Where unusual Polymarket wins concentrate

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

Christopher Geneth Kurg

Vultax Research, OmniOS OÜ, Tallinn, Estonia. research@vultax.com. ORCID: https://orcid.org/0009-0007-4902-8847

Version 1.0 published 8 October 2026 at https://vultax.com/research/polymarket-excess-wins-domain-concentration. This version 1.1: 27 October 2026. Historical observation window: 1 January 2025 to 28 April 2026.

DOI: https://doi.org/10.5281/zenodo.23696794

Suggested citation: Kurg, C. G. (2026). Where unusual Polymarket wins concentrate: An exploratory comparison of published earnings accounts, market-implied expectations and broader wallet screens. Vultax Research Note 2026-025, version 1.1. Vultax Research, Tallinn. https://doi.org/10.5281/zenodo.23696794

Licence: Creative Commons Attribution 4.0 International (CC BY 4.0).

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. A check added in this version shuffles auditor labels only among companies of the same SIC sector group and public-float size group; the observed excess on KPMG-labelled cycles is reached by none of 20,000 such shuffles. The note sets out the units, failed screens, the corrected concentration calculation and the checks still needed before any stronger interpretation.

Keywords: prediction markets; Polymarket; insider trading screens; earnings markets; calibration benchmarks; wallet concentration. **JEL classification:** G14, K42, D84.

Status. Exploratory research; not peer reviewed. The figures are checked against saved outputs of analyses completed on 5 and 6 October 2026. This publication is not a new raw-data rerun. Version

1.1 names the author and adds a sector-and-size check (section 6), acknowledgements, references, declarations and a citation block; no figure, table or result of version 1.0 changed (section 7).

Table 1. Three results to keep separate

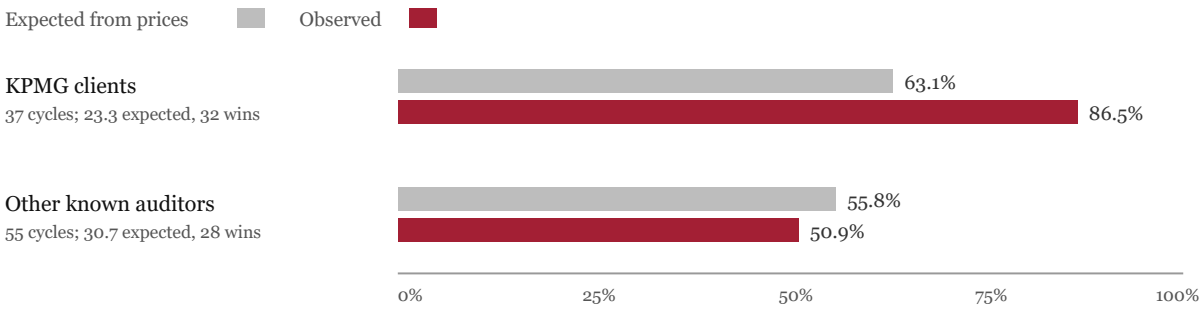
Question	Result	Interpretation
Where did a published group outperform its entry prices?	32 of 37 wins on KPMG clients; 23.3 expected. Other known auditors: 28 of 55; 30.7 expected.	A retrospective comparison within thirteen selected wallet addresses.
How often did a broader subject screen qualify wallets?	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.

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]. They were matched 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 and 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 in the companion package.

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. No confirmatory significance claim is attached 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. Section 6 answers the first half of that question for the published group. The later-period test remains proposed work, not a completed validation or a promised outcome.

3 The broader subject screen

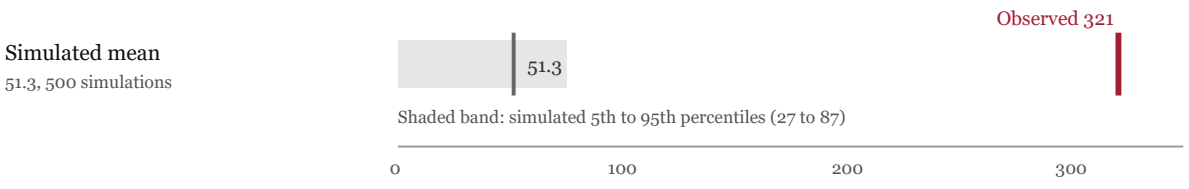
The full archive has 714,691,643 fills, 834,732 markets and 2,404,822 wallet addresses [1]. After excluding sports, esports and fixed-clock crypto, the subject search retains 135,284 markets (122,509 resolved within the tape) and 118,158,616 fills. Wallet addresses, fills, bets and people are different 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.

