Opening Wide, Moving Fast: Polymarket Spread Dynamics

Carl Jörnell[1*], Vicente Perez[1], and Oriol Saguillo[2]

[1] Lund University, School of Economics and Management, Lund, Sweden
{ca7170jo-s,vi5236pe-s}@student.lu.se
[2] IMDEA Networks Institute, Madrid, Spain
oriol.saguillo@networks.imdea.org

Abstract. Polymarket creates a new market for every event it lists, and each contract begins with an empty order book: no inherited orders, no assigned market maker, and no opening auction. Every contract therefore has a clear opening moment and an early period during which its trading conditions are still forming. Using roughly 160,000 contracts created between October 2025 and March 2026, we study what it costs to trade in a market's first hours of life and how that cost behaves across different kinds of events.

Over the first twenty-four hours of each market, the *genesis window*, we measure how wide spreads start and how quickly they narrow, how much it actually costs to trade immediately, and how much of that cost stays with the liquidity provider versus reflecting the price moving afterwards. We compare these patterns across eight event categories. Trading costs are strongly category-dependent from the very first trade: sports markets open cheaply with deep, balanced books, while other categories open much wider and take longer to settle, with most of the narrowing happening within the first hour. Centrally, we document a *dissociation*: ranking categories by how expensive they are to trade and by how much early trades move the price produces two orderings that do not agree. These are separate properties, not two views of the same thing.

Keywords: Prediction markets · Market microstructure · Polymarket

1 Introduction

Prediction markets have moved from academic research into venues that price billions of dollars in event contracts, and Polymarket is the largest of them. Unlike a conventional exchange, Polymarket creates a brand new market every time an event is listed. Each contract begins with an empty order book: no inherited orders, no assigned market maker, and no opening auction. Trading begins the moment one trader accepts a price another has posted. Every contract therefore has a clear opening moment, a first trade, and an early period during which

* C. Jörnell and V. Perez contributed equally to this work.

its trading conditions are still forming. Because trades settle on the Polygon blockchain, the full history of every market can be reconstructed.

Whoever trades first faces whatever prices the market happens to offer. The gap between the best bid and the best ask, the spread, is the cost a trader who wants to act immediately must accept, and in a market that has just opened it can be wide. A second property is equally observable: just after an early trade, the price sometimes barely moves and sometimes moves substantially in the direction of the trade. How large the opening cost is, how quickly it falls, and how the price behaves after early trades are all questions Polymarket's structure makes answerable with data.

Existing Polymarket research has studied the 2024 election market, wallet-level behaviour, wash trading, and venue-wide characteristics, but it has not measured what it costs to trade in a market's first hours of life, nor compared that cost across the kinds of events the venue hosts. This paper does both. Using roughly 160,000 contracts created between October 2025 and March 2026, we measure, over the first twenty-four hours of each market, the quoted spread, the effective half-spread paid by takers, and the split of that cost, once the price has adjusted, into a realised half-spread (retained by the liquidity provider) and price impact (reflected in subsequent price movement). We then compare these measures across eight event categories: monetary policy, economy, geopolitics, politics, elections, sports, crypto, and other.

The central finding is a mismatch. Ranking the categories by how expensive they are to trade at the open produces one ordering; ranking them by how much the price moves after early trades produces a different one. The two do not line up. The cost of entering a newly opened market and the degree to which early trades move the price are two separate properties, not two views of the same thing. Documenting this dissociation, and setting out which established microstructure explanations are consistent with it, is the main contribution of the paper. The study measures and describes these patterns; it does not attempt to identify which single force, informed trading, inventory risk, maker competition, or platform incentives, produces them.

The paper is organised as follows. Section 2 describes the institutional setting and develops the measurement framework. Section 3 covers the data and sample. Section 4 presents the results of our research questions. Section 5 conclusions.

1.1 Contributions

We organise the analysis around three research questions, each comparing a trading-cost measure across the eight event categories over the first twenty-four hours of a market's life (the *genesis window*).

RQ1 (Posted liquidity). How wide do spreads start, and how quickly do they narrow over the genesis window, and does this differ across categories?

RQ2 (Realised taker cost). How much does it actually cost to trade immediately, and how does that cost differ across categories and over time within the genesis window?

RQ3 (Cost composition). Of that cost, how much stays with the maker and how much reflects the price moving afterwards?

Our contributions are threefold. First, we provide the first systematic measurement of opening-phase trading costs on a prediction market, applying standard microstructure measures to roughly 160,000 Polymarket contracts created between October 2025 and March 2026. Second, we show that these costs are strongly category-dependent from the very first trade: sports markets open near the tick floor with deep, balanced books, while non-sports categories open six to seven times wider with liquidity parked defensively away from fair value, and most compression occurs within the first hour. Third, and centrally, we document a *dissociation*: ranking categories by opening cost (RQ1) and by post-trade price impact (RQ3) produces two orderings that do not agree. A category can open with a wide, defensive book yet show little subsequent price movement, while another opens at moderate width yet moves strongly after early trades. We set out which established microstructure mechanisms this pattern is consistent with, while stressing that the study is descriptive and does not identify a single causal channel.

1.2 Related Work

Spread measurement and decomposition. The bid-ask spread is a central object in empirical microstructure: the cost a trader pays for immediacy [5,16]. We use four measures that build on one another. The quoted spread captures posted liquidity but can mis-state realised cost, since trades may execute inside the spread or walk the book [15,3]. The effective spread, anchored to executions, captures what traders actually pay. Huang and Stoll [11,12] decompose it into a realised spread (retained by the liquidity provider) and price impact (subsequent price movement), separating trades that mainly pay for immediacy from those that move the price. These frictions reflect order-processing costs, inventory risk [2,10], and adverse selection against informed counterparties [8,14]. Our measures separate the observable stages of the trading-cost process rather than isolating these structural channels.

Liquidity formation at openings. Newly opened markets do not begin in a steady state: liquidity is thin, participation is still forming, and early quotes reflect uncertainty about fair value. Work on conventional venues shows that early quotes are highly informative and that the opening phase can be isolated within a defined window, IPO debuts [1], auction openings and pre openings [17,4], and scheduled macroeconomic releases, around which spreads widen and then contract [7]. Polymarket opens a contract with no underwriter, no auction, and no assigned market maker, but the central lesson holds: the early phase is where prices and liquidity form, and it can be studied within a window after opening.

Cross-market variation in liquidity costs. Spreads differ systematically across securities and environments. Adverse selection ties spread width to price impact through a single quantity, the share of informed order flow, so the simplest reading predicts that the two line up across markets [8,14]. Inventory and hedging risk add a second channel: positions that are harder to offload demand wider compensation [13]. Market design is a third: tick size and trading rules shape the observed spread distribution [9]. A fourth is maker competition, which governs how strongly these frictions appear in quotes. Observed cross-sections can reflect several at once, and our category comparisons document where costs differ rather than attribute them to one channel.

Prediction markets and Polymarket. Prediction-market prices can aggregate dispersed information [32], though this reading requires assumptions about trader beliefs and constraints [18]. Most early work focused on accuracy rather than microstructure. Recent Polymarket-specific research has begun to close this gap: Tsang and Yang [30] apply standard microstructure measures to the 2024 election market and document maturation; Mitts and Ofir [19] find wallet-level patterns consistent with informed trading in geopolitical and election markets; Sirolly et al. [29] show that raw volume is distorted by wash trading; and Saguillo et al. [28] document arbitrage deviations even in active markets. Our trade data come from the dataset of Wang et al. [31]. None of this work measures how liquidity costs form during the opening phase of newly created markets, or compares that formation across event categories, the gap this paper fills.

2 Background

Polymarket lists a contract for each real-world event that pays one USDC if a stated outcome occurs and zero otherwise, so its price can be read as an implied probability [25]. Each outcome is a pair of complementary YES/NO tokens whose prices sum to one absent frictions; categorical events (e.g. multi-candidate elections) are implemented through the NegRisk framework and treated here as separate contracts at the token level [24]. Trading runs through a central limit order book in continuous mode: a trader either posts a resting order (acting as *maker*, or liquidity provider) or executes against one (acting as *taker*) [27]. The minimum tick is one cent for most markets, which is a hard lower bound on the observable quoted spread [26]. Crucially for our design, Polymarket pairs off-chain order matching with on-chain settlement on Polygon, so every executed trade is permanently recorded and the full history of any market can be reconstructed [27]. This is the only feature of the venue's mechanics our analysis depends on; resolution occurs well after the genesis window and does not affect the trading-cost measures.

2.1 Liquidity Rewards, Fees, and Maker Incentives

Because each Polymarket market must attract its own liquidity from scratch, the platform actively shapes maker behaviour through two distinct schemes. These schemes are part of the Polymarket environment in which every spread we measure is produced, and they bear directly on how the cross-category results should be read.

The liquidity rewards programme. The rewards programme pays makers a share of a per-market reward pool in proportion to the quality of their resting orders, with the score penalising distance from the midpoint: the closer to the midpoint a maker quotes, the larger the reward [22]. Each market has a configured boundary, the `max_incentive_spread`, outside which orders earn no reward regardless of size, and a `min_incentive_size` setting how large an order must be to count. Both parameters are set per market by Polymarket and are observable through the platform API. There is no platform-wide spread target: because the eligibility boundary is configured market by market, the strength of the incentive to quote tightly varies across markets and, potentially, systematically across categories.

Fee structure introduced during the sample period. During the sample period Polymarket introduced two linked features: taker fees which fund maker rebates [20]. These were launched and rolled out category by category rather than all at once, beginning with crypto and later extending to sports. Both apply only to markets deployed on or after their category's activation date, so a market opened before its category went live is unaffected for its entire life. Within the sample period, fees reached crypto and sports markets; the remaining categories were fee-free throughout.

How it works. The taker on each fill pays a fee. The fees in each category are collected and pooled by the platform and category-specific share of that pool is redistributed daily as a rebate to the makers whose resting orders were filled, in proportion to their qualifying executed liquidity in that market. The rebate share is twenty percent in crypto and twenty-five percent in the other fee-enabled categories. Because the rebate is funded entirely by taker fees, a fee-free category generates neither.

Difference. This rebate is distinct from the liquidity rewards programme described above. They differ in funding source and in what triggers payment. The rewards programme is funded by Polymarket from a fixed pool and pays makers for posting tight resting quotes near the midpoint, whether or not those quotes are ever filled. The rebate is funded out of taker fees and pays makers only when their resting orders are executed. A maker can earn rewards on an unfilled quote but earns a rebate only on a completed trade.

Implications for the analysis. Three points follow. First, the quoted spreads we measure are *composite*: the rewards programme pulls quotes toward the midpoint within each market's configured boundary, so part of the cross-category dispersion in opening spreads may reflect differences in configured parameters or maker participation rather than event characteristics alone. We do not attempt to separate the reward channel from the other forces shaping liquidity. Second, the fee rollout affects only a limited part of the sample. The largest extension, Fee Structure V2, went live on 30 March 2026 [20],one day after our sample endpoint of 29 March 2026, which is precisely why the window ends there: extending it would mix fee-active and fee-free markets for most categories. Within the sample, fees touch only the short-duration crypto markets we exclude as recurring microbets and a small subset of sports markets opened after 18 February 2026, so the headline results are measured in an essentially fee-free environment. Third, persistently wide spreads remain informative despite these incentives: reward eligibility gives makers a reason to quote tightly but does not compel them, so a category that still opens wide, or compresses slowly, is empirically meaningful even when its cause cannot be assigned to event characteristics alone.

3 Data

The analysis combines three sources, each capturing a different layer of Polymarket activity. Executed trades and the standing order book are recorded through separate channels, trades settle on-chain, while resting orders live in Polymarket's off-chain infrastructure and are not preserved in the trade record, so both an on-chain trade dataset and an order-book snapshot feed are required, together with category metadata.

On-chain trades. The matched trade record comes from the open Polymarket dataset of Wang et al. [31], released by the Shanghai Innovation Institute through the HuggingFace data repository and reconstructed from Polygon blockchain data. It spans the platform's inception to 31 March 2026. We use two of its files. The `markets.parquet` file provides a static snapshot of 734,790 markets (identifier, slug, question text, and resolution outcome where available). The `quant.parquet` file contains 568,590,741 normalised peer-to-peer trade records between 21 November 2022 and 31 March 2026, each carrying a timestamp, executed price, quantity, wallet identifiers, and the linking market identifier. Every trade is represented from the YES-token perspective, which makes prices and directions directly comparable across the two complementary outcome tokens without manual conversion.

Order-book snapshots from Dome. On-chain data records what executed but not the resting liquidity a taker faced, so quoted spreads, midpoints, and depth are built from the Dome API, a developer-infrastructure service that recorded Polymarket order-book states while markets were live and retained them after close [6]. Two properties of Dome shape the entire spread analysis. First, each recorded state reports the best bid, the best ask, the depth available at each visible level, the midpoint, and a timestamp, enough to reconstruct the top of the book a taker traded against. Second, and unlike a fixed-interval feed, *Dome records a new state only when the observed order book changes.* This is not a sampling limitation but an informational feature: the absence of a new snapshot is itself evidence that the book has not moved, so carrying the last observed state forward to a later timestamp preserves the true state of the book rather than interpolating one. Dome began recording on 14 October 2025 and was discontinued on 28 April 2026, after the end of our sample window. The order-book snapshot dataset is open sourced at https://huggingface.co/datasets/vpcapitano/polymarket-genesis-liquidity.

Category tags. Because market metadata already comes from `markets.parquet`, the Polymarket Gamma API [21] is used only to attach the one-or-more category tags Polymarket assigns to each market (e.g. politics, sports, crypto, geopolitics). These tags are the starting point for the classification into eight analytical categories: monetary policy, economy, geopolitics, politics, elections, sports, crypto[3], and other, built from a rule-based pass over the tags, a language-model review of borderline cases, and a manual audit, following each market's *resolving event* rather than the wording of its title.

Sample window and eligibility. The window runs from 17 October 2025 to 29 March 2026, the period both data sources cover: it starts a few days after Dome coverage begins and ends one day before the Fee Structure V2 rollout (Section 2.1) and a month before the CLOB v2 upgrade [23], keeping the sample within a single, largely fee-free period. A contract is eligible if (i) its first executed trade falls inside the window, we define market *opening* as that first trade, denoted T_0; (ii) it has at least one Dome state within the twenty-four-hour genesis window; and (iii) it accumulated at least 5,000 USDC of cumulative volume. The volume filter serves two purposes: it kept data collection feasible within the fixed window left by Dome's discontinuation, and it sets aside markets with almost no order-book activity, for which the spread measures would carry little meaning. After these filters the universe falls from 497,506 to 162,342 contracts. Categorical events are treated as separate observations at the contract level, each with its own genesis window beginning at its own first trade. The analysis code is available at https://github.com/vpercap-rg/polymarket-genesis-liquidity.

4 Results

A few conventions apply throughout. Every reported number is a *median* across markets rather than a mean, because a small number of unusual markets would otherwise dominate any average. We suppress any cell built from fewer than thirty markets. The analysis is deliberately descriptive: we compare categories by the size of the gaps between their medians and do not run significance tests. Alongside the full sample we report a subset we call *Type A*: markets that traded actively enough early on for the spread measures to capture genuine price formation rather than a stale, untouched book. A market qualifies if it recorded more than ten trades in its first three hours and traded at least 1% of its volume

[3] Recurring, mechanical crypto contracts, e.g., five minutes, fifteen minutes, or hour, are excluded from the crypto category, because they repeat continuously rather than marking a distinct information event, they are set aside along with other mechanical formats such as mention- and tweet-count markets.

Table 1. Key variables and their sources.

Variable	Source	Description
Trade timestamp	Wang et al. (on-chain)	Execution time; minimum defines T_0
Trade price	Wang et al. (on-chain)	Executed price, YES-token perspective
Trade quantity	Wang et al. (on-chain)	Size of the executed trade
Wallet identifiers	Wang et al. (on-chain)	Counterparty addresses
Market identifier	Wang et al. (on-chain)	Links trade to market metadata
Question text	`markets.parquet`	Market title, used in classification audit
Best bid / best ask	Dome API	Top-of-book quotes at a recorded state
Level depth	Dome API	Quantity available at each visible level
Midpoint	Dome API	Mean of best bid and best ask
Snapshot timestamp	Dome API	Time of an order-book state change
Category tag	Gamma API	Polymarket subject tag(s) per market

within the first 24 hours. When a pattern holds in this subset, it is not an artefact of dormant markets. Sample sizes vary from table to table for a given category because each measure has its own eligibility and trade-to-snapshot matching requirements.

One fact recurs in every table and every figure: *sports sits at one extreme of essentially every distribution.* Because of that the more interesting story is the variation *among the other seven categories*. We organise the section around the 3 research questions: how wide markets open and how fast they tighten (RQ1), what it actually costs to trade early (RQ2), and where that cost goes (RQ3).

At the other end, the other category is a residual bucket rather than a coherent event type: it aggregates markets from many unrelated subjects, so its medians should be read as a summary of a heterogeneous mixture rather than the signature of a single kind of event.

4.1 How Wide Markets Open, and How Fast They Tighten (RQ1)

Opening spreads differ enormously across categories. At the very first trade, the median quoted spread runs from three cents in sports to twenty-one cents in economy, a seven times gap (Table 2, Left). Sports is in a class of its own, but the non-sports categories do not share a common level either. They fall into three bands: politics and geopolitics open tightest (8–10 ¢), monetary policy, elections, and "other" sit in the middle (13–14 ¢), and crypto and economy open widest (19–21 ¢). Restricting to Type A markets lowers every non-sports level, in some cases by half, but leaves the ordering essentially intact (Table 2, Right). So the wide full-sample numbers are partly inflated by dormant markets, but the *spread between categories* is not. Opening liquidity is category-specific, not a single venue-wide constant.

Whatever a market opens at, most of the tightening happens fast. By six hours, the non-sports categories have shed roughly half to two-thirds of their opening spread; by twenty-four hours, every category except sports sits between two

Table 2. Opening spread by market category. Left: full sample. Right: Type A.

Category	N	Open (¢)
Monetary policy	111	13.0
Economy	201	21.0
Geopolitics	1,403	10.0
Politics	605	8.0
Elections	771	13.0
Sports	25,357	3.0
Crypto	565	19.0
Other	1,801	14.0

Category	N	Open (¢)
Monetary policy	56	7.0
Economy	108	10.5
Geopolitics	1,133	7.0
Politics	516	7.0
Elections	437	10.0
Sports	19,309	3.0
Crypto	404	15.0
Other	1,195	11.0

and five cents (Fig. 1, Tables 3). The genesis window is therefore not a slow, steady flattening, it is a sharp drop in the first few hours followed by a long, quiet tail. The categories tied to scheduled events (elections, monetary policy) tighten the most gradually, which fits a reading in which the market waits for an announcement before settling on a price.

Table 3. Spread compression by category (full sample): absolute

Category	N	T_0	6h	12h	18h	24h
Mon. pol.	196	13.0	8.0	7.0	6.0	5.0
Economy	332	21.0	8.0	5.0	5.0	4.0
Geopol.	2,152	10.0	4.0	3.0	2.2	2.0
Politics	821	8.0	3.0	3.0	2.9	2.0
Elections	1,090	13.0	7.0	5.0	4.0	3.6
Sports	46,131	3.0	3.0	2.0	2.0	3.0
Crypto	880	19.0	7.0	5.0	3.0	3.0
Other	2,690	14.0	6.0	5.0	4.0	3.9

Sports is the apparent exception, but its near-zero compression is mechanical rather than economic: it opens at three cents, already a hair above the one-cent tick floor, with almost no room left to move. The Type A view sharpens the rest of the picture (Fig. 4): every non-sports category converges to a median near two cents by hour twenty-four, despite starting anywhere from seven to fifteen.

Why do the non-sports categories open so wide? The answer is in the *shape* of the opening book (Table 4). A non-sports market is not empty, it has visible resting liquidity, but that liquidity is parked defensively, far from fair value, where it fills only at prices that punish the taker. Three numbers make the contrast concrete. The median opening spread is about five cents in sports against thirty-plus cents in every non-sports category. Sports shows a median of 171 contracts resting at the best ask, versus just 32–73 elsewhere. And only a quarter of the visible sports ask-side depth sits at extreme prices (at or above 0.90),

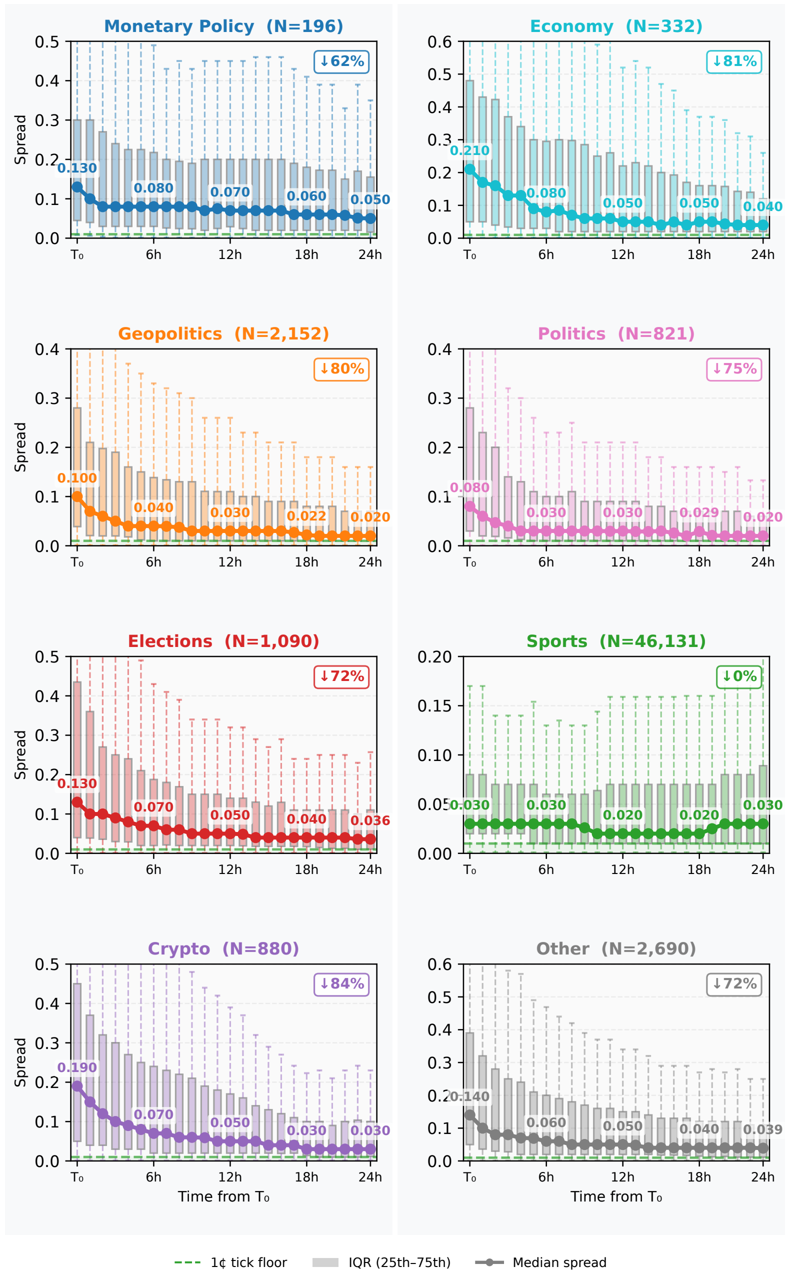

Fig. 1. Bid–ask convergence from T_0 by category. Bars show the 25th–75th percentile range, the solid line the median, the dashed line the one-cent tick floor; corner percentages denote proportional compression from T_0 to 24h.

Table 4. Order-book structure at T_0 by category. Spread in probability units; top-ask size in contracts; extreme ask is the share of top-ten ask depth at ≥ 0.90; stub is the share of markets with best ask ≥ 0.85 and best bid ≤ 0.15.

Category	N	Spread	Top-ask size	Extreme ask	Stub
Monetary policy	192	0.30	40	75.6%	9.9%
Economy	305	0.34	32	68.3%	10.2%
Geopolitics	1,811	0.32	50	80.0%	15.3%
Politics	734	0.35	41	88.3%	20.3%
Elections	1,076	0.34	33	67.6%	16.3%
Sports	45,838	0.05	171	23.9%	5.1%
Crypto	881	0.34	73	91.9%	15.2%
Other	2,498	0.36	37	75.9%	17.8%

against two-thirds to over ninety percent in the non-sports categories. In short, sports opens with a deep, balanced book sitting near fair value, while the others open with their liquidity pushed out toward the edges.

Does that opening spread actually cost the first trader anything? It does, it is a real bill, not a number on a screen. The median first trade in every category executes at exactly the price the visible book implies, an execution ratio of 1.00 (Table 8, Fig. 5). The typical first taker pays precisely what the posted spread says they will. The dispersion in opening spreads documented above therefore translates almost one-for-one into a dispersion in what the first real trader actually pays.

4.2 What It Costs to Trade Early, and for How Long (RQ2)

The cost of being early starts brutal and then collapses. Measured as the *effective half-spread*, what a taker actually pays per share, relative to the prevailing midpoint, execution costs fall by roughly a factor of ten over the first hour (Table 9, Fig. 2).

The worst moment is the very first minute. There, every non-sports category charges a median half-spread of 12 to 16 cents, while sports charges 1.5. On a contract that ultimately pays either zero or one dollar, paying twelve to sixteen cents for a single share is a striking cost, and the gap between sports and everyone else is at its widest at the instant a market opens.

Then it closes quickly. Within the 1–6 minute window, non-sports costs have roughly halved. By 30–60 minutes, every non-sports category is within a factor of two of the sports value. By the one-to-three-hour mark, they have essentially converged. The economic penalty for being early is therefore real but *concentrated in the opening minutes*, not smeared across the whole genesis day. (A few non-monotone cells in monetary policy and economy are small-sample effects, where individual buckets draw on fewer than fifty markets.)

Sports, once again, is flat: it begins at the tick floor and stays there. It has no liquidity-formation problem to solve. The other categories spend their first hour solving exactly that problem, and the takers who trade during that hour pay for the privilege.

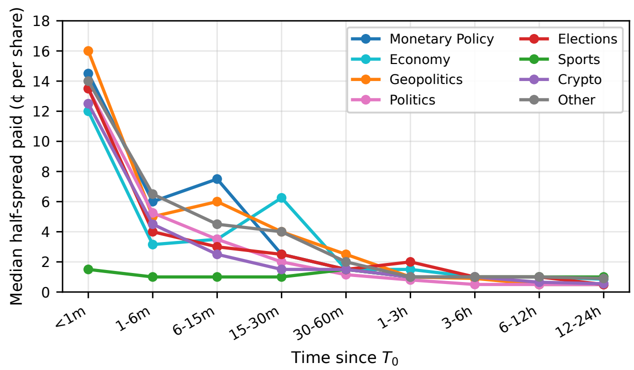

Fig. 2. Median effective half-spread paid by takers (cents per share) across nine time buckets over the first 24 hours after T_0, by category.

4.3 Where the Cost Goes: The Central Dissociation (RQ3)

So far we have measured how much early trading costs. The final question is what that cost actually *is*, and this is where the paper's main result appears.

The intuition. When a taker pays an effective half-spread, that payment divides into two parts. One part is compensation to the liquidity provider for standing ready to trade; the maker keeps it, and we call it the *realised half-spread*. The other part is the price moving against the maker right after the trade, because the trade turned out to carry information the market then absorbed; we call that *price impact*. A compact summary of the split is the *R/E ratio*: the share of the taker's cost that the maker gets to keep (realised divided by effective). An R/E near 1.0 means early trades were essentially uninformative, the maker pockets almost the whole spread. A low R/E means early trades moved the price, most of what the taker paid was really the market learning something.

Method, briefly. We sign each trade with the Lee and Ready (1991) quote rule, a trade above the midpoint is a buy, below it a sell, and simply drop the small share of trades that land exactly on the midpoint (2.8–4.8%, Table 5). We measure the split over the first six minutes after T_0, letting the midpoint adjust for one hour ($\tau = 1$h, Table 6, Fig. 3).

The result is not what a single-mechanism story predicts. If wider spreads simply meant more informed trading, the cost ranking and the price-impact ranking would line up. They do not. Sports pays 1.0 ¢, keeps 1.0 ¢, and shows zero median price impact, for an R/E of 1.00: early sports trades carry no information the market later incorporates. Geopolitics, politics, elections, and other are the mirror image, with R/E ratios of 0.25, 0.35, 0.29, and 0.33, most of what their takers pay is the price moving, not maker rent. Economy is the striking anomaly: it opens with one of the widest spreads in the whole sample, yet its makers keep most of it (R/E = 0.86). In *composition* economy looks like sports, even though in *cost level* it sits at the opposite extreme. Crypto is intermediate (0.54), and monetary policy posts the lowest ratio (0.15) but rests on only thirty-four markets and should be read as suggestive.

Table 5. Mid-touch rate by category. N_{mid} trades execute exactly at the prevailing midpoint and are excluded from the decomposition.

Category	N	N_{mid}	% mid
Monetary policy	7,254	210	2.89%
Economy	34,256	1,316	3.84%
Geopolitics	522,427	18,108	3.47%
Politics	214,438	8,020	3.74%
Elections	58,866	1,762	2.99%
Sports	10,962,722	465,776	4.25%
Crypto	143,945	6,842	4.75%
Other	258,282	7,207	2.79%

Table 6. Spread decomposition (¢) at $\tau = 1\text{h}$, first six minutes, Type A. Components are category-level medians computed independently, so they need not sum to the effective half-spread; the identity holds trade-by-trade.

Category	N	Eff. $\frac{1}{2}$	Realised	Impact	R/E	% pos.	P25–P75
Monetary policy	34	5.00	0.73	4.30	0.15	55.9%	$[-5.75, +11.50]$
Economy	221	3.50	3.00	2.00	0.86	82.4%	$[+0.25, +10.00]$
Geopolitics	1,864	12.00	3.00	5.00	0.25	67.7%	$[-0.50, +12.50]$
Politics	1,021	10.00	3.50	4.00	0.35	67.6%	$[-1.50, +12.50]$
Elections	981	5.00	1.45	3.50	0.29	67.4%	$[-1.00, +4.50]$
Sports	48,908	1.00	1.00	0.00	1.00	67.1%	$[-0.50, +2.50]$
Crypto	677	6.50	3.50	1.50	0.54	67.7%	$[-2.00, +16.50]$
Other	2,045	7.50	2.50	3.50	0.33	68.4%	$[-1.50, +11.50]$

This is the dissociation, and it is the central finding. Rank the categories by how expensive they are to enter (RQ1), then rank them by how much their early trades move the price (RQ3), and the two orderings disagree. Sports and

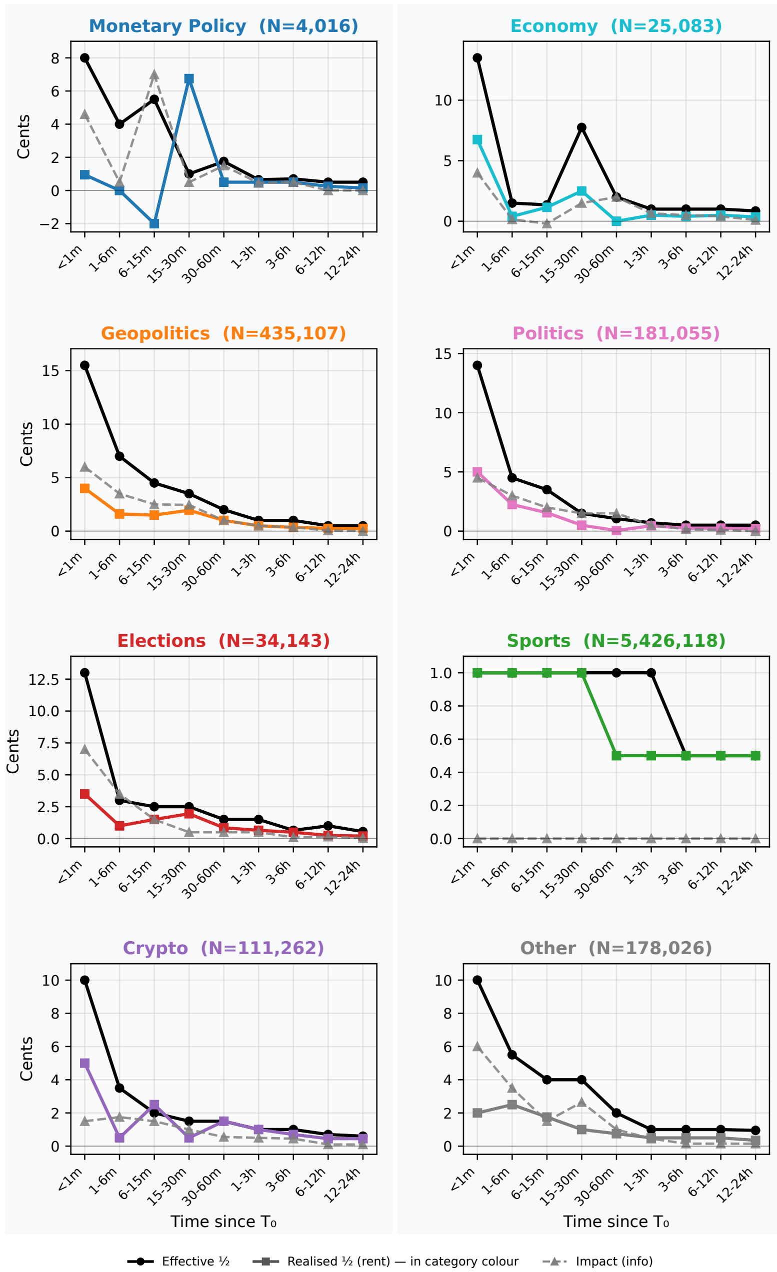

Fig. 3. Effective half-spread decomposed into realised half-spread and half price impact across nine time buckets, by category. Lines are category-level medians; adjustment horizon $\tau = 1\text{h}$; Type A markets.

economy both let the maker keep most of the spread, yet sit at opposite ends of the cost scale. Geopolitics, politics, elections, and other give most of the spread back as price impact, at differing cost levels of their own. How costly a market is to enter and how informative its early trades are turn out to be *two separate properties of a category, not two readings of the same thing.*

Table 7. Distribution of realised half-spread (¢) at $\tau = 1\text{h}$, first six minutes, Type A.

Category	N	P10	P25	P50	P75	P90	% neg.
Monetary policy	34	−10.40	−5.75	0.73	11.50	18.00	38.2%
Economy	221	−0.50	0.25	3.00	10.00	20.50	15.8%
Geopolitics	1,864	−5.50	−0.50	3.00	12.50	29.50	29.7%
Politics	1,021	−12.00	−1.50	3.50	12.50	30.50	31.9%
Elections	981	−2.00	−1.00	1.45	4.50	19.50	30.2%
Sports	48,908	−4.00	−0.50	1.00	2.50	9.50	27.6%
Crypto	677	−13.50	−2.00	3.50	16.50	37.50	30.1%
Other	2,045	−7.50	−1.50	2.50	11.50	26.30	30.9%

One caveat keeps the claim honest. These are medians over wide distributions. The realised half-spread is positive at the median in every category, but its lower quartile is negative in all categories except economy, and between 16 and 38 percent of trades show a negative realised half-spread at the one-hour horizon (Table 7). Because trade direction is inferred, and because the "future" midpoint absorbs whatever happened to move the market over the following hour, the R/E ratio describes post-trade midpoint behaviour rather than cleanly isolating informed trading. The dissociation itself, however, is robust to these caveats: the orderings disagree regardless of how the components are read.

5 Conclusion

We examined what it costs to be among the first to trade a Polymarket contract, and how that cost changes over a market's first day, across roughly 160,000 markets grouped into eight kinds of events.

The type of event matters from the very start. In sports, the opening orders sit close to a sensible price and there are plenty of them, so the first trader pays little; in every other category the orders are there but sit well away from where the market soon trades, so the first trader pays far more, several times more in economy than in sports. That cost then drops quickly, and by the end of the first hour most of it is gone. Being early is expensive, but mainly for a few minutes.

What that cost is made of is the more surprising part. Some of it is simply the market-maker's reward for offering to trade; the rest is the price reacting to the trade, as though the trade revealed something. Separating the two shows that the categories most expensive to enter are not the ones whose early trades move the price most. Sports and economy sit at opposite ends of the cost scale, yet in both the maker keeps nearly all of what the trader pays; in politics and geopolitics, much of the cost is instead the price reacting. A market's entry cost, then, says little about how informative its first trades are, the two are different things.

The event categories are, in the end, a convenient label for deeper differences. Sports contracts do not form their prices in isolation: they sit alongside mature betting exchanges and bookmakers, so a reference price exists from the moment the contract opens, and a reading consistent with our numbers is that sports opens tight and shows no drift because its price discovery has largely happened elsewhere. The categories that open wide and move afterwards have no such external anchor. What separates the categories may then be less the subject matter than whether an event is scheduled, how much external information is available, and how mature the surrounding market is, with the labels standing in for these underlying axes. Economy is the reminder that the two properties are separate: its events are largely scheduled public releases, so early takers are rarely better informed, and adverse selection stays low even though its opening spread is among the widest.

Our sample ends deliberately in a fee-free environment, one day before Fee Structure V2 took the platform to taker fees funding maker rebates. Because that change pays makers only on filled orders, it sharpens the incentive to quote close to fair value at the open, so a natural prediction is that rebates should compress opening spreads and push the R/E ratio upward, most visibly in the categories that currently open widest. Whether the dissociation survives that shift is a question our sample cannot answer, and a natural target for work on the post-fee platform.

The same price or volume figure carries different information depending on where and when it is observed: a cheap, liquid sports market and an expensive, thin geopolitics one are not comparable at face value, and neither is a market in its first minutes and the same market an hour later. