

Where unusual Polymarket wins concentrate

An exploratory comparison of published earnings accounts, market-implied expectations and broader wallet screens

8 October 2026 | Historical observation window: 1 January 2025 - 28 April 2026

Abstract

A successful wallet is easier to interpret when its record is compared with the probabilities it paid for and the subjects it traded. In thirteen previously published wallet addresses, pooled initial positions won in 32 of 37 company reporting cycles involving KPMG clients, against 23.331 wins implied by entry prices. The same group won in 28 of 55 cycles involving other known auditors' clients, against 30.704 expected. Separately, a subject-concentration screen selected 321 of 22,322 eligible wallets, compared with a mean of 51.252 in 500 simulations; weather and daily-temperature subjects accounted for 162 selections. These are retrospective departures from stated benchmarks. They do not establish insider trading, common ownership or a reliable way to identify informed traders in advance. No tested population rule retained all studied case wallets. We set out the units, failed screens, corrected concentration calculation and the next checks needed before stronger interpretation.

Three results to keep separate

Question	Result	Interpretation
Where did a published group outperform its entry prices?	32 / 37 wins on KPMG clients; 23.3 expected. Other known auditors: 28 / 55; 30.7 expected.	A retrospective comparison within thirteen selected wallet addresses.
How often did a broader subject screen qualify wallets?	321 of 22,322 eligible wallets; simulated mean 51.3.	A departure from a stated model, not an insider count.
Did the strict 48-hour screen beat matched established wallets?	1,316 bets; matched gap -5.40 cents per share.	The corrected average remains negative.

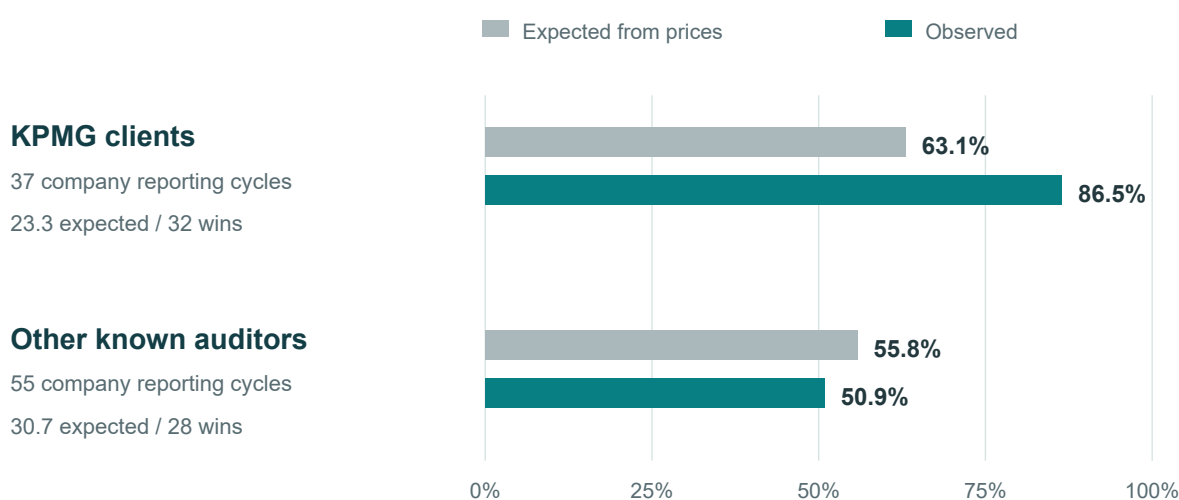
Status. Exploratory research; not peer reviewed. The figures are checked against saved outputs of analyses completed on 5-6 October. This publication is not a new raw-data rerun. No author or reviewer outside Vultax is credited or implied.

1. The earnings comparison

The thirteen addresses were published before this analysis: nine in the initial EventWaves account list and four in supplementary sections [3]. We matched those addresses to the fill archive. The group is a reproducible selection of published addresses; it is not verified common ownership, an exact reconstruction of every account mentioned in press reporting, or a finding about an audit firm.

