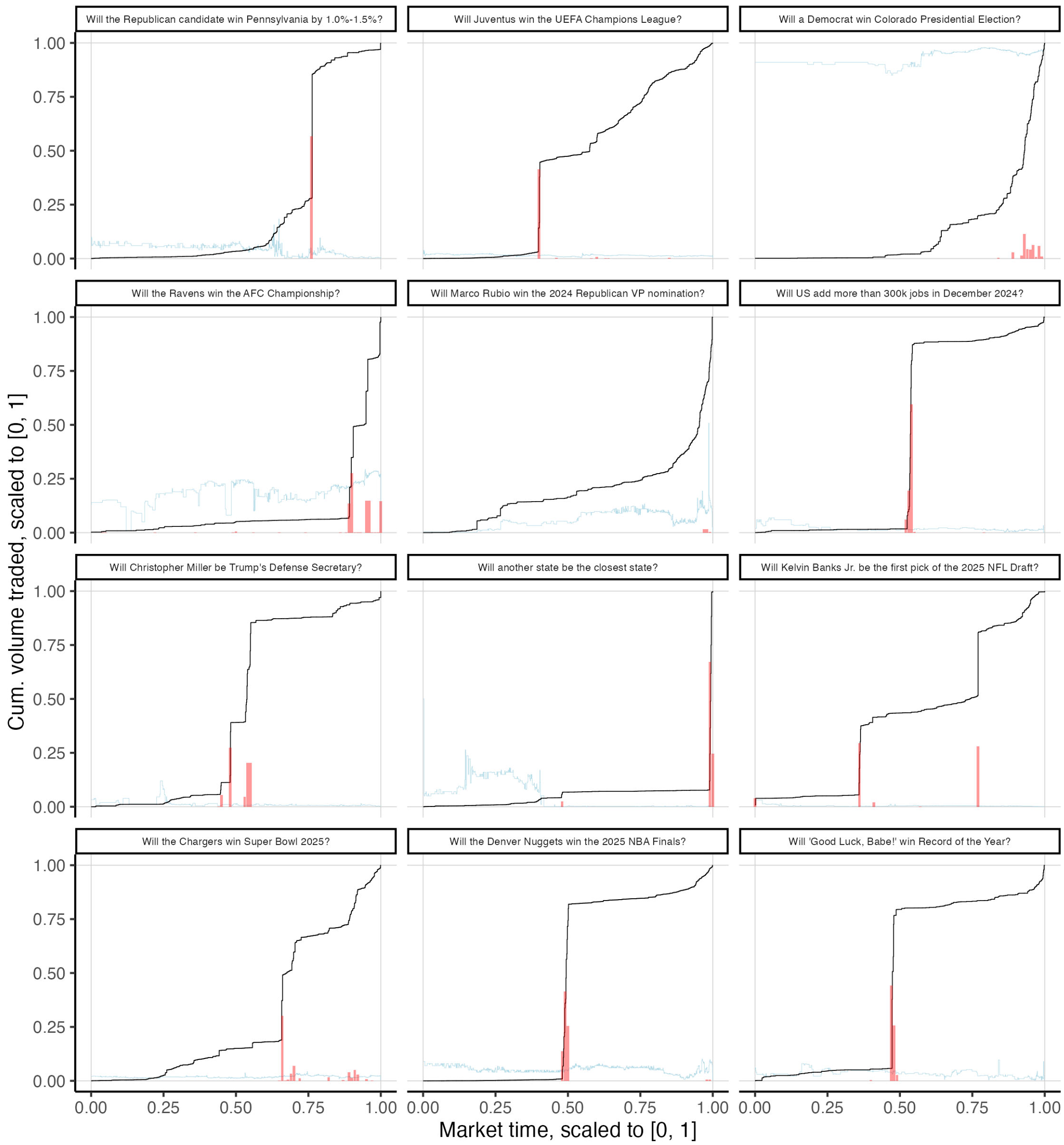


Figure 25: Time series of the cumulative share volume traded (black) and "yes" share price (light blue) for a set of example markets. Red bars—calculated in increments of 0.01 of the market duration—represent the fraction of total market volume classified as wash trading by Algorithm 2. In many cases, wash trades occur rapidly, relative to the duration of the market, and correspond to noticeable spikes in volume without significant price movements.

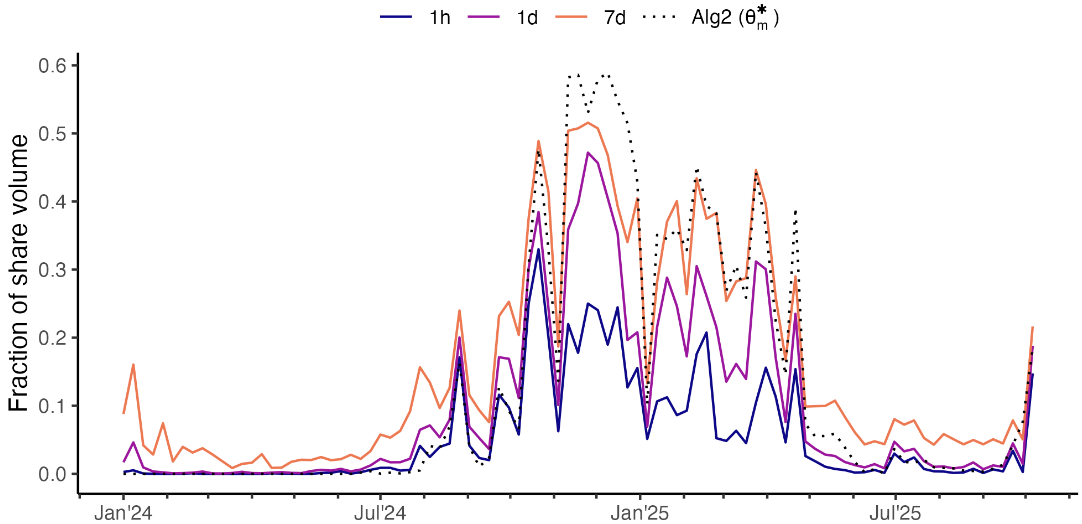


Figure 26: The weekly fraction of Polymarket share volume in which a wallet's `created_at` timestamp is within one hour (1h), one day (1d), or seven days (7d) of its counterparty's `created_at` timestamp, versus wash volume detected by Algorithm 2 (dotted).

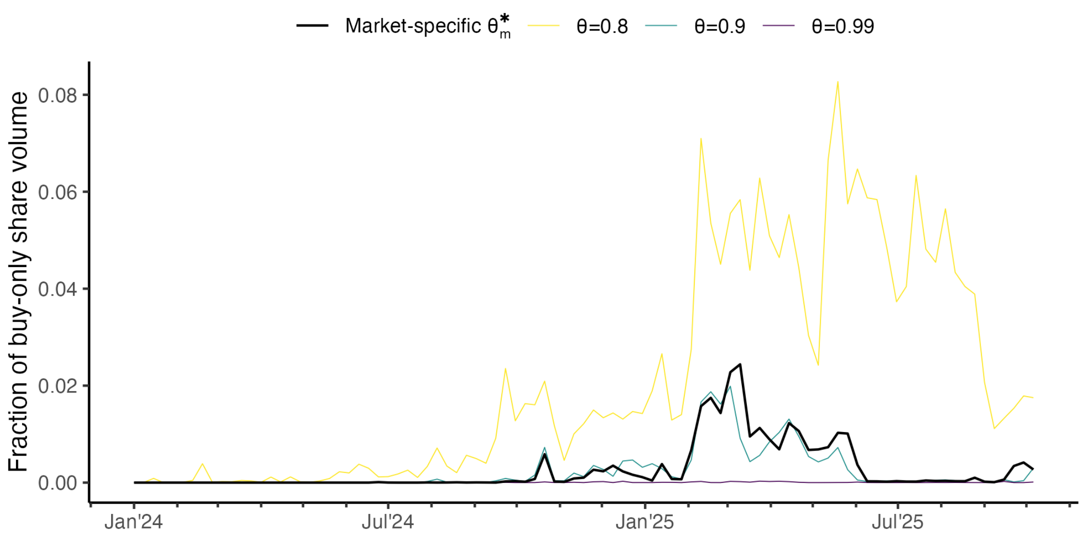


Figure 27: The weekly fraction of *buy-only* share volume that is classified as wash volume by Algorithm 2 (black) and, for reference, by Algorithm 1 under different fixed thresholds θ. (The buy-only volume excludes volume generated by wallets which *only buy*, i.e., never sell, in a given market.)

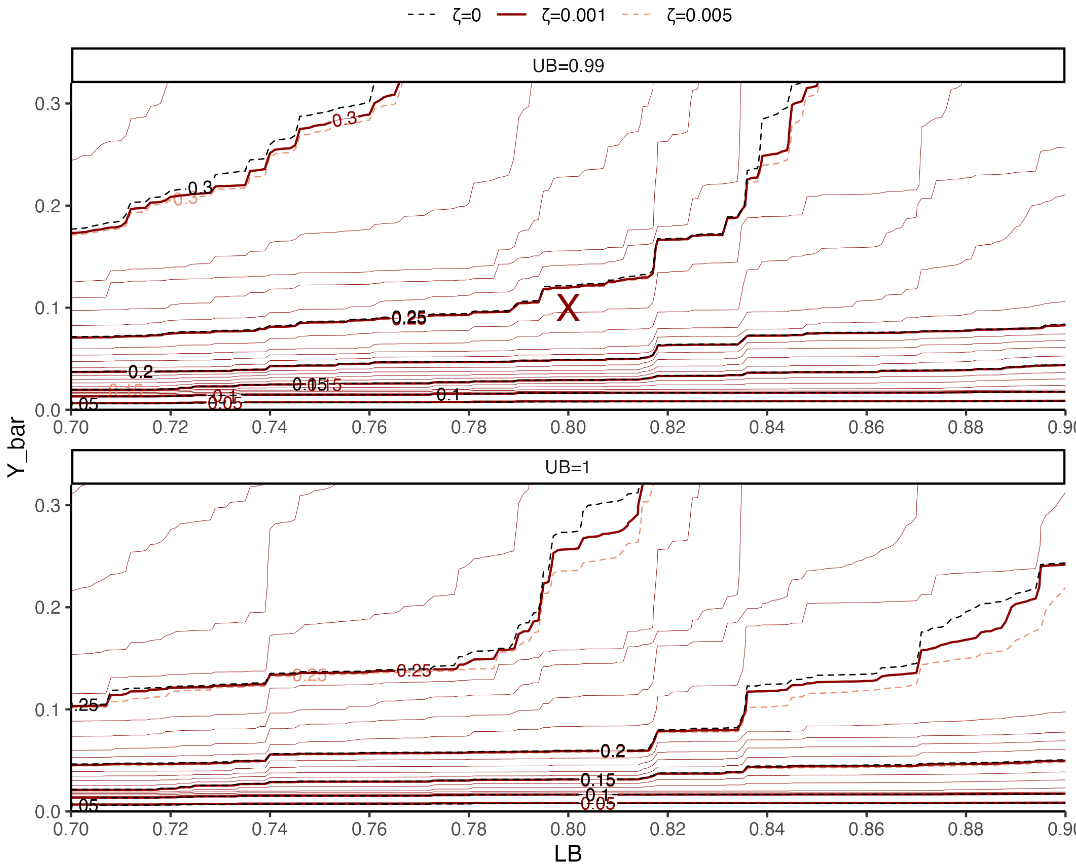


Figure 28: Sensitivity plot showing the aggregate fraction of Polymarket share volume classified as wash trading (contour lines) by Algorithm 2 as a function of threshold parameters $(\underline{\theta}, \overline{\theta}, \overline{Y}, \zeta)$ (note that LB and UB refer to $\underline{\theta}$ and $\overline{\theta}$, respectively). The parameters chosen for our implementation are (0.8, 0.99, 0.1, 0.001), marked by the red X. See "Market-Specific Threshold Selection" in Section 5 for details.