Table 2. Reference eligibility, before testing success

Condition	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. Wallets passing the same reference screen: simulated and observed



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. Source: population-sensitivity.csv and headline-data.csv in the companion package.

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 to 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.

Table 3. Eight corrected topic and 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. Source: strict-screen-comparisons.csv in the companion package.

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 studied case episodes include the Venezuela market of January 2026 and the Iran-strike markets of February 2026, whose trading was first set out publicly by Sethi [6, 7], and the published KPMG-client accounts [3].

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. No validated single screen that captures every studied case has been found.

Before a stronger claim

Match companies within sector and size groups (section 6 does this for the published group); 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. All steps after the first remain outstanding.

Companion material and audit trail

The companion package 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 on 5 and 6 October; version 1.0 checked and packaged them on 8 October. The input-file hashes in the manifest allow an editor to distinguish this version from later revisions. A changed figure

receives a new version and an explicit change note. Version 1.1 adds, beside the unchanged package, two aggregate tables and the script of the check in section 6.

6 A sector-and-size check

This section was added in version 1.1. It answers the first half of the question in section 2 for the published group: the company-label shuffle of section 2 preserved each audit firm's number of clients but not the sector or size mix of those clients, so a group that happened to trade companies in sectors or size classes where one firm audits many could show a concentration for that reason alone.

Method

Each of the 423 companies with a resolved main-stratum earnings market received its Standard Industrial Classification code from the SEC's company submissions and its latest public float reported on or before 28 April 2026 from the SEC's XBRL frames (dei:EntityPublicFloat) [8, 9]; 419 companies have a code and 402 a public float. Codes were collapsed into fifteen sector groups (for example chemicals and pharmaceuticals, electronics, finance and insurance, business services and software, wholesale and retail). Size groups are terciles of public float, with cuts at $5.7 billion and $41.4 billion; a median split is also reported. Companies without a code or a float form their own group.

The statistic is the one in Figure 1: the pooled thirteen's initial-direction wins on KPMG-labelled cycles minus the wins expected from entry prices, 32 − 23.331 = 8.669 over 37 cycles belonging to 27 companies. The auditor labels of all 423 companies were then shuffled 20,000 times (seed 20261027), each time only among companies of the same stratum, so that every stratum keeps its own count of each firm's clients. Entries, prices, timing and the once-per-cycle rule stay fixed. The labels derived from the saved auditor roster match the saved group records in all 92 cycles. Three quantities were compared with the observed values: the excess wins on KPMG-labelled cycles, the exact Poisson-binomial tail of that record (observed 4.43×10^{-5}), and the gap between the inside and outside excess rates (observed 28.3 percentage points).

Table 4. Company-label shuffles of the pooled thirteen's record, within sector and size groups

Labels shuffled within	Strata	KPMG cycles, mean	Excess wins: mean / 95th pct / largest	At or above 8.669 wins	Tail at or below observed	Gap at or above 28.3 points
None (as in section 2)	1	16.1	1.06 / 3.87 / 8.11	0	4	183
Sector group	15	17.6	1.39 / 4.17 / 7.73	0	1	140
Size tercile	4	16.1	1.30 / 4.06 / 7.52	0	4	238
SIC division, size tercile	24	16.7	1.36 / 3.96 / 8.69	1	4	146
Sector group, size median	35	17.5	1.62 / 4.39 / 8.31	0	2	211
Sector group, size tercile	48	17.3	1.67 / 4.30 / 7.98	0	1	177

20,000 shuffles per row; the last three columns count shuffles out of 20,000. Strata include the groups for companies without a code or a float. A stratum of one company keeps its label: there is one such stratum in row 4, three in row 5 and eight in row 6. Source: sector-size-check.csv and sector-size-composition.csv, published with this note.

Result and limits

Stratification moves the benchmark a little, not the finding. Under the null the mean excess on KPMG-labelled cycles rises from 1.06 wins without stratification to 1.67 wins within sector and size groups, because the group traded in sectors where KPMG audits a larger share of companies; the group's 92 cycles fall in eleven of the fifteen sector groups, with wholesale and retail (14 KPMG-labelled cycles against 11 others) and business services and software (8 against 12) the largest. The observed excess of 8.669 wins is reached by none of 20,000 shuffles in five of the six specifications and by one shuffle in the sixth. The inside-minus-outside gap is matched in 0.7% to 1.2% of shuffles.

The limits of section 2 are unchanged. The thirteen addresses were selected after their activity was reported, so this is a check on one alternative explanation, not a test of a prediction. The shuffle preserves sector and size but not liquidity, the number of contracts per company, event difficulty or each company's reporting calendar. The sector code and the public float are the SEC's current fields, not values frozen at each trade date. The six specifications are not adjusted for being six. Thirty-seven cycles from 27 companies is a small record, and the companies and reporting periods are not independent.

What a later-period test still needs

The second half of the question in section 2, a rule frozen in advance and scored on a later period, has not been run. It needs fills after 28 April 2026, where the archive used here ends, for at least one full reporting season; the company and auditor roster carried forward with the same Form AP rule; and the rule written down before the first flagged position settles. A forward rule of that kind for new accounts on one audit firm's clients was drafted on 6 October 2026 and has not been frozen or run. Until it has, the concentration in Figure 1 is a retrospective observation about thirteen published addresses.

7 Acknowledgements and version note

The author thanks the correspondents whose questions shaped the tight 48-hour screen in section 4 and the comparison of the studied cases, and the readers to whom version 1.0 was sent. None of them reviewed this version, and all errors are the author's. The fill archive is the work of Boka Qin and Rui Yang [1].

Version 1.1 (27 October 2026) names the author; adds section 6, a check proposed in version 1.0 and computed on 27 October 2026 from the saved group records and public SEC fields; adds this section, references [6] to [9], the declarations in section 8 and the citation block in section 9; and carries the series number. No figure, table or result of

version 1.0 of 8 October 2026 changed; the companion package and its hashes are unchanged.

8 Declarations

Competing interests. The author is a co-founder of OmniOS OÜ, Tallinn, Estonia, the company that operates Vultax Research and publishes this series. Vultax Research publishes its research free of charge, offers a free research terminal on prediction markets, and may provide paid data and research services to businesses; no such service was involved in the work reported here. The author holds no positions in the markets studied and has received no payment, data or other consideration from Polymarket, Kalshi or any other venue or data vendor beyond their public interfaces and published datasets, which are cited in the references.

Funding. This work received no external funding.

Use of AI. Generative AI tools (Anthropic's Claude) assisted with drafting, data processing and editing of the study and of this note; the author directed the work, checked every figure and source, and takes full responsibility for the content.

Data and materials. The companion package (aggregate tables, methodology, calculation script and SHA-256 manifest) is published with the study and deposited with the Zenodo record, together with the two aggregate tables and the script of section 6. Sources keep their own terms; the fill archive is CC BY 4.0 [1].

Licence. This note is licensed under the Creative Commons Attribution 4.0 International licence (CC BY 4.0).

9 How to cite

Kurg, C. G. (2026). Where unusual Polymarket wins concentrate: An exploratory comparison of published earnings accounts, market-implied expectations and broader wallet screens. Vultax Research Note 2026-025, version 1.1. Vultax Research, Tallinn. https://doi.org/10.5281/zenodo.23696794

```
@techreport{kurg2026_025,
  author      = {Kurg, Christopher Geneth},
  title       = {Where unusual Polymarket wins concentrate: An exploratory comparison of
published earnings accounts, market-implied expectations and broader wallet screens},
  institution = {Vultax Research},
  type        = {Vultax Research Note},
  number      = {2026-025},
  address     = {Tallinn, Estonia},
  year        = {2026},
  month       = oct,
  note        = {Version 1.1, 27 October 2026; version 1.0 published 8 October 2026. CC BY 4.0},
  doi         = {10.5281/zenodo.23696794},
  url         = {https://vultax.com/research/polymarket-excess-wins-domain-concentration}
}
```

References

1. Qin, B. and Yang, R. 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 (2026). Polymarket insider trading: the charged cases, and what a screen catches. Earlier long-shot and tight-screen study, read with the correction in this note. https://vultax.com/research/polymarket-longshot-insider-screen-base-rate. Zenodo: https://doi.org/10.5281/zenodo.23694013
6. Sethi, R. (2026a). The Venezuela Windfall. Imperfect Information, 8 January 2026. https://rajivsethi.substack.com/p/the-venezuela-windfall
7. Sethi, R. (2026b). Trading on Violence. Imperfect Information, 2 March 2026. https://rajivsethi.substack.com/p/trading-on-violence
8. U.S. Securities and Exchange Commission. EDGAR application programming interfaces: company submissions and XBRL frames. https://www.sec.gov/search-filings/edgar-application-programming-interfaces
9. U.S. Securities and Exchange Commission. Standard Industrial Classification (SIC) code list. https://www.sec.gov/search-filings/standard-industrial-classification-sic-code-list

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.