The earnings roster contains 891 markets, of which 756 resolved, covering 423 companies and 756 company reporting cycles. Annual auditors come from PCAOB AuditorSearch/Form AP [2]. Company and ticker identity used SEC public entity records and exact issuer-name matches. An annual-auditor match does not identify quarterly-review staff or establish personal access to information.

Figure 1. Initial-direction win rates, pooled once per company reporting cycle



The two rows use the same thirteen addresses. Expected rates sum the recorded initial entry probabilities and divide by the cycle count. Multiple group wallets in one company cycle do not create additional trials. Source: saved Vultax F2 results; headline-data.csv.

Counting and benchmark

Take the earliest directional entry made by any group member in each company reporting cycle, and count the binary result of that initial direction once. For entry probabilities p and results y , expected wins $E = \sum(p)$, observed wins $W = \sum(y)$, and excess wins $= W - E$. The KPMG-client difference is $32 - 23.331 = 8.669$ wins; the comparison difference is $28 - 30.704 = -2.704$.

These are not counts of profitable trading histories. Later exits, hedges, position sizes and fees do not enter this initial-direction win statistic. Summing expectations does not require independence, but probability-tail calculations and uncertainty estimates do require assumptions about dependence and selection.

2. What the comparison can establish

The same selected group has an observed win rate of 86.5% on KPMG clients, against 63.1% implied by entry prices: a difference of 23.4 percentage points. Elsewhere the rates are 50.9% and 55.8%, a difference of -4.9 points. We do not attach a confirmatory significance claim to the difference between those gaps.

The concentration is also visible in directional acquisition cost: 80.55% of the group's main-earnings cost falls on KPMG clients. KPMG accounts for 74 of the 423 mapped companies, or 17.49%. Company share is context, not a uniform random-choice model; companies have different numbers of contracts, liquidity and trading opportunities.

Sensitivity checks already completed

For each auditor, twelve threshold combinations cross two probability-tail ceilings (0.01 and 0.001), three minimum cycle counts (3, 5 and 8), and two cost-share gates (50% and 80%). The selection counts are compared with 1,000 company-label permutations. A company's quarters move together and each auditor's company count is preserved.

Five of the twelve KPMG cells exceed their own shuffled 95th percentile; none of the other auditors' cells does. The cells overlap, and the comparison is not adjusted for searching across them. The shuffles do not preserve sector or company-size mix. The corresponding earnings-call-mentions analysis produces no KPMG selections; that separate roster contains 906 markets, 885 resolved, and 56 companies.

The important limits

Selection: the thirteen addresses were already identified for unusual activity. They are not a random sample or a held-out test set.

Information: wallet transactions and auditor labels do not establish a person's job, access, knowledge, identity or intent.

Comparability: entry prices may be miscalibrated, and company groups can differ in size, sector, liquidity and event difficulty.

Dependence: pooled company cycles reduce duplicate counting but do not make related companies, reporting periods or market outcomes independent.

The useful next question

Does the concentration remain when companies are compared within sector and size groups, with entry-price and timing differences retained? If it does, freeze a rule and evaluate it on a later period that has not been examined. That test is proposed work, not a completed validation or a promised outcome.

3. The broader subject screen

The full archive has 714,691,643 fills, 834,732 markets and 2,404,822 wallet addresses [1]. After excluding sports, esports and fixed-clock crypto, the subject search retains 135,284 markets (122,509 resolved within the tape) and 118,158,616 fills. Wallet addresses, fills, bets and people are different units. Metadata are a current Gamma snapshot [4], not historically frozen tags.

Choose each wallet's dominant retained subject by its share of observed lifetime directional acquisition cost. Market and event tags are combined after removing broad and administrative labels. The cost denominator includes excluded categories and unresolved positions. It is directional turnover, not deposits or maximum capital at risk.