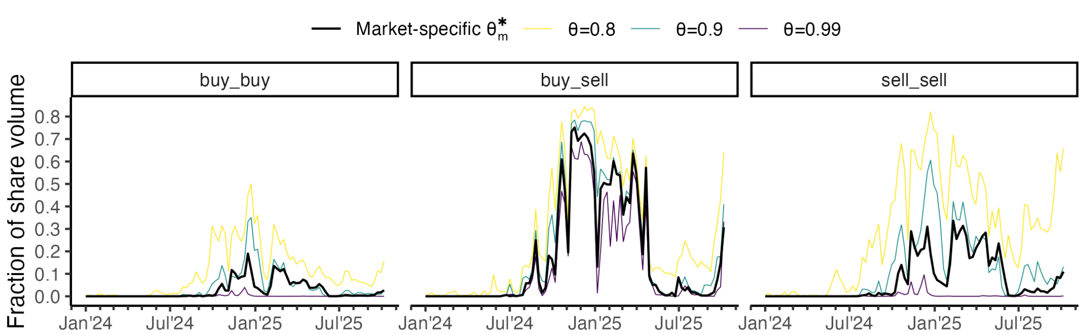


Figure 29: The weekly fraction of Polymarket share volume, by trade type, that is classified as wash volume by Algorithm 2 (black) and, for reference, by Algorithm 1 under different fixed thresholds θ.

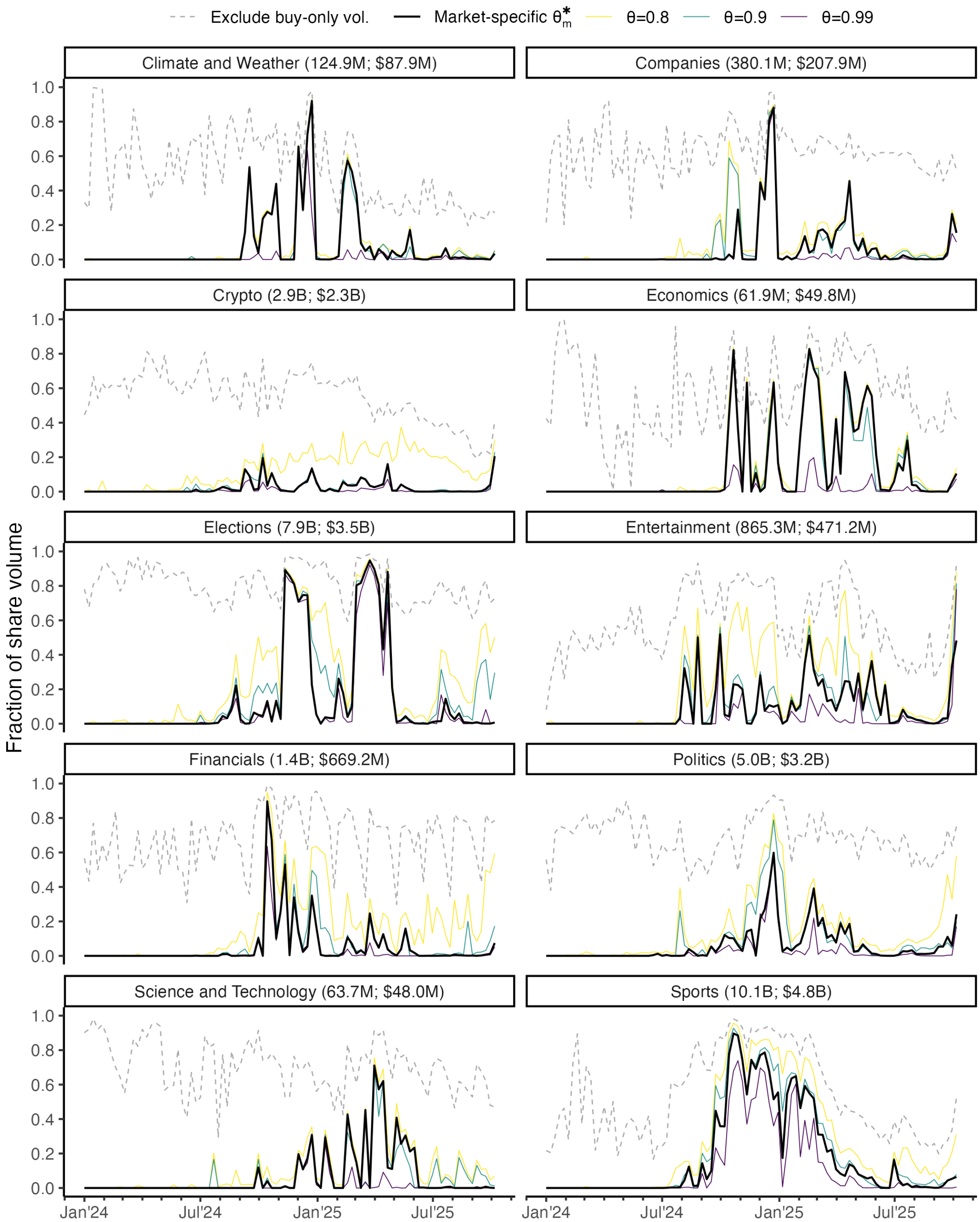


Figure 30: The weekly fraction of Polymarket share volume, by event category, that is classified as wash volume by Algorithm 2 (black) and, for reference, by Algorithm 1 under different fixed thresholds θ. (Event categories are labeled using OpenAI's ChatGPT 4o-mini due to inconsistent coverage of labels in the Polymarket API.) The dashed gray line is the fraction of share volume which excludes volume from wallets which *only buy* shares in a given market. Aggregate share and dollar volume in parentheses.

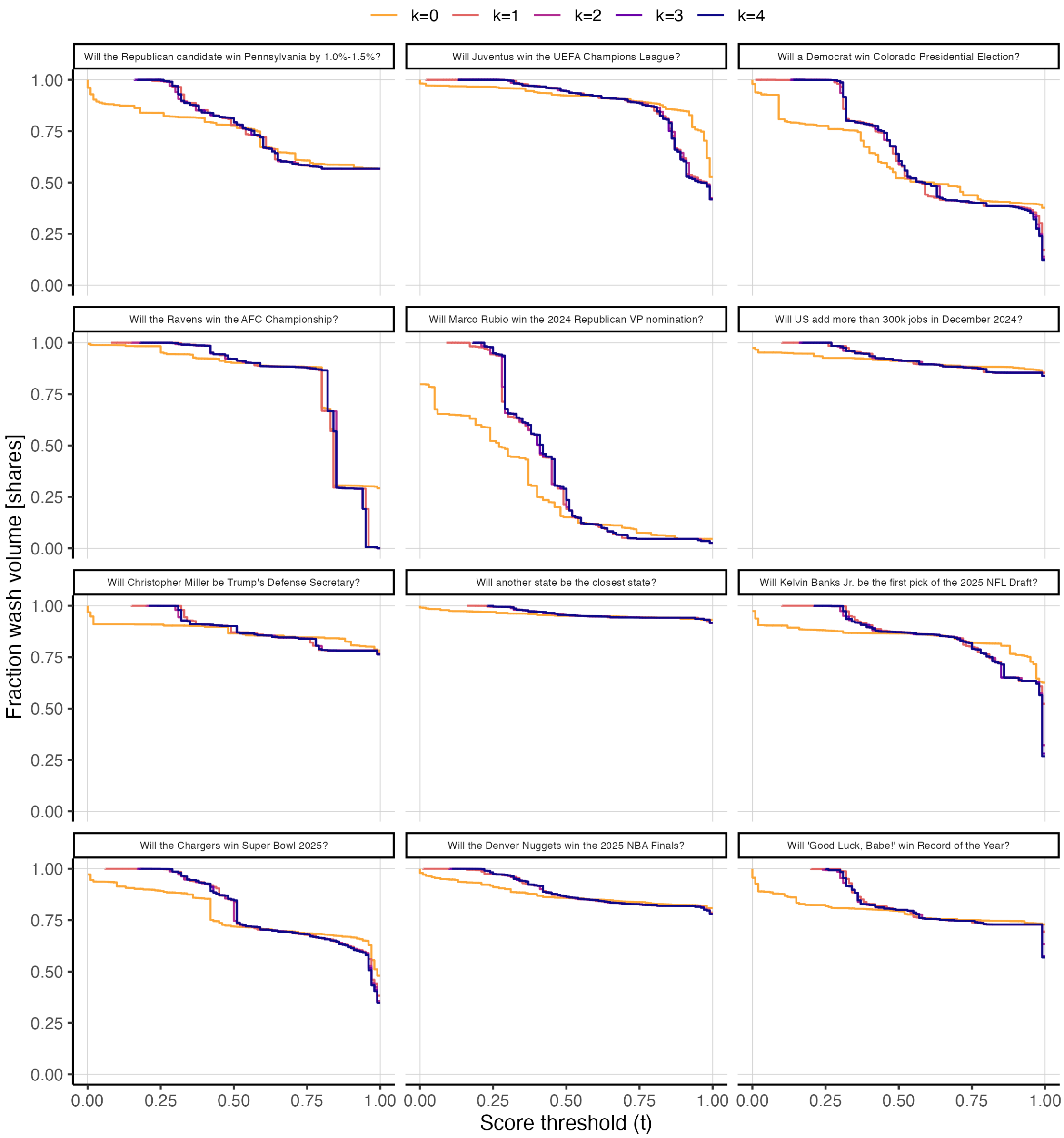


Figure 31: The fraction of share volume classified as wash trading in each example market, by Algorithm 1 iteration (k) and score threshold θ. A trade is labeled a wash trade if both long and short wallets have score $x_i^{(k)} \geq \theta$, hence a higher θ corresponds to lower wash volume.

Question	Market volume [CLOB]		θ_m^*	$\Upsilon_m(\theta_m^*)$	Est. wash %	
	Shares (M)	Dollars ($M)			Alg2	VW21
Will Donald Trump win the 2024 US Presidential Election?	1 568.7	1 184.0	1.000	–	0.0%	0.1%
Will Kamala Harris win the 2024 US Presidential Election?	1 072.0	634.8	1.000	–	0.0%	0.2%
Will Donald Trump be inaugurated?	400.4	324.2	1.000	–	0.0%	0.8%
Will the Sacramento Kings win the 2025 NBA Finals?	378.0	34.6	0.968	0.009	93.0%	4.2%
Will Nicolae Ciucă win the 2024 Romanian Presidential election?	326.5	2.6	0.833	0.004	98.5%	0.0%
Will Zelenskyy wear a suit before July?	242.2	156.9	1.000	–	0.0%	0.0%
Will any other Republican Politician win the 2024 US Presidential Election?	241.7	33.9	0.974	0.091	9.7%	2.0%
Kamala Harris wins the popular vote?	163.8	118.3	1.000	–	0.0%	0.5%
Will the Toronto Raptors win the 2025 NBA Finals?	154.2	15.6	0.800	0.035	96.3%	0.9%
Will Michelle Obama win the 2024 US Presidential Election?	153.4	35.6	1.000	–	0.0%	2.5%
Will Robert F. Kennedy Jr. win the 2024 US Presidential Election?	141.6	36.5	1.000	–	0.0%	2.1%
Will the Panthers win Super Bowl 2025?	139.3	14.3	0.800	0.038	94.8%	22.8%
Fed increases interest rates by 25+ bps after November 2024 meeting?	134.0	8.9	0.936	0.030	87.0%	0.9%
Will Aston Villa win the UEFA Champions League?	133.1	17.7	0.800	0.025	96.8%	2.7%
Will the Washington Wizards win the 2025 NBA Finals?	130.2	16.0	0.800	0.058	93.1%	1.4%
Will the Utah Jazz win the 2025 NBA Finals?	129.9	19.8	0.800	0.052	94.3%	4.4%
Will the Raiders win Super Bowl 2025?	124.0	13.7	0.800	0.023	96.3%	18.4%
Will Donald Trump win the popular vote in the 2024 Presidential Election?	119.9	88.4	0.977	0.091	0.6%	0.2%
TikTok banned in the US before May 2025?	119.7	89.7	0.945	0.018	1.0%	0.4%
Will the Titans win Super Bowl 2025?	119.5	15.6	0.800	0.033	95.0%	3.1%
Will the Charlotte Hornets win the 2025 NBA Finals?	116.7	12.8	0.800	0.051	94.4%	1.4%
Will any other Democratic Politician win the 2024 US Presidential Election?	116.6	28.8	1.000	–	0.0%	3.4%
Will the Browns win Super Bowl 2025?	115.3	5.7	0.800	0.023	96.7%	1.8%
Will the Democratic candidate win Pennsylvania by 1.5%-2.0%?	109.1	0.2	0.817	0.001	99.7%	62.1%
Will the Giants win Super Bowl 2025?	107.9	16.2	0.957	0.043	79.1%	3.0%
Will Nikki Haley win the 2024 US Presidential Election?	107.5	21.8	1.000	–	0.0%	2.3%
Nottingham Forest wins the Premier League?	101.4	15.4	0.800	0.058	92.6%	2.5%
Will Hillary Clinton win the 2024 US Presidential Election?	93.3	9.3	1.000	–	0.0%	1.3%
Will Gavin Newsom be D-nom for VP on Election Day?	92.9	1.0	0.990	0.015	94.0%	92.7%
Southampton wins the Premier League?	88.2	16.2	0.800	0.068	89.7%	3.5%
Will Lee Jae-myung be elected the next president of South Korea?	88.0	73.3	1.000	–	0.0%	0.0%
Will the Miami Heat win the 2025 NBA Finals?	86.0	10.2	0.974	0.027	77.4%	1.1%
Will Inter Milan win the UEFA Champions League?	83.0	77.7	0.975	0.001	91.0%	86.1%
Will Hunor Kelemen win the Romanian presidential election?	80.9	0.3	0.960	0.003	98.7%	0.0%
Will Trump launch a coin before the election?	76.9	51.7	0.966	0.007	0.0%	0.0%
Will the Patriots win Super Bowl 2025?	74.3	10.8	0.800	0.038	93.0%	10.4%
Will the Indiana Pacers win the 2025 NBA Finals?	73.6	28.8	0.972	0.025	44.7%	8.2%
Will Kim Moon-soo be elected the next president of South Korea?	73.2	37.5	1.000	–	0.0%	0.2%
Will Kamala Harris be inaugurated?	72.3	34.5	0.980	0.099	0.3%	1.3%
Will Joe Biden win the 2024 US Presidential Election?	72.2	20.5	1.000	–	0.0%	0.0%
Will Real Betis win La Liga?	71.8	4.5	0.800	0.030	96.9%	0.2%
Will the Charlotte Hornets win the Eastern Conference?	71.3	6.1	0.958	0.008	95.0%	27.0%
Manchester United wins the Premier League?	71.1	20.6	0.974	0.060	48.0%	8.4%
Will Hillary Clinton win the popular vote in the 2024 Presidential Election?	70.2	10.5	0.986	0.024	5.5%	6.8%
Will the LA Clippers win the 2025 NBA Finals?	69.8	48.8	0.964	0.012	82.6%	56.4%
Will Red Star Belgrade win the UEFA Champions League?	69.3	22.2	0.801	0.021	97.8%	38.5%
Will Pierre Poilievre be the next Canadian Prime Minister?	67.4	52.1	1.000	–	0.0%	0.0%
No change in Fed interest rates after September 2025 meeting?	67.4	38.4	1.000	–	0.0%	0.1%
Brighton & Hove Albion wins the Premier League?	67.3	20.3	0.990	0.092	27.2%	5.3%
Will Austin Scott be the first elected Speaker of the House for the 119th congress?	67.3	2.5	0.863	0.002	99.7%	0.1%

Table 13: The market-specific threshold θ_m^*, relative spillover $\Upsilon_m(\theta_m^*)$, and estimated wash fraction of share volume—under Algorithm 2 and the volume-matching algorithm of Victor and Weintraud ((2021))—for the 50 largest markets by share volume (together accounting for 29.4% of overall historical share volume, and 23.0% of dollar volume).

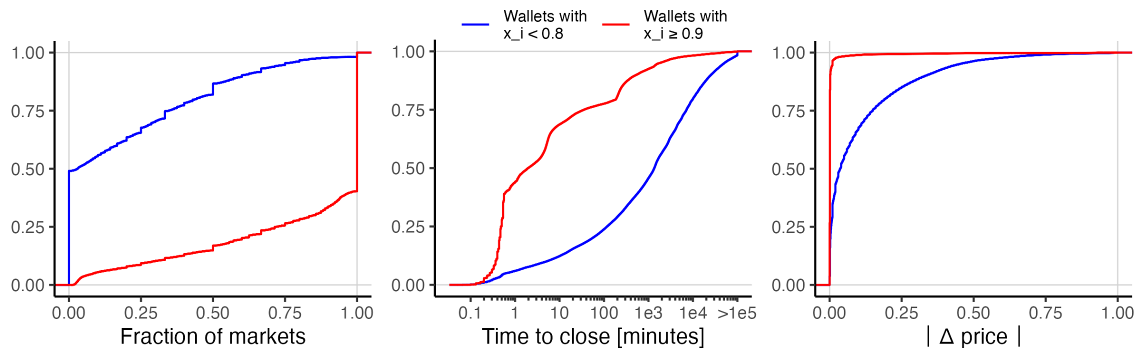


Figure 32: CDFs for the fraction of markets in which a wallet closed its position at least once (left); the median time in minutes until a wallet's open position in a market is closed, conditional on closing at least once (middle); and the median absolute price change between the opening and closing of positions, conditional on closing at least once (right), for wallets with final iteration score $x_i \geq 0.9$ versus wallets with $x_i < 0.8$.

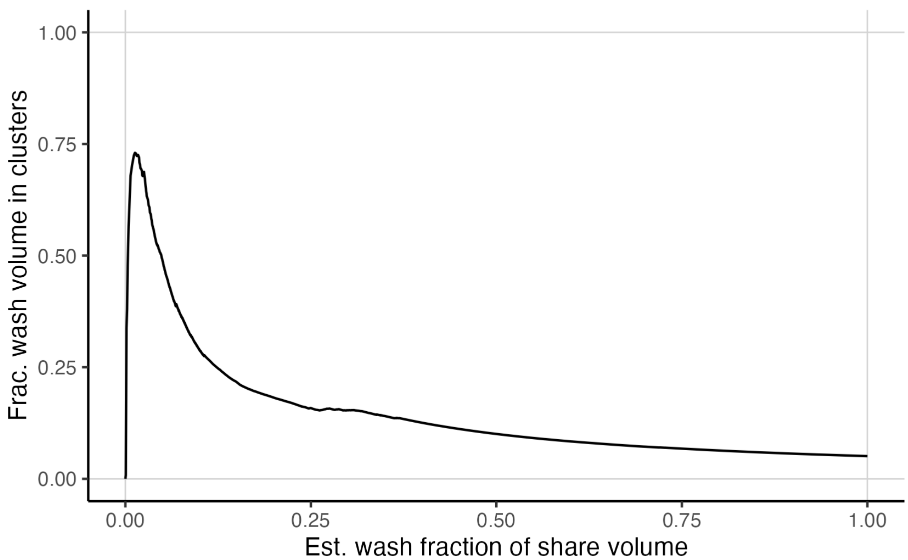


Figure 33: The fraction of detected wash volume that belongs to the clusters identified in Table 10, as a function of the overall wash fraction of share volume by Algorithm 1, as the score threshold θ is varied from 1 to 0. (E.g., at $\theta = 0$, all volume is classified as wash, and within-cluster volume represents approximately 5% of this total.)

display_name	wallet_id	first_funded	funded_by	funding_amt
fengchu	0xc469af64e33e8e1cabe2cb761ad1c3552f29dd61	2024-10-22 06:28:50	?	36599.20
fdetdddw	0x89e3586dcba73e14cfc3b5ffa58b4c10c696e78e	2024-10-24 23:35:57	fengchu	5000.13
duichong	0x70e0896be377fe10f851c4f2222410261823e809	2024-10-24 23:54:55	fengchu	5000.12
DuiChong1	0xbc9d3c09c22165bfe64485bc76e59614984c5b17	2024-10-24 23:56:41	fengchu	5000.12
duic	0xd70736442ede63aebe803e5fa45560afb9832121	2024-10-25 00:19:43	fengchu	5000.12
miya	0x176d7d1186c4728cff1ddb7c812b745265864a26	2024-10-25 02:54:09	fengchu	5000.20
DuiDui	0xea831a531e8eb046efdc3d27864f4f995e08d96b	2024-10-25 03:13:13	fengchu	5000.22

Table 14: Direct USDC transfers relating to the wallet called fengchu (URL: https://polymarket.com/profile/0xc469af64e33e8e1cabe2cb761ad1c3552f29dd61), viewable on Polygonscan. The funding amount may be split across multiple transactions, in which case the time of initial funding is shown.

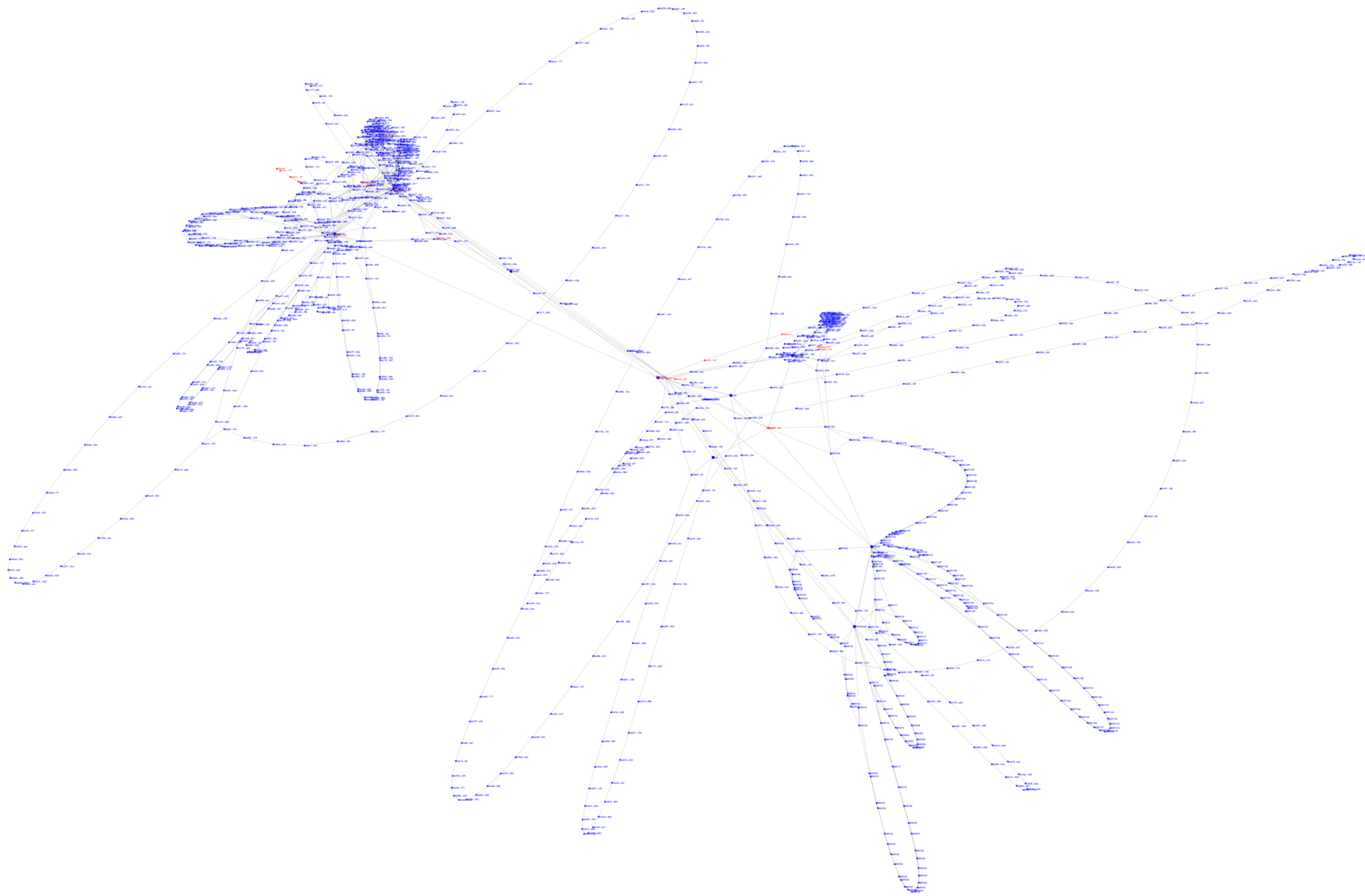


Figure 34: The graph of direct USDC transfers (i.e., which do not involve Polymarket) for the large wallet cluster originating from the wallet with display name “fengchu” (`0xc469af64e33e8e1cabe2cb761ad1c3552f29dd61`). Wallets which trade on Polymarket are colored blue, while non-trading wallets (which may represent token exchanges which facilitate transfers) are colored red.

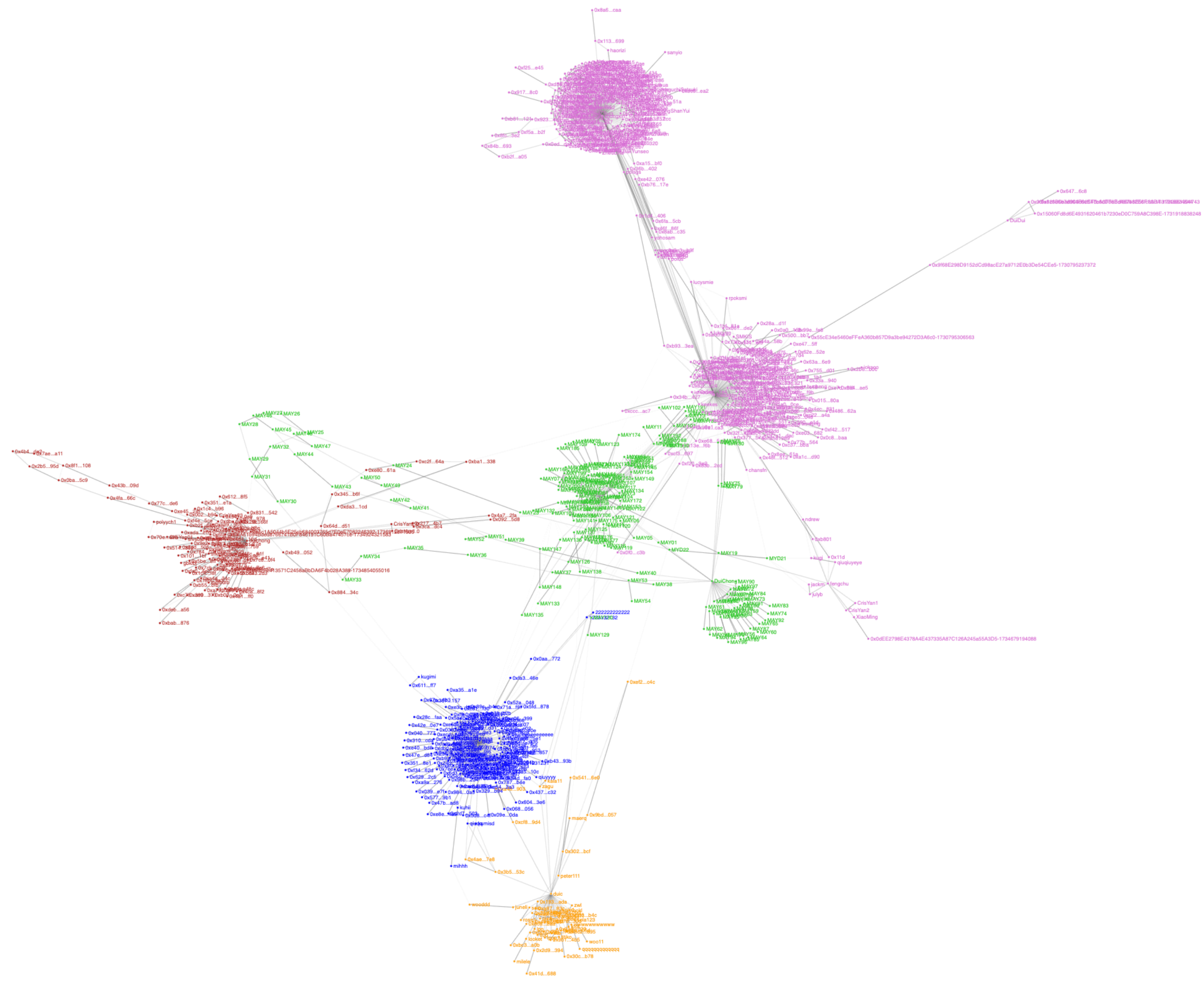


Figure 35: The trade graph for the large wallet cluster originating from the wallet with display name "fengchu" (`0xc469af64e33e8e1cabe2cb761ad1c3552f29dd61`), which includes clusters "MAY" (green), "miya" (blue), "duic" (orange), "duichong" (red), and "zhongxin" (orchid); see Table 10. A directed edge with weight $w_{ij} \in (0, 1]$ indicates that a fraction w_{ij} of i's total share volume traded was traded with counterparty j.

Appendix C Proofs

Proposition 1. *The sequence of score vectors $\{\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \dots\}$ in Algorithm 2 converges to the $\boldsymbol{x}$ which satisfies the stationary equation*

$$\boldsymbol{x} = \frac{1}{2}(\boldsymbol{x}^{(0)} + \boldsymbol{B}\boldsymbol{x}).$$

Proof. Let $\boldsymbol{x}$ be the unique point satisfying the stationary equation. The iterative update is given by $\boldsymbol{x}^{(k+1)} = \frac{1}{2}(\boldsymbol{x}^{(0)} + \boldsymbol{B}\boldsymbol{x}^{(k)})$.

Subtracting the first equation from the second gives the evolution of the error, $\boldsymbol{e}^{(k)} = \boldsymbol{x}^{(k)} - \boldsymbol{x}$:

$$\boldsymbol{x}^{(k+1)} - \boldsymbol{x} = \frac{1}{2}(\boldsymbol{x}^{(0)} + \boldsymbol{B}\boldsymbol{x}^{(k)}) - \frac{1}{2}(\boldsymbol{x}^{(0)} + \boldsymbol{B}\boldsymbol{x}) = \frac{1}{2}\boldsymbol{B}(\boldsymbol{x}^{(k)} - \boldsymbol{x}).$$

This defines a fixed-point iteration on the error, $\boldsymbol{e}^{(k+1)} = \frac{1}{2}\boldsymbol{B}\boldsymbol{e}^{(k)}$.

By the Gershgorin Circle Theorem, the spectral radius $\rho(\boldsymbol{B}) \leq 1$, therefore $\rho(\frac{1}{2}\boldsymbol{B}) < 1$, such that $\boldsymbol{e}^{(k)} \to \boldsymbol{0}$ as $k \to \infty$. Thus, $\{\boldsymbol{x}^{(k)}\}$ converges to $\boldsymbol{x}$. $\square$

Proposition 2. *The volume-weighted score $\boldsymbol{v}^\top \boldsymbol{x}^{(k)}$, where element v_i of $\boldsymbol{v}$ is wallet i's total share volume traded, is preserved across iterations $k \in \{0, 1, \dots\}$ of Algorithm 2.*

Proof. Consider again the iterative score update

$$\boldsymbol{x}^{(k+1)} = \frac{1}{2}\left(\boldsymbol{x}^{(0)} + \boldsymbol{B}\boldsymbol{x}^{(k)}\right),$$

and notice that $\boldsymbol{v}^\top \boldsymbol{B}_{\cdot j}$ is the total volume traded by counterparty j, such that $\boldsymbol{v}^\top \boldsymbol{B} = \boldsymbol{v}^\top$, i.e., $\boldsymbol{v}$ is a left eigenvector of $\boldsymbol{B}$. Multiplying both sides of the iterative score update by $\boldsymbol{v}$,

$$\boldsymbol{v}^\top \boldsymbol{x}^{(k+1)} = \frac{1}{2}\left(\boldsymbol{v}^\top \boldsymbol{x}^{(0)} + \boldsymbol{v}^\top \boldsymbol{B}\boldsymbol{x}^{(k)}\right) = \frac{1}{2}\boldsymbol{v}^\top \left(\boldsymbol{x}^{(0)} + \boldsymbol{x}^{(k)}\right). \tag{10}$$

The result can be shown using induction. For the base case, observe that

$$\boldsymbol{v}^\top \boldsymbol{x}^{(1)} = \frac{1}{2}\boldsymbol{v}^\top \left(\boldsymbol{x}^{(0)} + \boldsymbol{x}^{(0)}\right) = \boldsymbol{v}^\top \boldsymbol{x}^{(0)}.$$

Applying the inductive hypothesis to (10), we have for all $k \in \{0, 1, \dots\}$

$$\boldsymbol{v}^\top \boldsymbol{x}^{(k+1)} = \frac{1}{2}\boldsymbol{v}^\top \left(\boldsymbol{x}^{(k)} + \boldsymbol{x}^{(k)}\right) = \boldsymbol{v}^\top \boldsymbol{x}^{(k)} = \boldsymbol{v}^\top \boldsymbol{x}^{(0)}.$$

$\square$

Appendix D Additional Trade Graph Examples

As in Figure 5, each node i is a wallet with circular area proportional to $\sqrt{\text{share volume}_i}$, and an edge with weight $w_{ij} \in (0, 1]$ indicates that a fraction w_{ij} of i's volume traded in the market was traded with counterparty j. (Note that this means there are overlapping edges pointing in opposite directions.) As specified by Algorithm 1, the calculation of wallets' scores uses information from *all* markets, not just the market shown. In market m, wallets with $r_{im} \geq \theta_m^*$ are colored red; all other wallets are colored blue. (See (3) for the definition of r_{im}.) Wash trades are represented by edges between red nodes.

Figure 36: Will Marco Rubio win the 2024 Republican VP nomination? (January 30, 2024–July 15, 2024)

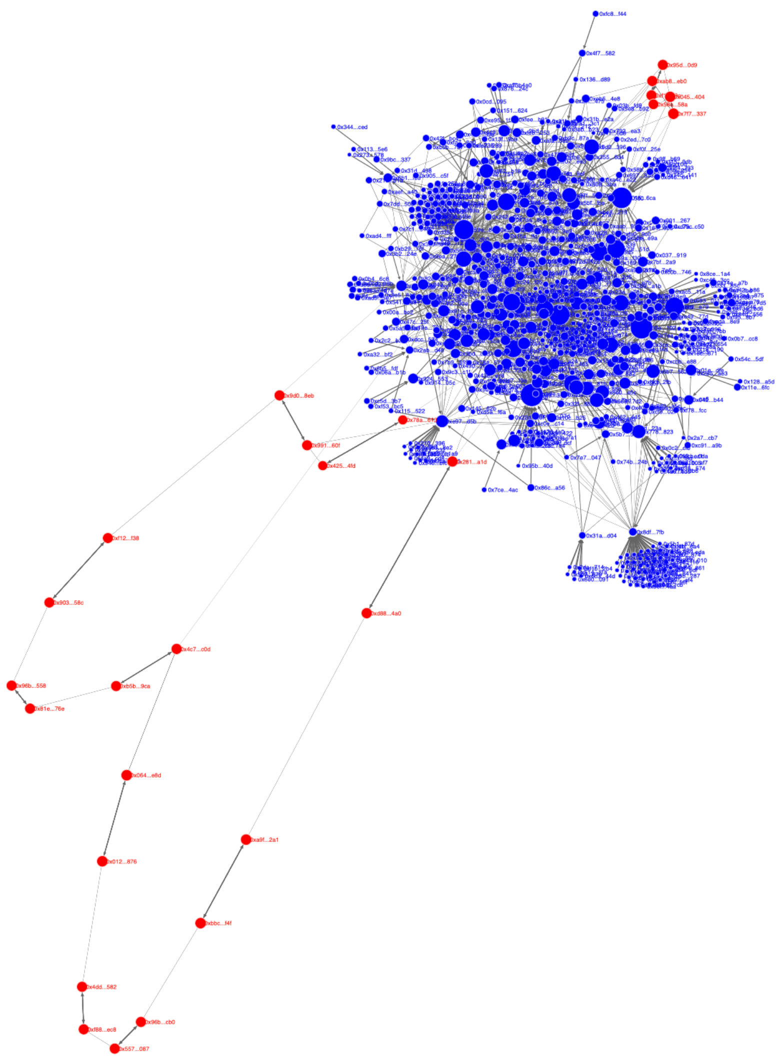


Figure 37: Will another state be the closest state [in the 2024 U.S. Presidential Election]? (October 8, 2024–December 17, 2024)

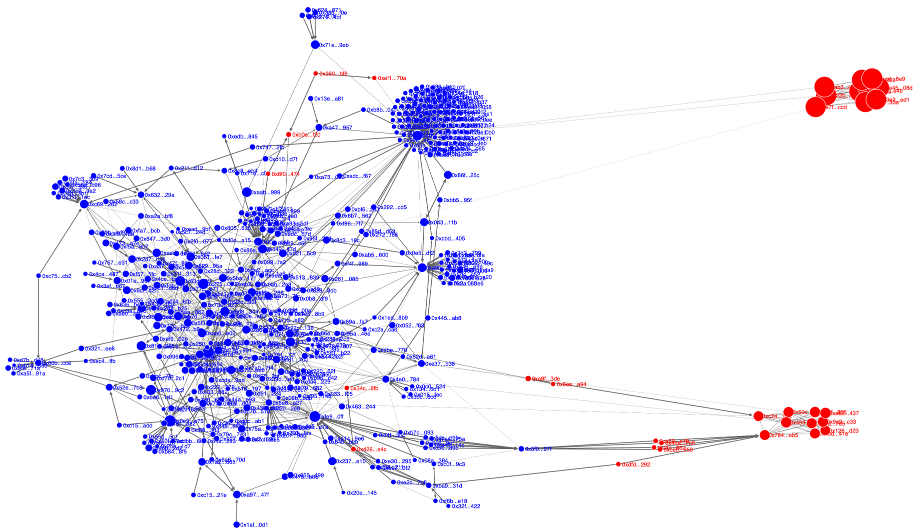


Figure 38: Will "Good Luck, Babe!" win Record of the Year? (November 20, 2024–February 3, 2025)

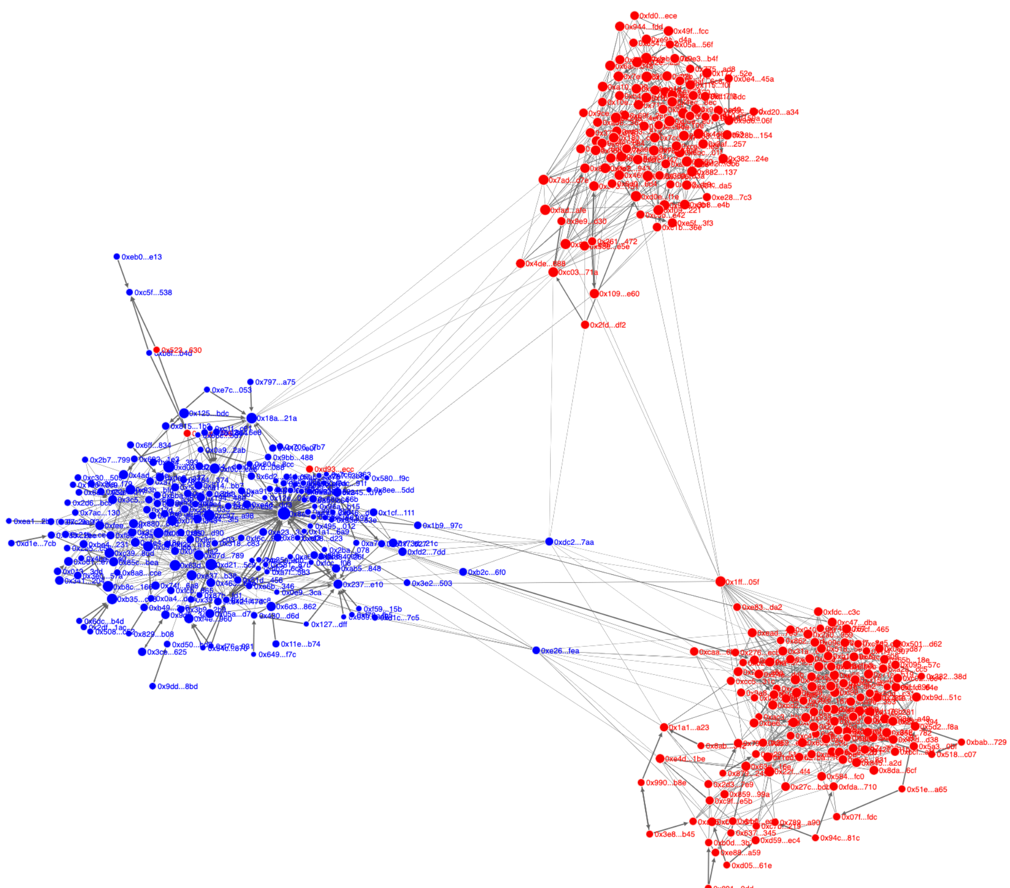


Figure 39: Will AS Roma win the Serie A? (October 15, 2024–April 27, 2025)

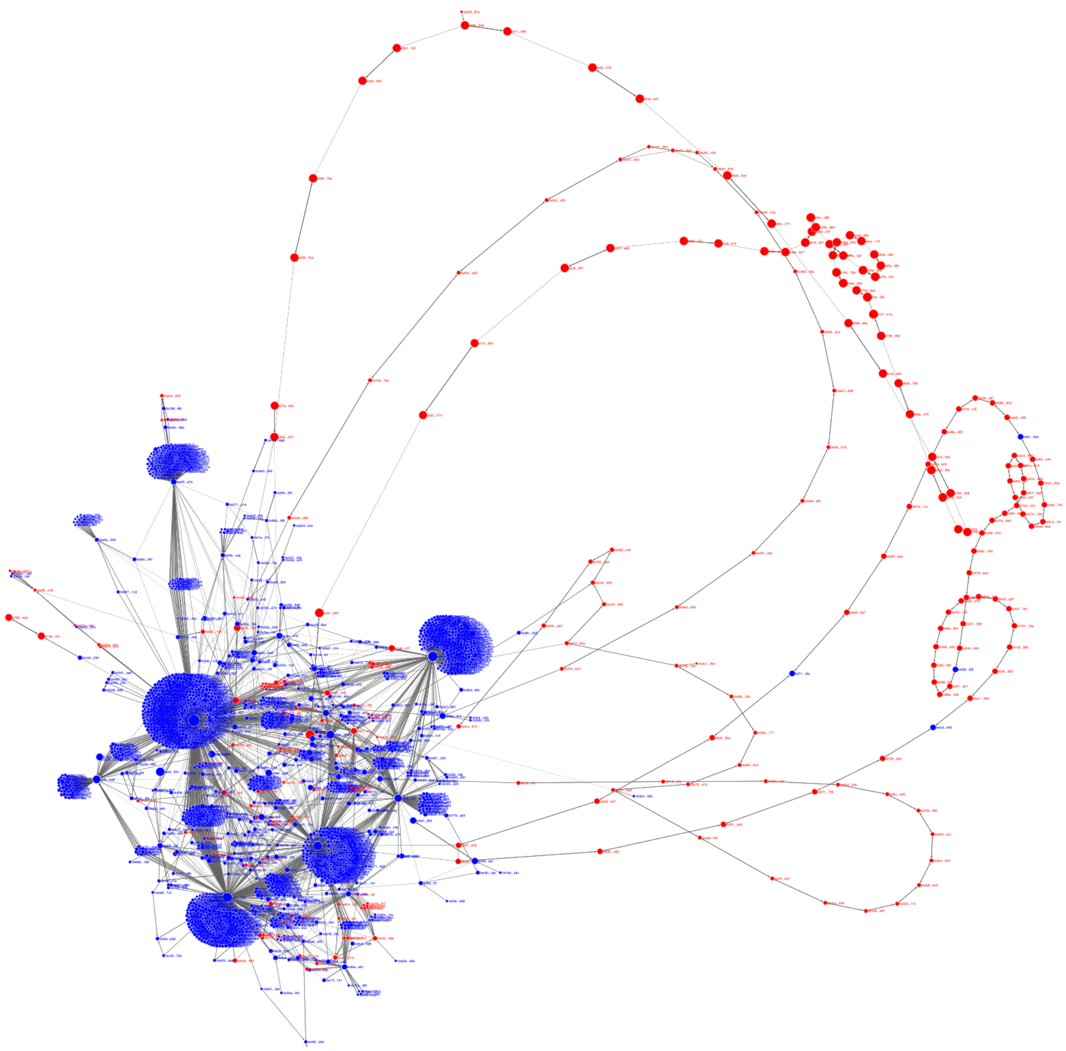


Figure 40: Will SPD, FDP, and Greens form the next German Government? (January 28, 2025–May 6, 2025)

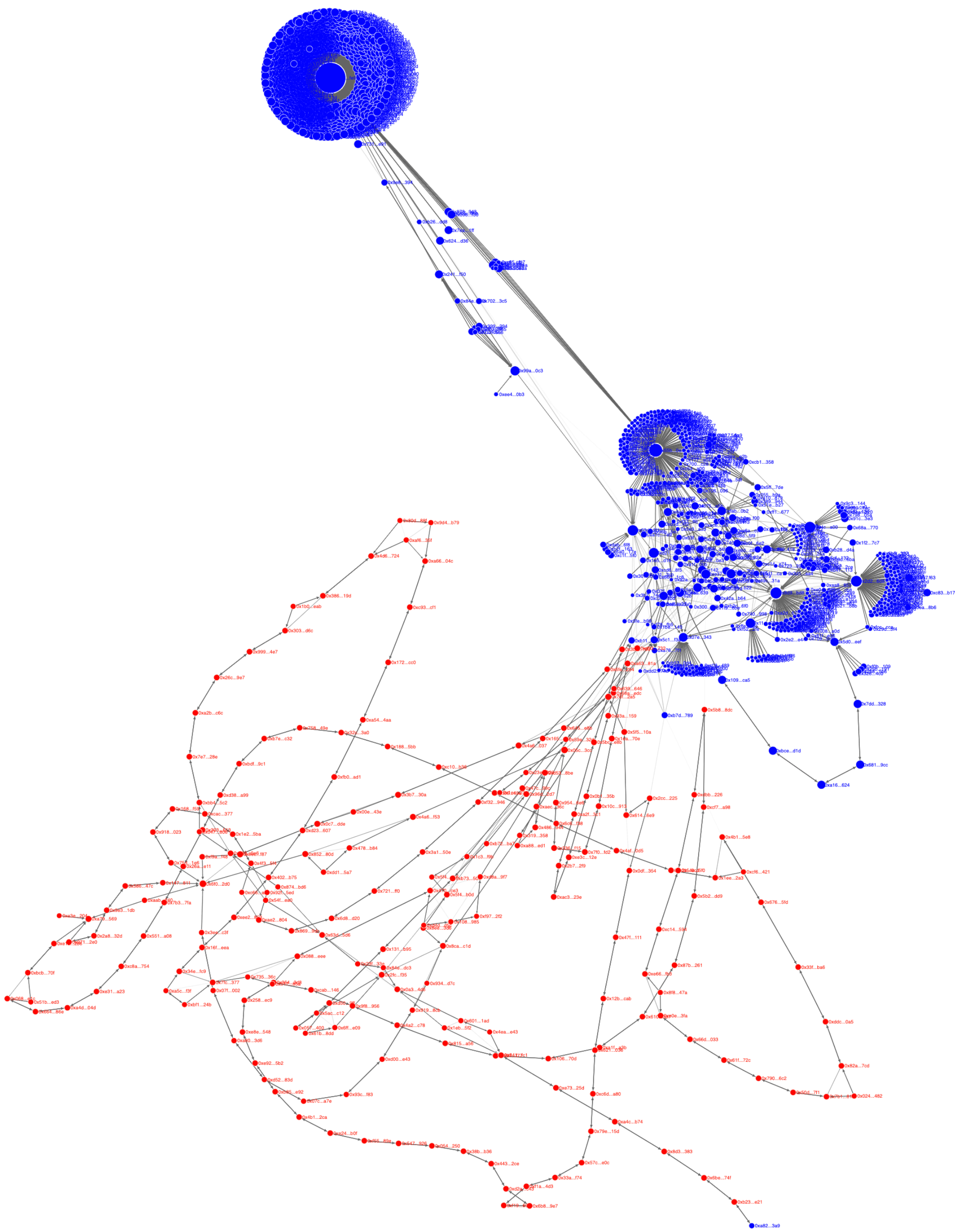


Figure 41: Will Phil Murphy be the next DNC chair? (November 19, 2024–February 2, 2025)

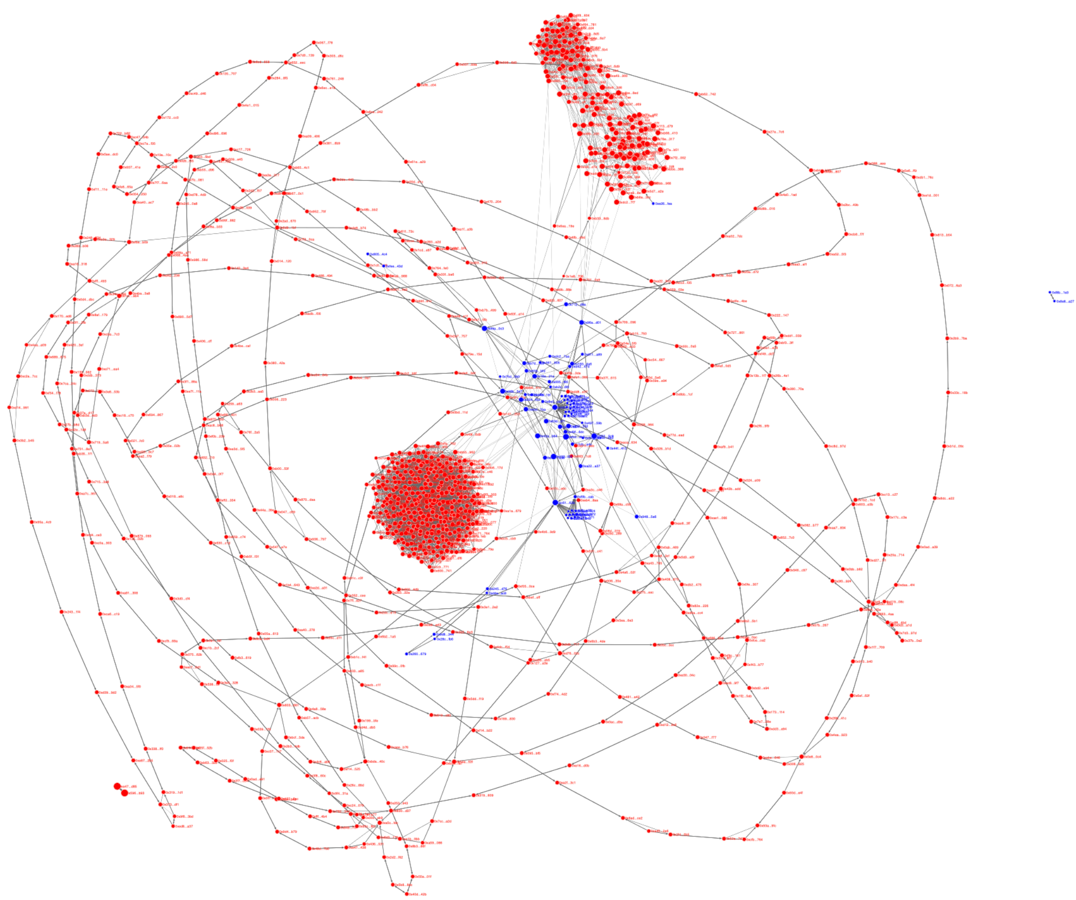


Figure 42: Will the Denver Nuggets win the 2025 NBA Finals? (September 24, 2024–May 18, 2025)

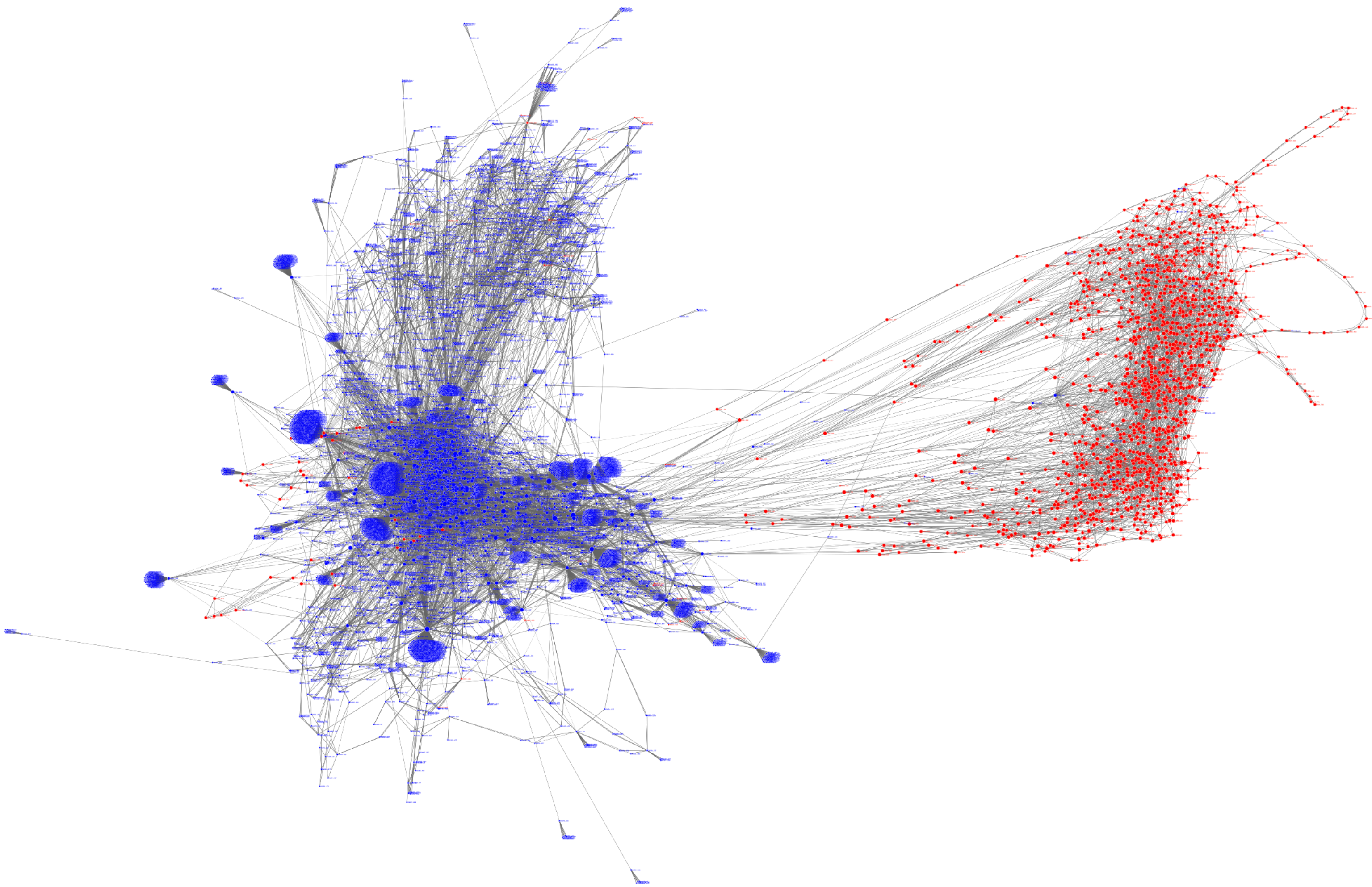


Figure 43: Will a Democrat win Colorado in the 2024 U.S. Presidential Election? (March 28, 2024–November 6, 2024)

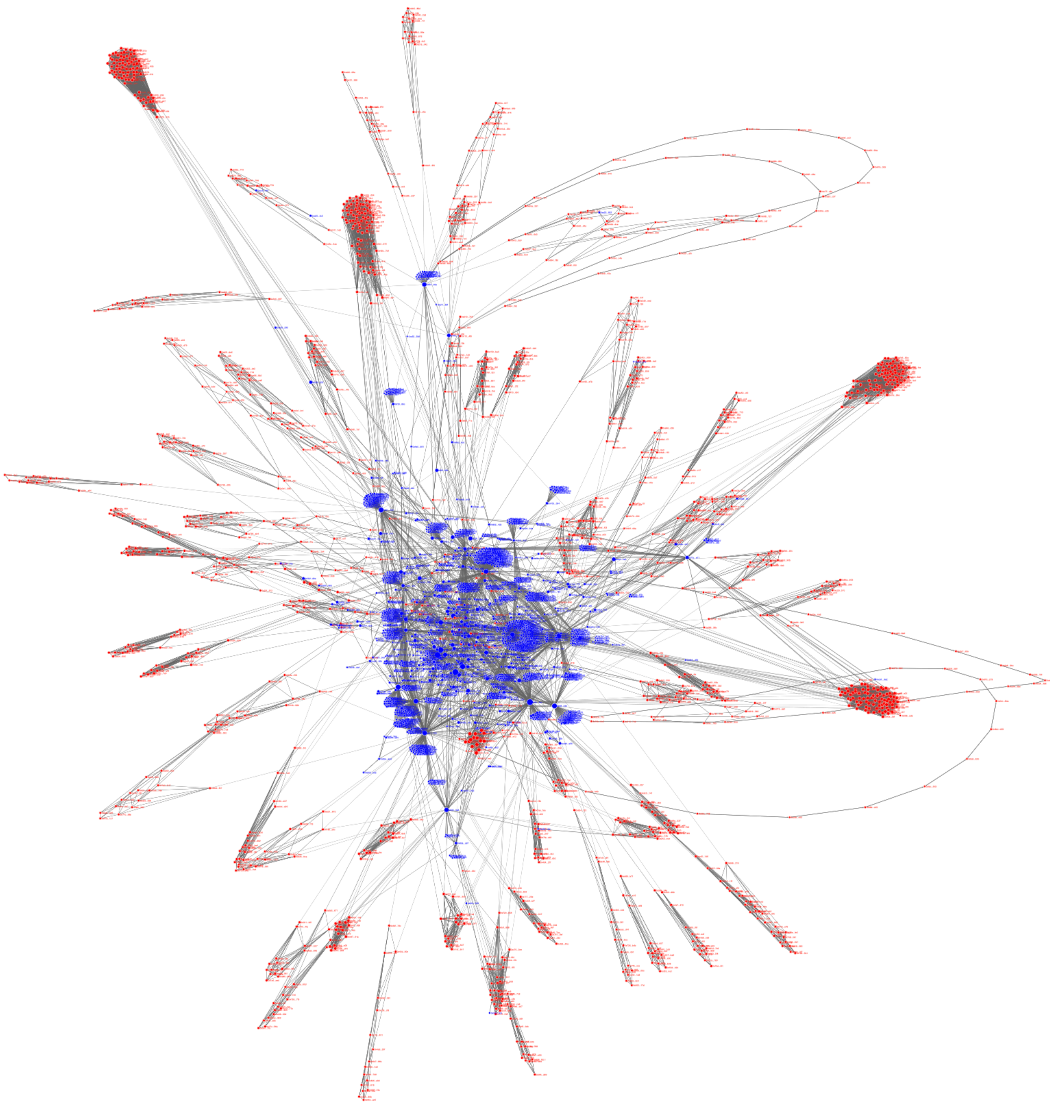


Figure 44: Will Brad Lander win the Democratic Primary for Mayor of New York City? (December 30, 2024–June 25, 2025)

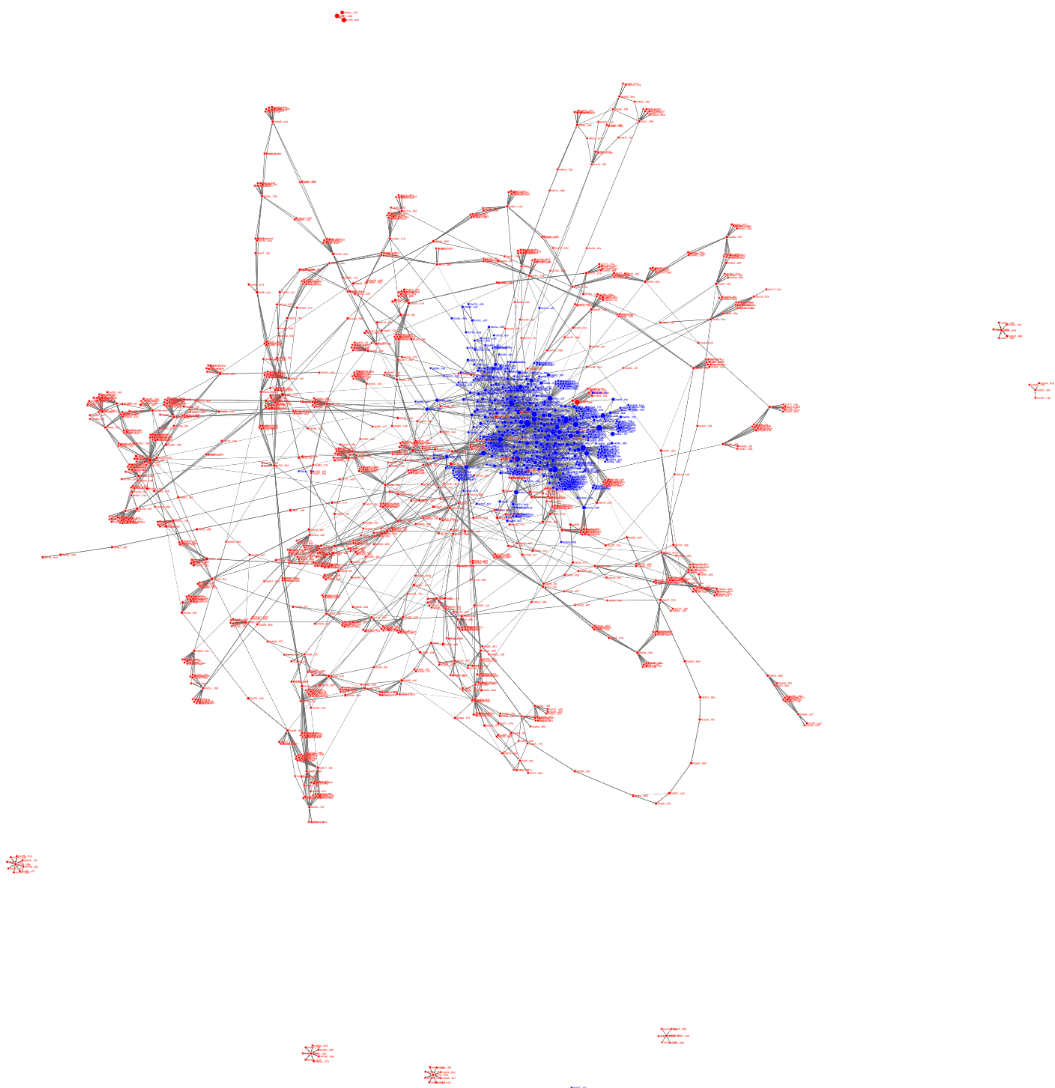


Figure 45: Will the San Jose Sharks win the 2025 Stanley Cup? (October 8, 2024–June 23, 2025)

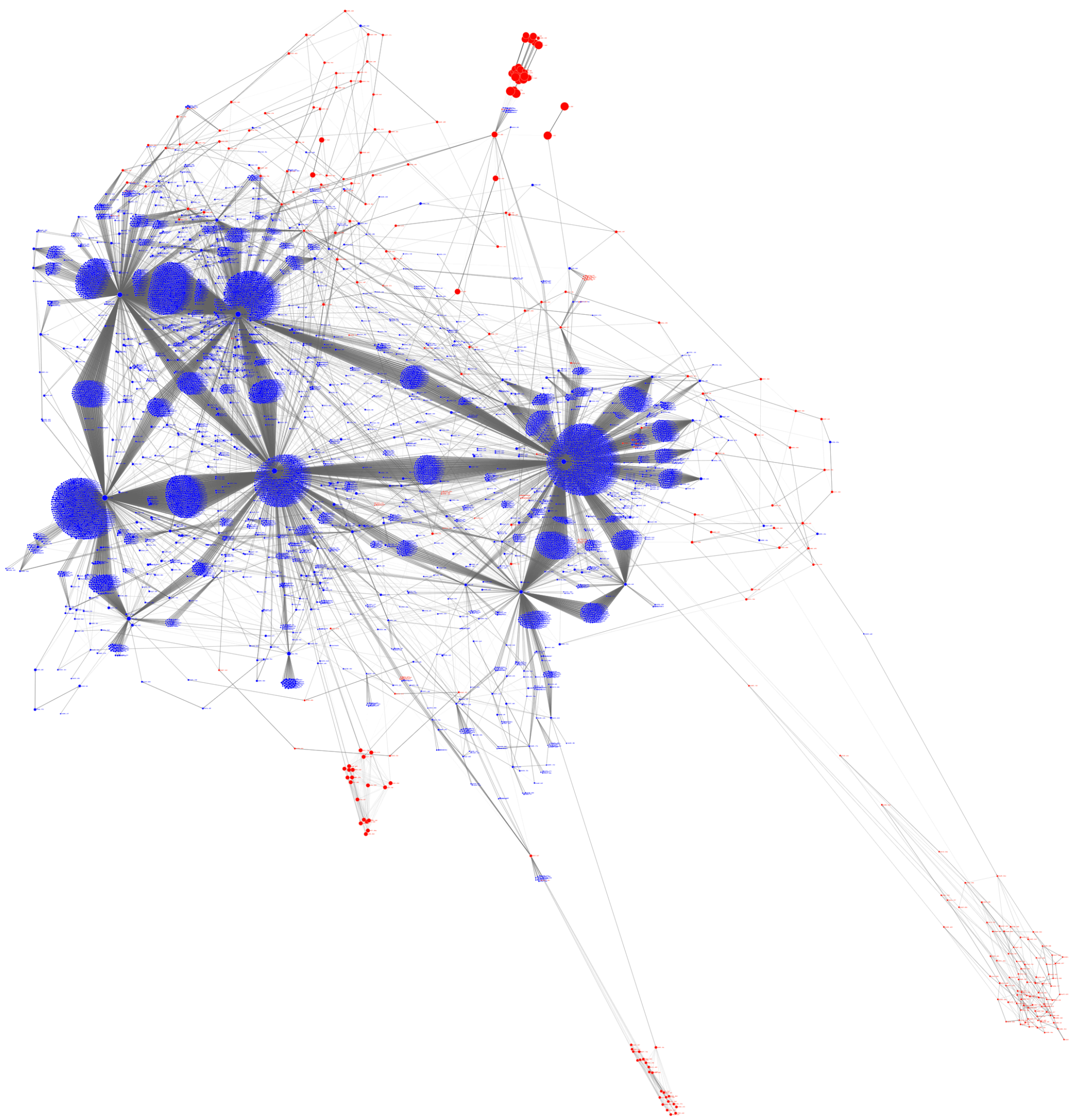


Figure 46: Will Nevada be the tipping point state [in the 2024 U.S. Presidential Election]? (July 16, 2024–November 5, 2024)

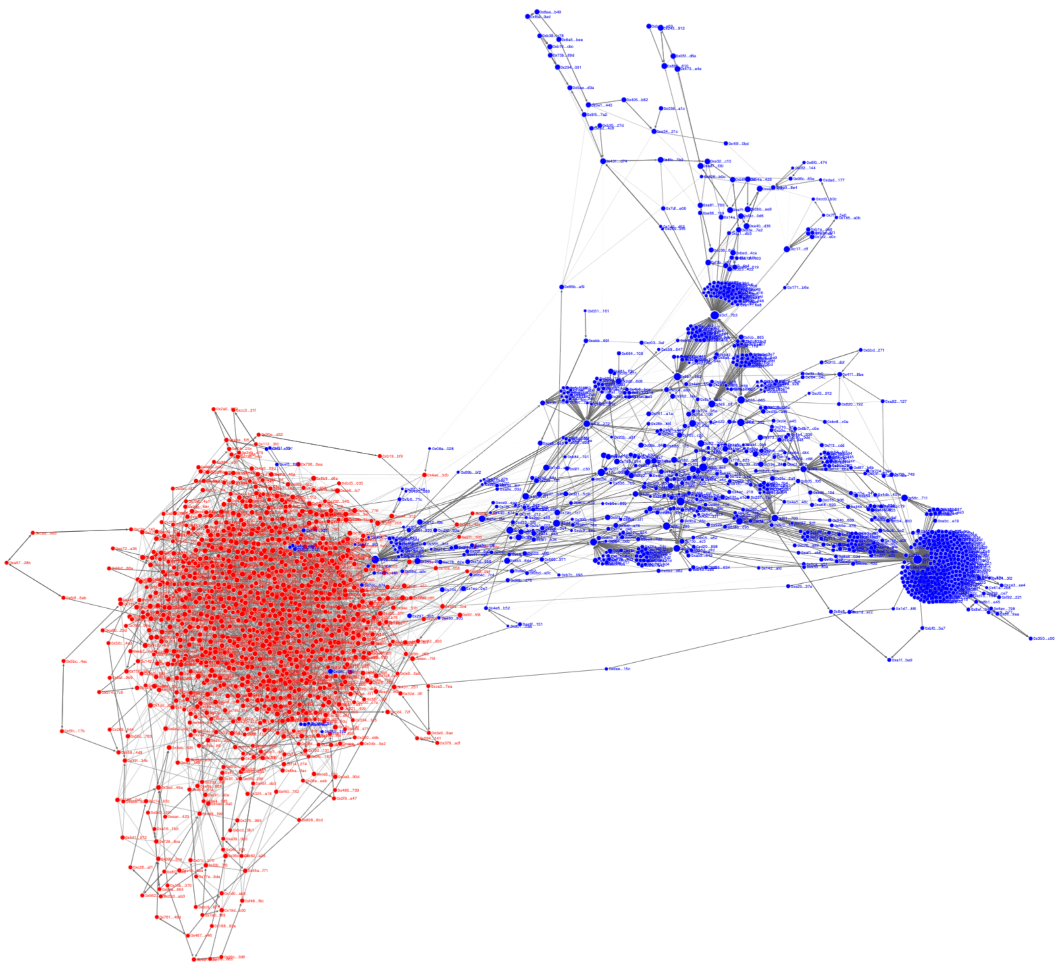


Figure 47: Will Lee Jun-seok win 2nd place in the South Korean presidential election? (May 5, 2025–June 3, 2025)

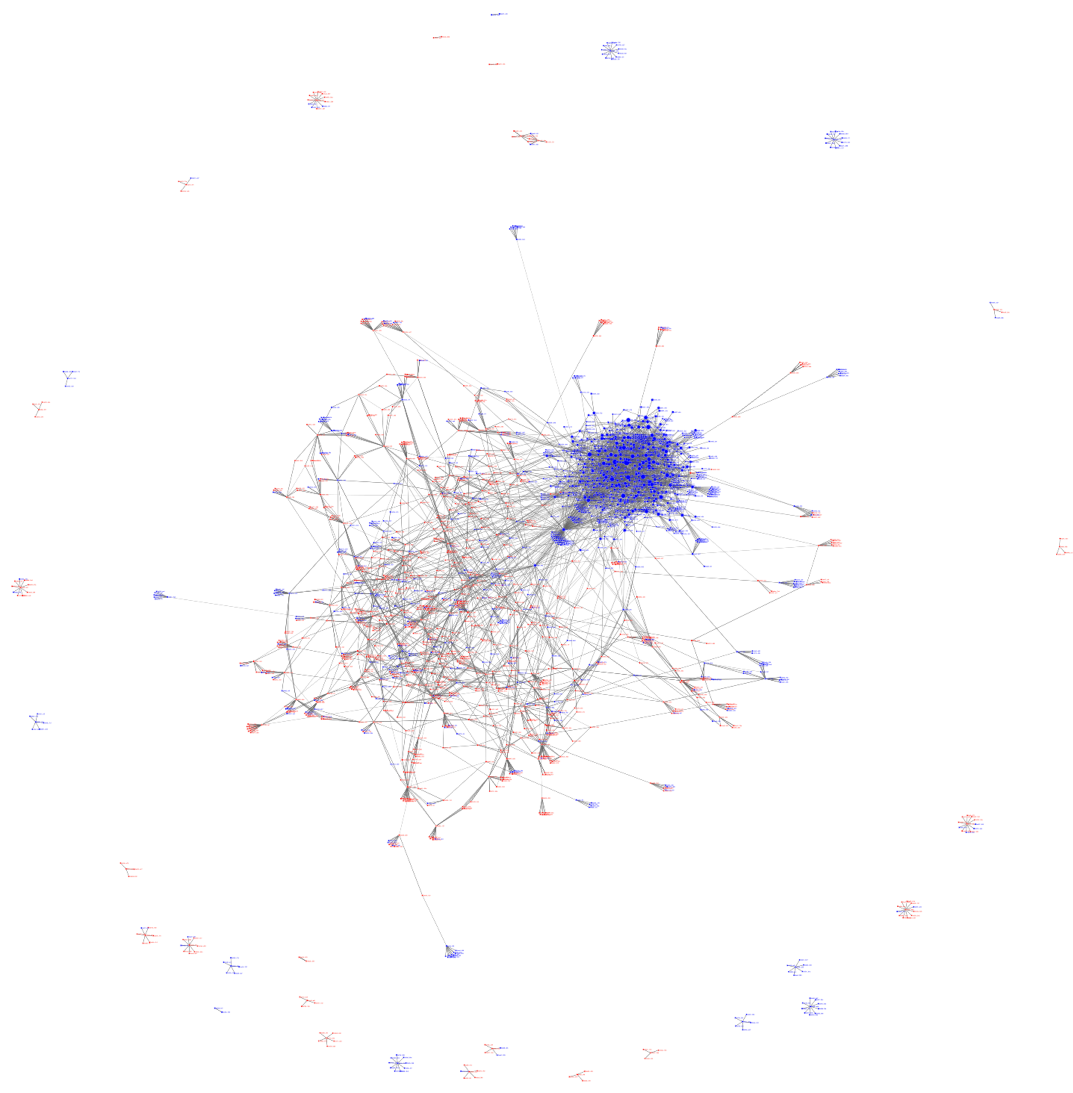


Figure 48: Israel x Hamas ceasefire before May? (March 19, 2025–April 30, 2025)

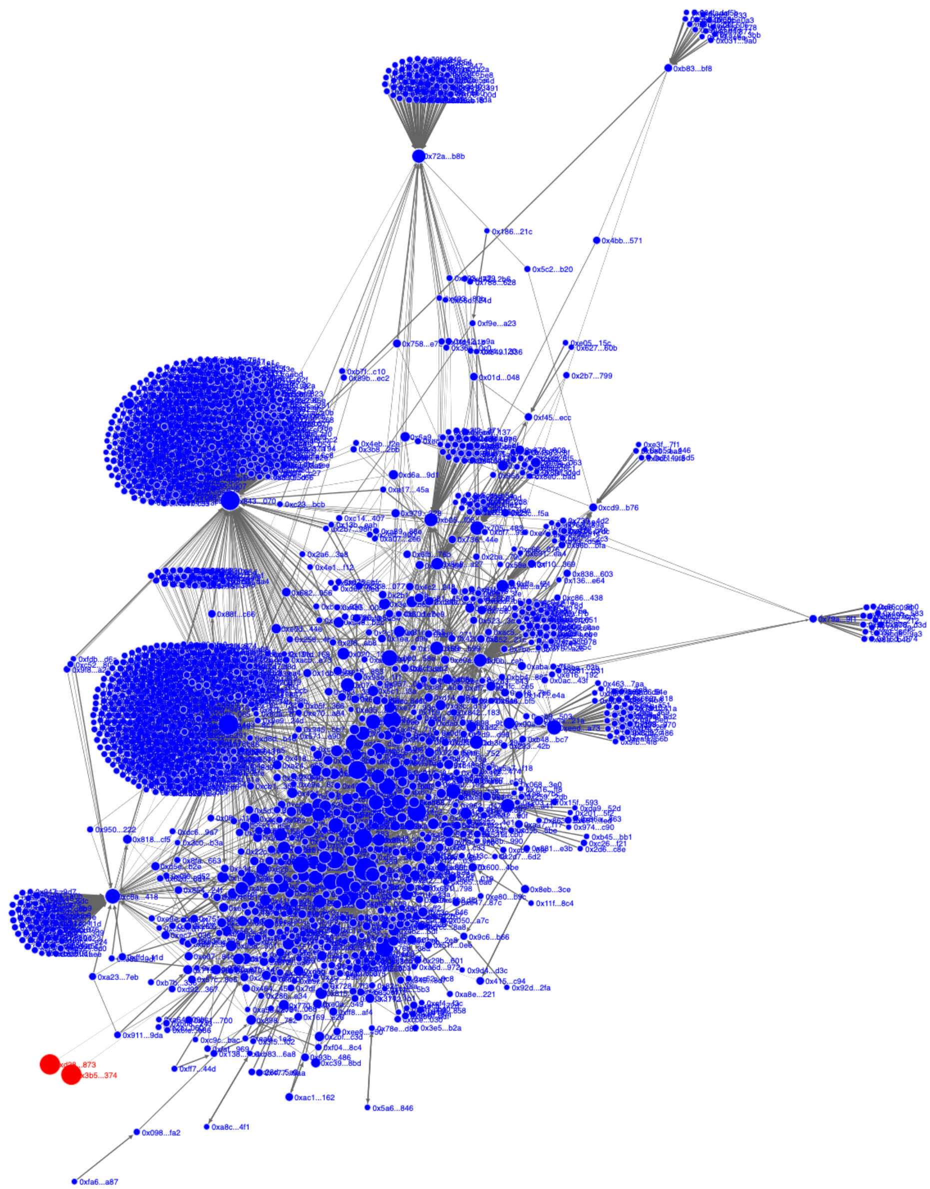


Figure 49: Will Maine be the tipping point state [in the 2024 U.S. Presidential Election]? (July 16, 2024–November 5, 2024)

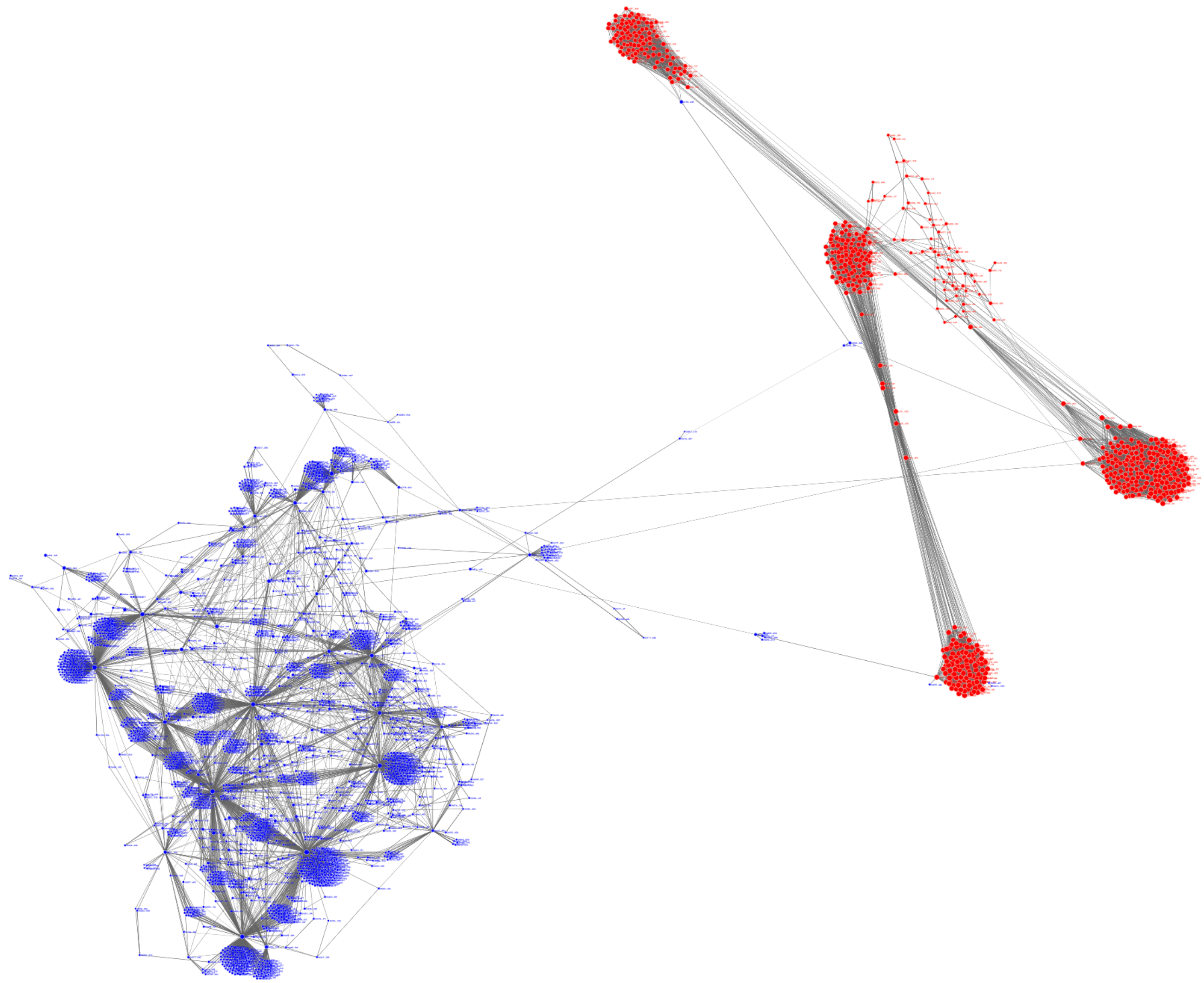


Figure 50: Will Austin Scott be the first elected Speaker of the House for the 119th congress? (November 6, 2024–June 30, 2025)