

Actionable Signals in Prediction Markets: Evaluating Public and Private Information through Systematic Trading

Oriol Saguillo
oriol.saguillo@imdea.org
IMDEA Networks & Universidad
Carlos III de Madrid
Madrid, Spain

Fabio Giacomelli
fabio.giacomelli@uniroma2.it
University of Tor Vergata
Rome, Italy

Lucianna Kiffer
lucianna.kiffer@networks.imdea.org
IMDEA Networks
Madrid, Spain

Abstract

Prediction markets aggregate the private beliefs of many participants into publicly observable probability estimates. Beyond prices alone, the trading behavior they generate may reveal *public* and *private* signals that are actionable for automated trading agents. We study whether these observable market signals reliably identify profitable trading opportunities. Specifically, we evaluate two complementary classes of trading signals: *Risk Arbitrage*, which exploits periods where publicly available information has effectively determined an outcome before prices fully converge, and *Information Asymmetry*, which exploits large positions accumulated while market prices still reflect substantial uncertainty.

Using the complete history of matched trades from Polymarket, spanning 26,964 markets between January 2024 and January 2026, we evaluate the reliability of these two classes of signals. We find that they exhibit markedly different levels of reliability. *Risk Arbitrage* identifies substantial opportunities around effective resolution but suffers significant losses from false positives, yielding a net loss of \$23.4M despite \$15.3M in gross profit. *Information Asymmetry* is considerably more concentrated, generating \$139.3M in gross profit but only \$7.8M net, with activity overwhelmingly driven by the 2024 U.S. presidential election. Across both regimes, profits concentrate among a small number of accounts, while position size reflects capital commitment more strongly than predictive skill.

CCS Concepts

• Security and privacy → Economics of security and privacy.

Keywords

Prediction Markets, Algorithmic Trading, Market Efficiency

1 Introduction

Prediction markets aggregate beliefs from diverse participants into probability estimates for future events, producing well-calibrated forecasts across a wide range of domains [20, 21]. The market’s price for an outcome is its *consensus signal*: a public, continuously

updated estimate of an outcome’s probability, formed by aggregating the private beliefs of many traders. As platforms such as Polymarket [15] have grown to handle billions of dollars in trading volume, a central question is not only how faithfully prices reflect underlying probabilities, but also whether the signals emitted by the market remain informative enough for automated or strategic trading agents to exploit.

Participants in these markets do not interact on equal footing. Traders differ in their access to information and in their ability to act upon it. Some opportunities arise from *public information* that has not yet been fully incorporated into market prices, while others arise from *private information* held by only a subset of participants. These two information regimes give rise to distinct observable market signals. A *public signal* emerges when an outcome has effectively been determined by publicly available information, yet prices continue to imply residual uncertainty. A *private signal* emerges when a trader accumulates a substantial position while prices still reflect considerable uncertainty, suggesting access to information not yet incorporated into market prices. Because small trades are often noisy, we focus on signals backed by substantial capital commitment, where larger positions provide stronger evidence of informational advantage than routine speculative trading.

Polymarket, a prediction market built on the Polygon blockchain, has emerged as a prominent platform for forecasting future events, driven in part by unprecedented participation during the 2024 U.S. presidential election, during which it processed over \$3.7 billion in trading volume. Because every matched trade is publicly recorded on the blockchain, the complete trading history is available to any observer, allowing the reconstruction of individual market positions over time. This level of transparency provides a unique opportunity to study large-scale trading behavior and evaluate how public and private signals manifest in prediction markets.

In this paper, we evaluate whether the public and private signals described above are actionable for systematic trading. Our premise is that if a signal contains actionable information, then an automated strategy acting on that signal should consistently generate profits. Profitability therefore serves as a measure of signal reliability rather than trader skill. We study this through two complementary trading regimes. *Risk Arbitrage* evaluates whether observable market behavior alone can identify that an outcome has effectively resolved before prices fully converge, while *Information Asymmetry* evaluates whether unusually large positions accumulated under uncertainty provide an actionable signal of informational advantage before it is incorporated into market prices.

This paper makes the following contributions:

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
Conference acronym 'XX, Woodstock, NY

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-XXXX-X/2018/06
<https://doi.org/XXXXXXXXXXXXXXX>

- **A framework for evaluating trading signals:** We introduce a framework for identifying public and private signals from executed trades and evaluating their actionability by accounting for both successful and unsuccessful detections.
- **Large-scale empirical evaluation:** Across 26,964 Polymarket markets spanning January 2024–January 2026, we characterize the reliability of public and private signals, showing that they differ substantially in prevalence, profitability, and concentration.

1.1 Related Work

Prediction market prices serve as accurate probabilistic forecasts, often outperforming expert- or survey-based methods [2, 20]. Under standard assumptions, prices can be read as wealth-weighted average beliefs over outcomes [21], forming a public consensus signal that aggregates the dispersed private beliefs of many traders [11], though this is complicated by heterogeneous risk preferences and budget constraints [10]. The accuracy of this consensus is not uniform, degrading where information is thin or where frictions delay its incorporation into prices. We treat the profitability of trading against the near-resolution consensus as a *measurement* of its residual calibration error, rather than assuming calibration a priori.

Classical microstructure theory predicts informed agents reveal private signals through strategic trading, incorporating information into prices [4, 7]; in betting markets, Shin formalized how prices adjust to protect against informed bettors [17]. Position size is itself informative: the Kelly criterion prescribes bet sizes proportional to an agent's edge [6, 9], and in Kelly-populated markets prices converge to accurate belief aggregates [3]. Whether large Polymarket positions reflect such an informational edge, rather than merely scaled-up conviction, remains untested at scale; the question our second regime, Information Asymmetry, addresses directly.

Our first regime is a prediction-market analogue of *risk (merger) arbitrage* in traditional finance, where an announced but not-yet-closed deal leaves the target trading at a discount to its terminal value, and arbitrageurs earn the spread by bearing residual completion risk and the time cost of capital [12]. The same structure arises between a market's *effective* resolution and its *formal* settlement: once an outcome is publicly decided, the winning contract may still trade below \$1, and traders capture the remaining spread by holding to settlement. As in the merger case, the edge is structural rather than informational, available to anyone observing the public consensus, so the returns it leaves on the table measure how completely the consensus has converged. Decentralized prediction markets such as Polymarket make this measurable at scale through the transparency of their on-chain records [16, 18], yet systematic analysis of trader behavior in these settings remains limited. In particular, the extent to which large traders' profitability reveals miscalibration of the public consensus versus the pricing of private signals has not been studied at scale.

2 Background

Prediction market platforms such as Polymarket [15], Kalshi [5], and Limitless [8] enable participants to trade shares tied to the outcomes of future events. We study Polymarket, a limit-order-book prediction market in which each binary market, referred to

as a *condition*, corresponds to a single yes/no outcome (e.g., "Will Candidate X win the election?"). Every condition has two tradable outcome shares, YES and NO, of which exactly one settles to \$1 while the other becomes worthless after the condition resolves. Consequently, outcome share prices can be interpreted as market-implied probabilities of the corresponding outcomes, and participants trade accordingly: a trader who believes an outcome is more likely than the price implies buys the corresponding share, while one who believes it less likely sells.

Orders are submitted to a centralized server and matched according to price and quantity. Once executed, trades are recorded on the Polygon blockchain, providing a complete and publicly verifiable history of matched trades. In addition to peer-to-peer trade executions, positions may also be created and redeemed through complementary token minting and merging operations, against \$1 of collateral, which determines the total supply of outstanding shares for a condition; these operations are likewise recorded on-chain. Together, these events allow the complete reconstruction of trader positions over time.

Some real-world events admit multiple mutually exclusive outcomes, such as elections or sports tournaments. Polymarket represents each possible outcome as a separate binary condition while grouping related conditions under a common event identifier. In our analysis, each condition is treated independently, allowing us to study price convergence, trader behavior, and signal reliability at the level of individual binary markets.

3 Public and Private Market Signals

Prediction markets expose two observable sources of information about future outcomes. The *public consensus signal* is reflected in market prices, which aggregate the information available to all participants. The *private signal* is reflected in the positions accumulated by individual traders before their information becomes incorporated into prices. Because the underlying information itself is unobservable, we evaluate the actionability of these signals through two complementary trading behaviors. *Risk Arbitrage* captures public signals by identifying late-stage purchases after an outcome has effectively been determined but before formal settlement. *Information Asymmetry* captures private signals by identifying unusually large positions accumulated while market prices still reflect substantial uncertainty. In both cases, realized profitability serves as a measure of signal reliability: if a signal is actionable, then a systematic strategy based solely on that signal should generate consistent profits.

3.1 Risk Arbitrage

Risk Arbitrage captures a public consensus signal that arises when it becomes publicly known that a market will resolve to either YES or NO, while formal resolution takes additional time. This may occur due to the structure of the market (for example, when resolution depends on a public announcement or certification, such as the final counting of mail-in votes in an election), or due to discrepancies or delays in oracle resolution. During this interval, traders may acquire the asset corresponding to the eventual outcome at prices below its terminal payoff and hold it until settlement, earning a small but predictable margin.

We characterize this signal using the following assumption and classification heuristics:

ASSUMPTION 1 (LATE-STAGE MARKET CONVERGENCE). *Consider a binary prediction market with creation time t_0 and resolution time t_n . There may exist a time $t^\dagger \in (t_0, t_n)$ such that, conditional on publicly available information at $t^\dagger$, the probability of the eventual outcome is arbitrarily close to one. During the interval $(t^\dagger, t_n)$, the asset corresponding to the eventual outcome has an expected terminal payoff close to one dollar, while its market price remains below one due to the remaining time until resolution. As a result, bids placed in this interval offer low downside risk and a small expected return equal to the remaining price–payoff gap.*

We take the executed **buy** order as the basic unit of analysis. A buy is considered a *Risk Arbitrage buy* if it is executed during the late-stage interval described in Assumption 1, when remaining uncertainty is low and the expected payoff is close to its terminal value. Because the underlying probability of the outcome is not directly observable, we characterize this behavior using three observable dimensions that reflect the payoff profile of Risk Arbitrage:

- **Price level (marginal return):** buys placed at prices close to one have limited remaining upside, directly capturing the small expected return of the strategy.
- **Timing relative to resolution:** Risk Arbitrage buys occur late in the market, so that profits primarily reflect waiting for settlement rather than early speculation.
- **Late-stage price stability (low risk):** Risk Arbitrage should occur only after the market has converged toward the eventual outcome and prices remain persistently elevated through resolution. Substantial downward repricing during this interval would indicate that the apparent certainty of the outcome had not yet been established.

Fix a price threshold $\tau \in (0, 1)$, which controls how close prices must be to the terminal payoff. For each market, we model this behavior by first identifying $x^* \in \{\text{YES}, \text{NO}\}$ as the realized winning outcome and t_n as the market resolution time. We define $t^\dagger = \max\{t < t_n \mid p_{x^*}(t) < \tau\}$ as the final moment the winning outcome’s price traded below a threshold τ . This timestamp establishes the *late-stage window* $(t^\dagger, t_n)$. Within this framework, a buy is classified as a *Risk Arbitrage buy* if it satisfies three concurrent conditions: (1) it involves the purchase of the winning outcome x^* , (2) it is executed strictly within the late-stage window, and (3) it is held through market resolution rather than offset by a prior sale.

Intuitively, Risk Arbitrage buys represent purchases made after the market has effectively converged to the realized outcome but before formal settlement. Their profitability therefore reflects the reliability of this late-stage public consensus signal.

3.2 Information Asymmetry

Information Asymmetry captures a private signal that arises when some traders possess information about the eventual outcome of an event that is not yet reflected in market prices. Such behavior manifests as unusually large accumulation of a position while market prices still reflect substantial uncertainty, followed by substantial realized profit as prices adjust. Recent controversies surrounding markets on events such as military strikes in Iran [13] and political

instability in Venezuela [19] have raised concerns that some participants may trade on information before it becomes reflected in public market prices.

ASSUMPTION 2 (ASYMMETRIC INFORMATION ACCUMULATION). *Consider a binary prediction market with outcome $x \in \{\text{YES}, \text{NO}\}$. There may exist periods during which one or more traders possess information indicative of the realized outcome x , while the market price $p_x(t)$ continues to reflect substantial uncertainty. During such periods, informed traders may accumulate unusually large positions in x prior to subsequent market repricing. As information becomes public and prices adjust toward their terminal value, these positions realize substantial profit.*

Because asymmetric information is not directly observable, we characterize this signal using the following observable dimensions:

- **Accumulation under uncertainty:** the position is accumulated while the market price of the realized outcome still reflects substantial uncertainty.
- **Large cumulative positions:** traders acting on asymmetric information are expected to accumulate unusually large positions relative to the overall market supply.
- **Substantial realized profit:** positions consistent with asymmetric information are expected to realize substantial profit following subsequent market repricing.

Information-Asymmetric Positions.

We take the trader–market position¹ as the unit of analysis and identify positions that exhibit this private signal.

For each market:

- (1) Let $x^* \in \{\text{YES}, \text{NO}\}$ denote the realized winning outcome. Once cumulative traded supply exceeds a minimum threshold $S_{\min}$, we repeatedly examine the distribution of trader balances in x^* over time.
- (2) A trader position is flagged as a *position outlier* if its balance lies in the upper tail of the balance distribution at any observation point, as determined by threshold β . This captures unusually large accumulation of the realized winning outcome relative to the rest of the market.
- (3) For each position outlier, we compute the volume-weighted average purchase price (VWAP) of the accumulated position. We require that this VWAP remain below a maximum threshold $\theta_{\max}$, ensuring that accumulation occurred while the market still reflected substantial uncertainty regarding the outcome.
- (4) We then compute the realized profit associated with the position through subsequent liquidation or eventual market resolution.
- (5) Finally, we require positions to remain open for at least a minimum holding period $\Delta_{\min}$.

A trader–market position satisfying the above conditions is classified as *information-asymmetric*.

We set $S_{\min} = 10,000$ tokens as a minimum supply threshold, ensuring the balance distribution used for outlier detection contains sufficient participation; β is then defined via the Tukey IQR

¹Defined as the cumulative balance in a given outcome held by a trader at a given point in time.

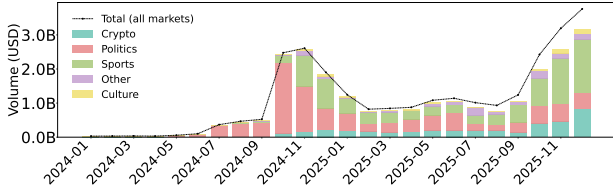


Figure 1: Polymarket monthly trading volume during the study period compared with the subset of conditions retained after dataset filtering. The retained dataset accounts for the large majority of economically meaningful platform activity.

method,² flagging accounts holding disproportionately large positions relative to the rest of the market. We require accumulation to occur at a VWAP below $\theta_{\max} = 0.70$, ensuring positions are established while the market still reflects substantial uncertainty. Finally, we impose a minimum holding duration of $\Delta_{\min} = 2$ days to exclude short-lived trading activity and focus on sustained position accumulation. These bounds are intentionally conservative, casting a wide initial net of candidates; the empirical distributions in Section 6.1 motivate tighter data-driven thresholds that refine this raw set to a high-conviction sample.

4 Data Collection

We analyze all prediction markets created and resolved on Polymarket between January 1, 2024 and January 1, 2026. Our dataset combines information from two sources: the Polymarket GAMMA API [14] and on-chain event data recorded on Polygon. From the GAMMA API, we collect condition metadata information including token identifiers, creation and resolution timestamps, and market categories. We reconstruct trading activity from on-chain events following the methodology of [16] by querying Polymarket’s exchange contracts via Alchemy [1]. Specifically, we reconstruct executed buy and sell trades, token minting and redemption events (split/merge operations), and token transfers, mapping all activity to the corresponding condition.

Across the study period, the raw dataset contains 243,155 conditions, 280,348,267 reconstructed trade executions,³ 1,745,600 unique trading accounts, and approximately \$28.0B in total trading volume. Our methodology focuses on identifying large position accumulation and late-stage trading activity. We therefore apply a sequence of filters designed to retain conditions with sufficient liquidity and participant activity to support these behaviors.

We first exclude conditions without a definitive resolution (0.1% of raw conditions)⁴ and conditions with fewer than 100 executed trades ($\approx 55\%$ of raw conditions), leaving 108,996 conditions. We then apply three sequential economic filters, each a conservative

²A position is flagged as a supply outlier if its balance exceeds $Q_3 + 1.5 \cdot \text{IQR}$ of the cross-sectional balance distribution at any observation point, where Q_3 is the third quartile and $\text{IQR} = Q_3 - Q_1$.

³Trade executions are reconstructed from emitted exchange events and include matched buy and sell executions associated with a condition. Split and merge operations are recorded separately as token minting and redemption events.

⁴Specifically, we exclude conditions in which both YES and NO resolved to 0.5, such as ambiguous markets including “Is Strawberry Gummy a Man or a Woman?”. Since these conditions lack a unique winning outcome, they cannot be used to evaluate directional profitability or token convergence dynamics.

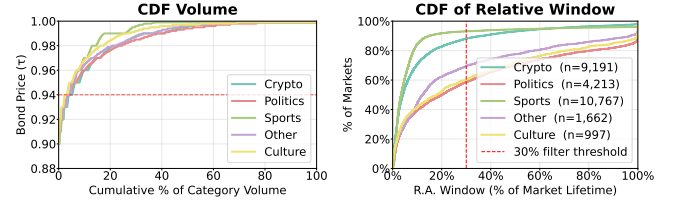


Figure 2: Threshold and window-length selection. Left: fraction of winning-outcome purchase volume retained as the candidate threshold τ increases, by topic; at $\tau^* = 0.94$, $\approx 96\%$ is retained. Right: Distribution of late-stage window lengths induced by $\tau^* = 0.94$, expressed as a fraction of the total market lifetime; markets exceeding 30% are excluded.

lower bound chosen to exclude the lower tail of minimally active conditions while retaining the economically active core: a minimum peak supply of \$20,000 (the maximum collateral simultaneously locked in a condition over its lifetime), at least \$100,000 in cumulative trading volume, and at least 100 unique trading accounts. The resulting analysis dataset contains 26,964 conditions and, despite excluding the large majority of raw conditions by count, preserves most of the platform’s economically meaningful trading volume throughout the study period (Figure 1).

To characterize behavioral patterns across market types, we assign each condition a primary topic derived from its Polymarket tag labels. After excluding generic structural tags (Recurring, Hide From New, and All) that carry no semantic information, we map the remaining tags into five categories: Sports (10,831 conditions, \$8.73B volume), Crypto (9,223, \$3.81B), Politics (4,235, \$9.87B), Other (1,673, \$1.45B), and Culture (1,002, \$0.71B). Sports and Crypto dominate by condition count due to recurring game and price markets, while Politics accounts for the largest trading volume.

5 Risk Arbitrage

Risk Arbitrage (RA) evaluates the reliability of a public market signal indicating that an outcome has become effectively certain. Traders with external knowledge of the underlying event can identify these opportunities directly, but an automated trading agent observing only market behavior must instead infer such opportunities from late-stage price dynamics. We therefore ask whether observable market behavior provides a sufficiently reliable proxy for effective resolution to support systematic Risk Arbitrage, or whether apparent convergence produces too many false signals to be profitable in aggregate.

5.1 Identifying the Late-Stage Regime

The primary parameter governing the Risk Arbitrage heuristic is the price threshold τ , which determines when observable market behavior indicates that a market has entered the late-stage regime. Because effective resolution is not directly observable from prices alone, τ defines the point at which an automated agent would begin treating the public consensus signal as actionable. To select an operational value, we first define a broad reference set of late-stage activity using a permissive threshold of $\tau = 0.90$, treating

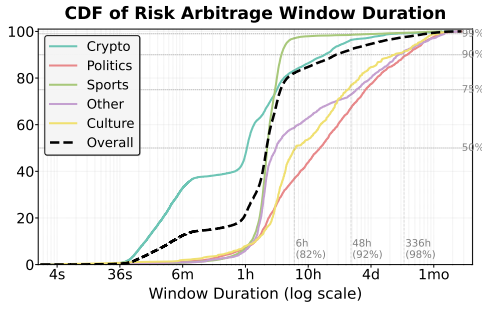


Figure 3: CDF of Risk Arbitrage window duration by market topic. While Sports and Crypto markets typically converge shortly before settlement, Politics, Culture, and Other often remain in the late-stage regime for longer periods.

winning-outcome purchases above this level as candidate end-of-market activity. We then examine the fraction of this purchase volume retained as τ varies across market categories (Figure 2, left). We adopt $\tau^* = 0.94$, representing a high-confidence definition of the late-stage regime while still retaining approximately 96% of candidate late-stage purchase volume across the major topic categories.

The late-stage windows induced by τ^* vary substantially across conditions. While most Sports and Crypto markets enter the late-stage window only shortly before resolution, Politics and Culture contain a heavier tail of conditions that remain near-certain for extended periods prior to settlement. These cases typically correspond to markets whose eventual outcome was effectively decided long before formal resolution.⁵ To exclude such structurally atypical conditions, we remove markets in which the late-stage window exceeds 30% of the total market lifetime (Figure 2, right), retaining approximately 83% of conditions with RA windows.

5.2 Prevalence of Risk Arbitrage Activity

Having identified the baseline threshold τ^* , we evaluate the prevalence and economic scale of Risk Arbitrage across market topics. Risk Arbitrage is widespread, appearing in 69% of filtered conditions and across all five topics. Yet within RA-active conditions it accounts for only a modest share of trading volume, ranging from 3.1% in Politics to 10.0% in Culture. Its share of *outstanding supply* is substantially larger, increasing from 3.1% of trading volume to 24.0% of supply in Politics, and from 4.5% to 15.5% in Sports. This disparity indicates that once the public signal is identified, positions are typically established and held through settlement rather than actively traded.

The duration over which the public signal remains actionable also varies systematically across market categories (Figure 3). Risk Arbitrage windows in Sports and Crypto are typically short, often lasting only hours before settlement, whereas Politics, Culture, and Other frequently exhibit substantially longer windows, extending from days to over a month. Longer windows increase the time over which capital remains committed, reducing its availability for

⁵For example, “Will Amazon be the largest company by market cap on December 31?”, for which the no outcome commanded near-certain probability throughout most of the market’s lifespan.

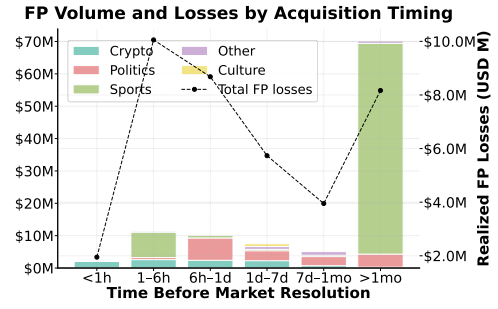


Figure 4: False-positive buy volume by time before resolution (stacked bars by category, left axis) and total realized losses (dashed line, right axis), restricted to buys at or above $\tau^* = 0.94$. Volume peaks for positions entered more than a month before resolution, but the *loss rate* is inverted: it falls from 88.6% for buys within an hour of resolution to 11.6% for buys entered over a month out.

other trading opportunities, and provide a greater interval during which apparent market convergence may reverse before formal resolution. These differences reflect the nature of the underlying events: rapidly resolving markets on one end, and slow-moving or foregone-conclusion markets on the other.

5.3 False Signals and Residual Risk

The reliability of the public signal depends not only on how often it identifies genuine late-stage convergence, but also on how often it produces false signals. We evaluate these failures through purchases of the losing outcome executed at or above the operational threshold $\tau^* = 0.94$ (*false positives*), measured on the same footing as RA’s winning-side profit so the two are directly comparable. Figure 4 shows both the volume and the realized loss of such false-positive purchases, bucketed by how long before resolution they were acquired. Volume is dominated by early entries: \$70.3M of the \$107.8M total false-positive volume (65.2%) was acquired more than a month before resolution. Yet these early positions carry the *lowest* loss rate of any bucket, just 11.6% (\$8.17M realized against \$70.3M deployed), suggesting that many early entrants successfully unwind as new information arrives. The pattern reverses sharply close to settlement: positions entered within a day of resolution (\$2.2M within the final hour, \$11.8M within six hours, \$10.7M within a day) realize 81–89% of their capital as loss, since a token still trading above τ^* within hours of settlement offers almost no window to exit before the outcome is finalized.

When convergence fails, participants are roughly evenly split between active unwinding and passive loss realization: 55.2% of false-positive buyer accounts sold their positions following acquisition, while 44.8% held their tokens to a zero payoff. Aggregate losses at τ^* total \$38.71M across 62,457 accounts (Table 1), with Sports accounting for the largest share (\$12.84M), followed by Politics (\$13.25M) and Crypto (\$8.51M). Despite this scale, failed convergence is concentrated in relatively few markets: the top 1% (31 of 3,150 markets) account for 72.6% of false-positive volume and

Topic	Win \$M	Win Accts	Loss \$M	Loss Accts	Net \$M
Politics	7.20	167,282	13.25	15,928	−6.05
Crypto	3.11	210,850	8.51	14,460	−5.40
Sports	2.91	46,114	12.84	28,430	−9.93
Other	1.61	63,073	2.67	5,150	−1.06
Culture	0.49	31,666	1.44	2,801	−0.95
TOTAL	15.32	358,825	38.71	62,457	−23.40

Table 1: Risk Arbitrage profitability by topic at $\tau^* = 0.94$. Despite substantial gross profits, every topic is net-negative once false-positive losses are included.

the top 5% (157 markets) for 85.0%. Even so, not every early, high-conviction entry escapes unscathed.⁶

Although false signals are concentrated in a relatively small number of markets, they exhibit a fundamentally asymmetric payoff. Genuine Risk Arbitrage opportunities offer only a small remaining spread to settlement, whereas a false signal can erase nearly the entire value of the position. Consequently, a small number of failed convergence events has the potential to outweigh profits accumulated across many successful trades, a hypothesis we quantify in the following subsection.

5.4 Profitability

Measured at its own operational threshold, Risk Arbitrage is a net loser. The winning side generates \$15.32M across 358,825 accounts (0.84% yield on \$1.82B deployed), but false-positive buys—losing-outcome purchases at the same $\tau^* = 0.94$ —produce \$38.71M in realized losses across 62,457 accounts, for a net of **−\$23.40M**. The loss is universal across topics (Table 1): Sports shows the largest shortfall (−\$9.93M, losses 4.4× profit), while Politics leads on gross profit (\$7.20M) yet is still net-negative (−\$6.05M). Despite being widespread, the public signal is therefore not sufficiently reliable to support systematic Risk Arbitrage at $\tau^* = 0.94$: the cost of false signals outweighs the gains from successful opportunities across every market topic.

Widely profitable, thinly so. Despite being net-unprofitable in aggregate, Risk Arbitrage is individually profitable for most participants: of 391,663 accounts with at least one RA position, 85.5% are net-positive. These results are reconciled by scale rather than sign. The median account nets just \$0.07, while a concentrated minority of large losses (mean net −\$59.74) dominates the aggregate. Both sides are top-heavy: the top 1% of winners capture 93.2% of profit, and the top 1% of losers capture 81.0% of losses. Failure is also rare at the market level. Of 25,559 RA-active markets, only 12.3% contain a false-positive position, but those failures are costly (average loss per false-positive market \$12,289 versus \$684 profit per true-positive market, an 18× asymmetry).

A further pattern suggests that successful and unsuccessful Risk Arbitrage are facets of the same underlying behavior rather than separate populations. A total of 29,619 accounts hold both winning and losing RA positions at τ^* , and 93.0% of false-positive markets also contain a genuine winning-side position in the *same* market,

⁶For example, in the market on PSG’s Champions League elimination, buyers entered at high prices well before resolution; the unexpected survival of PSG reversed what had appeared to be stable convergence toward elimination, and unlike the typical early entrant, these buyers did not exit in time.

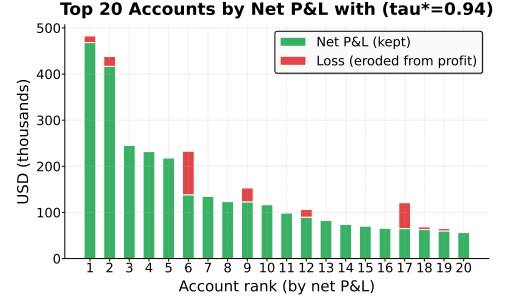


Figure 5: Top 20 Risk Arbitrage accounts ranked by net profit. Despite the strategy being net-unprofitable in aggregate, a small number of accounts consistently realize substantial positive returns while incurring comparatively small false-positive losses.

with both outcomes trading above τ^* before one ultimately resolved to zero. These are contested markets, not markets where the crowd was uniformly wrong.

Ranking accounts by net P&L rather than gross profit provides a clearer view of sustained performance. Although gains remain highly concentrated (the top 1% of net-positive accounts capture 93.1% of total profit), the highest earners are distinguished not by larger winning positions, but by remarkably limited false-positive exposure. The top net account earned approximately \$468K while incurring less than \$15K in losses, and several of the top ten accounts experienced no false-positive losses at all (Figure 5). This suggests that successful Risk Arbitrage depends not only on identifying the public signal, but also on distinguishing genuine convergence from false convergence.

6 Information Asymmetry

Information Asymmetry (IA) evaluates the reliability of a private signal inferred from market behavior. Rather than observing private information directly, an automated trading agent can only observe the accumulation of unusually large positions while the public consensus still reflects substantial uncertainty. We therefore ask whether such positions provide a reliable indication of information not yet incorporated into market prices, or whether they instead reflect concentrated, high-conviction speculation. Because Sports and Crypto markets are primarily driven by continuously unfolding public information rather than discrete event resolution, we restrict this analysis to Politics, Culture, and Other markets, where informational asymmetries are more plausibly reflected in large position accumulation.

6.1 Identifying Candidates

Applying the IA heuristic of Section 6, which identifies statistically large (Tukey–IQR) position outliers accumulated while the market remains uncertain ($\theta_{\max} = 0.70$), yields 131,317 candidate positions across 2,285 resolved conditions, tracking accumulation on the token that ultimately resolved to \$1. By construction, this subsection and Section 6.2 consider only the *winning* side of large-position accumulation and therefore describe the strongest candidate private signals. Section 6.3 applies the identical pipeline to losing positions,

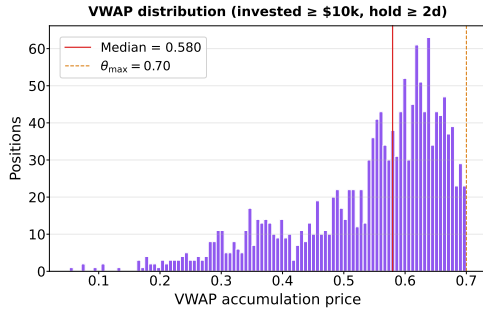


Figure 6: Distribution of volume-weighted accumulation prices for filtered IA positions. Large positions are typically established well before the market approaches the uncertainty threshold ($\theta_{\max} = 0.70$).

allowing us to evaluate whether similarly large accumulations are equally common on unsuccessful outcomes.

As a sanity check, we verify that flagged positions arise in markets with substantial participation and liquidity: candidate markets contain a median of 845 traders, and winning-token supply at detection has a median of 1,004,804 tokens, with 81.8% of positions detected in markets above 100,000 outstanding tokens, confirming that detection concentrates in liquid markets rather than thin ones.

We refine the candidate set to focus our analysis on economically meaningful positions. We require a minimum committed capital of \$10,000 (cost basis at detection), excluding positions too small to represent a practically actionable signal regardless of eventual outcome. We additionally impose a minimum holding duration of 48 hours, excluding positions liquidated shortly after accumulation while retaining longer-lived exposure under sustained uncertainty. Finally, Figure 6 shows that filtered positions are accumulated well below the $\theta_{\max} = 0.70$ ceiling: the median volume-weighted accumulation price is 0.580, consistent with entry under genuine uncertainty rather than late-stage convergence.

6.2 Scale and Persistence of IA Activity

The high-conviction sample comprises 1,522 positions across 326 conditions and 1,123 unique accounts. IA positions are predominantly medium- to long-term: the median holding duration is 11.8 days, with 66.6% held for more than one week and 22.9% for longer than 30 days. Election markets exhibit slightly longer holding durations than non-election markets (median of 12.4 vs. 9.4 days), consistent with their longer resolution timelines. [Examining the subsequent trading history of these positions (Section 6.3) shows that 63.5% are never sold on the open market before resolution, while the remaining 36.6% are partially or fully liquidated beforehand.] Together, these observations suggest that most detected positions represent sustained directional exposure under uncertainty rather than short-horizon trading.

Activity is concentrated in time and category. A disproportionate share occurs in 2024 US presidential election markets, which we analyze separately throughout, and participation is predominantly event-specific: the majority of accounts (934 of 1,123) appear in a

Sample	Outcome	Pos.	Accts	Total (\$M)	Top 10
Election	Win	1,062	853	125.2	47.8%
Election	Loss	1,456	1,155	116.0	22.9%
Other	Win	460	336	14.1	33.2%
Other	Loss	562	452	15.4	18.0%

Table 2: Winning and losing IA positions identified under identical detection criteria. Election markets account for the majority of both profits and losses.

single market, consistent with isolated informational advantages rather than platform-wide strategies.

6.3 False Signals and Signal Reliability

The reliability of the private signal depends not only on the existence of profitable large positions, but also on how often similarly large positions accumulate on the wrong outcome. Applying the IA heuristic to the losing side of each market under identical thresholds yields 2,018 large losing positions across 372 conditions and 1,551 accounts, exceeding the 1,522 winning positions identified in Section 6.1. Aggregate realized losses total \$131.5M against \$139.3M in profit (Table 2), leaving a net surplus of just \$7.8M. Median position size is similar in magnitude (\$20,606 loss vs. \$18,686 profit), although profits are substantially more concentrated than losses: the top 10 accounts capture 33.2–47.8% of profits compared with 18.0–22.9% of losses.

To evaluate the reliability of this private signal, we pool all detected positions from both sides (dropping the dollar filter, which biases comparison across price levels⁷) and compare each position’s VWAP entry price against its outcome. Across 203,936 pooled positions, win rate tracks price closely at the extremes but deviates in the middle: underperforming by 9–11 points in the 30–50% band, overperforming by 5–8 points in the 50–70% band (Figure 7). Predictability varies more by category than by whether a market is election-related: Culture, non-election Politics, and Other markets range from Brier 0.199–0.211, while 2024 US Election markets are only modestly less predictable (Brier = 0.222 vs. 0.215 pooled). Large accumulators generate the most profit and the most loss precisely where their entry prices are least reliable.

The near-parity between aggregate profit and loss is not driven by a uniform pattern across entry prices. Below VWAP 0.2, losing positions are individually larger than winners at the same price (median share 0.21 vs. 0.19 at [0.1, 0.2]), reflecting a small number of concentrated losses. Above roughly 0.2–0.3, the pattern reverses: winners hold the larger share, while losing positions are more numerous but individually smaller and more concentrated (Gini 0.81 vs. 0.74 at [0.6, 0.7]). The overall balance between profit and loss therefore emerges from distinct concentration dynamics across different price regimes rather than a single underlying mechanism.

A possible explanation for the observed profitability is that informed traders exit losing positions before the market fully reprices. We therefore compare exit behavior across winning and losing positions. While 36.6% of winning positions and 28.2% of losing positions are liquidated before resolution, conditional on exiting,

⁷Profit scales as $(1 - \text{VWAP}) \times \text{balance}$, loss as $\text{VWAP} \times \text{balance}$; a shared dollar floor admits cheap wins and expensive losses disproportionately.

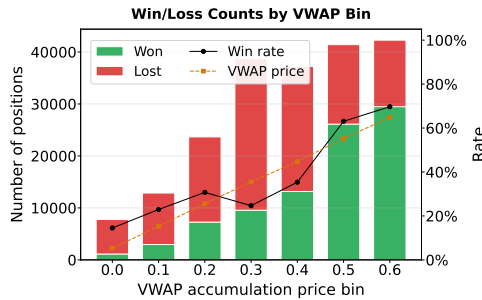


Figure 7: Number of winning (green) and losing (red) IA positions by VWAP entry-price bin, (left axis), with realized win rate and mean VWAP price overlaid (right axis; $n = 203,936$). The largest calibration errors occur in the middle price bins.

the timing is nearly identical (median 14.1d vs. 13.4d, $p = 0.23$). The difference therefore lies in *whether* positions are exited rather than *when*. Moreover, early exits from losing positions recover little value (median 5.8% of cost basis), suggesting that liquidation typically occurs only after prices have already collapsed.

6.4 Profitability and Concentration

The high-conviction winning sample comprises 1,522 positions across 326 conditions and 1,123 unique accounts. Aggregate realized profit totals \$139.3M (median \$18,686, mean \$91,521 per position) and is highly concentrated among a small set of participants. The top 10 accounts, fewer than 1% of traders, capture 43.0% of total profit and 43.6% of accumulated supply, while the top 50 account for 68.1% and 67.1%, respectively. Profit and supply shares track closely across quantiles, indicating that realized gains are driven primarily by position scale rather than selective timing.

A disproportionate share occurs in 2024 US presidential election markets, which account for \$125.2M of the \$139.3M total (Table 2), with substantially stronger concentration than elsewhere: the top 10 accounts capture 47.8% of election-market profit versus 33.2% in all other markets. Despite this concentration, participation remains predominantly event-specific, with 934 of 1,123 accounts appearing in only a single market. Yet profitability rises with repeat participation, from a median \$16.0K for single-market accounts to \$569.1K for accounts active in 11 or more markets — though that top bucket now contains only 3 accounts, so the latter figure should be read as illustrative rather than a stable estimate.

Taken together, these results indicate that the largest realized profits arise from a small number of highly concentrated positions, particularly in 2024 US presidential election markets. However, these gains coexist with a comparably large population of similarly sized losing positions and a market segment exhibiting the weakest calibration. While a small number of traders realize substantial profits, the aggregate evidence is more consistent with concentrated high-conviction betting than with a broadly reliable or systematically decodable private signal.

7 Conclusions

We examined the reliability of two observable trading signals in prediction markets. Across both settings, evaluating only successful instances substantially overstates the reliability of these signals. Accounting for equally observable false signals fundamentally changes the picture. Risk Arbitrage becomes net unprofitable once failed convergence events are incorporated, while the apparent profitability of Information Asymmetry is almost entirely offset by comparably large losing positions. Observable market behavior therefore contains informative signals, but neither late-stage price convergence nor large position accumulation alone reliably identifies actionable trading opportunities.

More broadly, our results highlight the importance of evaluating trading signals symmetrically. Analyses based solely on profitable positions risk overstating both the reliability of public consensus and the informational content of large trades. Incorporating false signals provides a more complete picture of market calibration and suggests that concentrated profits alone should not be interpreted as evidence of systematic predictive advantage.

Aspects of this analysis could be further refined in future work. The operational thresholds used to identify both strategies are heuristic choices designed to capture economically meaningful behavior, and alternative parameterizations may reveal different operating points. Our analysis also treats blockchain addresses as independent participants and does not account for cross-market hedging or multi-wallet ownership. Finally, while our dataset spans two years of Polymarket activity, a substantial fraction of the observed behavior is driven by the 2024 US presidential election. Extending this methodology to additional election cycles, prediction market platforms, and richer trader-level reconstructions would help assess the generality of these findings.

References

- [1] Alchemy Insights, Inc. 2026. Alchemy: The Web3 Development Platform. <https://www.alchemy.com/>. Accessed: 2026-05-27.
- [2] Joyce E Berg, Forrest D Nelson, and Thomas A Rietz. 2008. Prediction market accuracy in the long run. *International Journal of Forecasting* 24, 2 (2008), 285–300.
- [3] Alina Beygelzimer, John Langford, and David M. Pennock. 2012. Learning performance of prediction markets with Kelly bettors. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems - Volume 3 (Valencia, Spain) (AAMAS '12)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 1317–1318.
- [4] Sanford J. Grossman and Joseph E. Stiglitz. 1980. On the impossibility of informationally efficient markets. *American Economic Review* 70, 3 (1980), 393–408.
- [5] Kalshi. 2026. Kalshi: Prediction Market for Trading the Future. <https://kalshi.com/>. Accessed: 2026-05-28.
- [6] J. L. Kelly. 1956. A new interpretation of information rate. *The Bell System Technical Journal* 35, 4 (1956), 917–926. doi:10.1002/j.1538-7305.1956.tb03809.x
- [7] Albert S. Kyle. 1985. Continuous Auctions and Insider Trading. *Econometrica* 53, 6 (1985), 1315–1335. <http://www.jstor.org/stable/1913210>
- [8] Limitless Exchange. 2026. Limitless Exchange: Prediction Markets. <https://limitless.exchange/>. Accessed: 2026-05-28.
- [9] Leonard C MacLean, Edward O Thorp, and William T Ziemba. 2010. Good and bad properties of the Kelly criterion. *Risk* 20, 2 (2010), 1.
- [10] Charles F. Manski. 2006. Interpreting the predictions of prediction markets. *Economics Letters* 91, 3 (2006), 425–429. doi:10.1016/j.econlet.2006.01.004
- [11] Blake Martin, Mithun Chakraborty, and Sindhu Kutty. 2021. Timing is Money: The Impact of Arrival Order in Beta-Bernoulli Prediction Markets. In *Proceedings of the Second ACM International Conference on AI in Finance (ICAIF '21)*. Association for Computing Machinery, doi:10.1145/3490354.3494406
- [12] Mark Mitchell and Todd Pulvino. 2001. Characteristics of risk and return in risk arbitrage. *The Journal of Finance* 56, 6 (2001), 2135–2175.
- [13] Polymarket. 2025. Insider trading on Iran geopolitical strike markets. Tweet on X (@Polymarket). <https://x.com/Polymarket/status/1936628432542060890>. Accessed: 2026-05-28.

- [14] Polymarket. 2026. Polymarket CLOB API Reference: Introduction. <https://docs.polymarket.com/api-reference/introduction>. Accessed: 2026-05-18.
- [15] Polymarket. 2026. Polymarket: The World's Largest Prediction Market. <https://polymarket.com/>. Accessed: 2026-05-28.
- [16] Oriol Saguillo, Vahid Ghafouri, Lucianna Kiffer, and Guillermo Suarez-Tangil. 2025. Unravelling the Probabilistic Forest: Arbitrage in Prediction Markets. In *7th Conference on Advances in Financial Technologies (AFT 2025)*. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 27–1.
- [17] Hyun Song Shin. 1991. Optimal betting odds against insider traders. *The Economic Journal* 101, 408 (1991), 1179–1185.
- [18] Allen Sirolly, Hongyao Ma, Yash Kanoria, and Rajiv Sethi. 2025. Network-Based Detection of Wash Trading. *Available at SSRN* (2025).
- [19] U.S. Department of Justice, Office of Public Affairs. 2026. U.S. Soldier Charged With Using Classified Information To Profit From Prediction Market Bets. Press Release No. 26-397. <https://www.justice.gov/opa/pr/us-soldier-charged-using-classified-information-profit-prediction-market-bets>. Accessed: 2026-05-28.
- [20] Justin Wolfers and Eric Zitzewitz. 2004. Prediction markets. *Journal of economic perspectives* 18, 2 (2004), 107–126.
- [21] Justin Wolfers and Eric Zitzewitz. 2006. Interpreting prediction market prices as probabilities.