Reference eligibility, before testing success	Requirement
Dominant-subject share	At least 80% of observed lifetime directional cost
Timing and history coverage	Subject entry within 3 days of first observed activity; first observed activity on or after 1 February 2025
Trials and size	At least 3 binary resolved subject events and \$2,500 resolved subject cost
Success rule	Poisson-binomial upper tail at most 0.01, using recorded entry probabilities

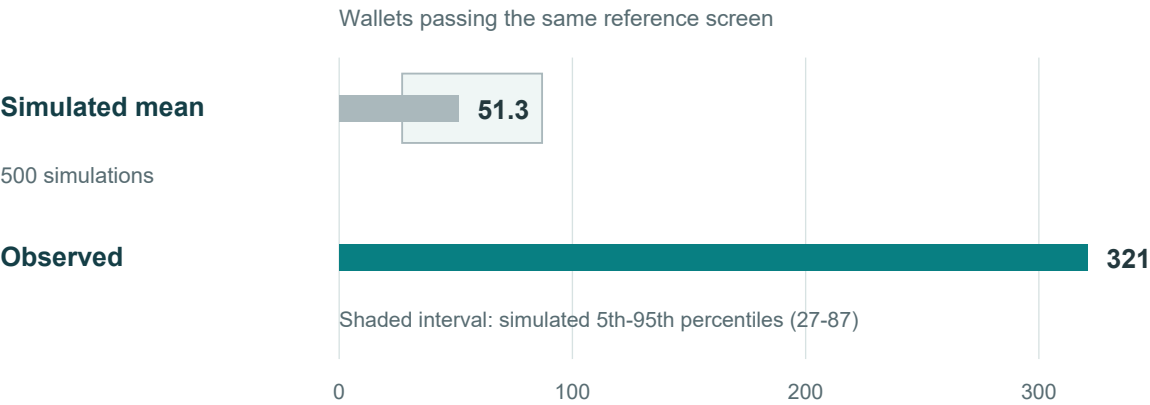


Figure 2. Reference cell 28. There are 22,322 eligible wallets; 321 qualify. The observed count is about 6.3 times the simulated mean of 51.252. Weather and daily-temperature subjects account for 162 selections. The shaded band describes simulated counts, not a causal effect.

In each of 500 simulations (seed 20261005), eligibility, subjects and entries remain fixed. One uniform draw per Gamma event is reused across wallets: the reference direction wins when $U < p$ and the opposite when $U \geq 1 - p$. Different events receive independent draws. The success rule is reapplied to each wallet. At different entry prices this construction preserves the probability marginals and some shared-event dependence, but does not recreate one literal settlement for every entry.

Cross-event dependence, adaptive trading and price miscalibration are absent from that model. The excess can therefore reflect model mismatch, skill or public information. A 0.01 screen threshold is not a 99% probability of misconduct.

4. Failed screens and a correction

The strict rule requires at least \$2,500 of rolling-hour stake at 5-35 cents, observed history under 14 days (first seen from 1 February), at most one unrelated prior market, at least 90% directional-cost concentration, and entry within 48 hours of the retrospective outcome clock. The original concentration gate used raw token notional. Each wallet leg instead needs shares times its own held-outcome price; the two legs can have very different costs.

The correction removes 24 bets and adds 9, changing 1,331 to 1,316. The corrected group has 1,307 wallets and 716 markets. It wins 264 times (20.06%) against 297.36 expected from prices (22.60%). The matched average gap is -5.40 cents per share, with a 95% market-bootstrap interval of -10.53 to -0.07 cents. The matched comparison covers 93.3% of the selected bets.

Eight corrected topic/time comparisons

Window / topic grouping	Bets	Matched gap, cents	95% interval, cents
48h / One market	1,316	-5.40	-10.53 to -0.07
48h / One event	1,637	-2.12	-8.07 to +3.97
48h / Event or series	3,297	-0.25	-3.45 to +2.98
48h / Shared tag	10,136	-0.69	-2.13 to +0.87
72h / One market	1,445	-4.62	-9.36 to +0.22
72h / One event	1,806	-1.58	-7.00 to +4.26
72h / Event or series	3,469	-0.08	-3.29 to +3.08
72h / Shared tag	10,440	-0.56	-2.05 to +0.98

Established controls have at least 90 observed days and ten prior markets. Matching uses category, price band, half-year and time-to-outcome bins; intervals resample markets 2,000 times. Related markets can remain dependent. The comparisons use a retrospective clock: the start of the final price run that stays at or above 95 cents, not formal settlement.

Broader topic definitions recover studied examples excluded by the strict single-market rule, but the average matched gaps do not clearly turn positive. This retrospective clock is not an executable real-time forecasting rule.

The first population grid contains 216 combinations of topic, concentration, event count, tail threshold and cost. It retains none of the studied named case wallets. The improved overlapping-subject grid has 72 primary cells and 144 additional history-coverage rows. No one cell retains every studied case wallet. The reference cell in Figure 2 was specified in the exploratory protocol; it is not an independent held-out test.

This note corrects the recomputed strict and topic/time results. It does not validate every older age-confirmed, independent-implementation or other sensitivity table in the earlier article [5]. Those older figures must not be silently combined with the corrected denominator.

5. Reproduction, interpretation and sources

What the evidence supports

The earnings comparison locates a difference within one selected group; the population screen locates an excess relative to a model. Together they motivate studying where unusual success concentrates. Neither identifies a cause, an insider population or a profitable strategy. We have not found a validated single screen that captures every studied case.

Before a stronger claim

Match companies within sector and size groups; retain repeated-company and shared-event structure; test alternative price-calibration and dependence assumptions; report selection, coverage and false alarms. Freeze the rule and evaluate a later untouched period. Report the outcomes even if the initial contrast disappears. These steps remain outstanding.

Companion material and audit trail

The accompanying ZIP contains headline-data.csv, strict-screen-comparisons.csv, earnings-sensitivity.csv, population-sensitivity.csv, subject-breakdown.csv, methodology.txt, reproduce-headlines.py and a SHA-256 manifest. It exposes aggregates and calculation rules. It does not redistribute the full fill archive, private correspondence, unpublished identity mappings or a complete raw-data rerun environment.

The calculation script rebuilds the headline rates, excess wins and observed-to-simulated ratio from the aggregate rows. Original research outputs were completed 5-6 October; the present note checks and packages them on 8 October. The input-file hashes in the manifest allow an editor to distinguish this version from later revisions. A changed figure should receive a new version and an explicit change note.

References

- [1] Boka Qin and Rui Yang. TimeSeventeen/Polymarket-v1 dataset and data definitions (CC BY 4.0). <https://huggingface.co/datasets/TimeSeventeen/Polymarket-v1>
- [2] PCAOB. AuditorSearch and Form AP: public annual-auditor records. <https://pcaobus.org/resources/auditorsearch>
- [3] EventWaves. KPMG-Audited Companies are Trading Strangely on Polymarket. Source of the initial nine and supplementary four published addresses. <https://eventwaves.substack.com/p/kpmg-audited-companies-are-trading>
- [4] Polymarket. Gamma API documentation: event metadata. <https://docs.polymarket.com/api-reference/events/list-events>
- [5] Vultax Research. Earlier long-shot and tight-screen study. Read with the correction in this note. <https://vultax.com/research/polymarket-longshot-insider-screen-base-rate>

Citation

Vultax Research (2026). *Where unusual Polymarket wins concentrate: An exploratory comparison of published earnings accounts, market-implied expectations and broader wallet screens*. Research note, version 1.0, 8 October 2026. vultax.com/research/polymarket-excess-wins-domain-concentration

Attribution and status. Underlying fill archive: Boka Qin and Rui Yang, TimeSeventeen/Polymarket-v1, CC BY 4.0. Analysis and presentation: Vultax Research. Publication does not imply peer review, endorsement by the source authors or the people whose public work is cited, or permission to infer personal identities from wallet statistics.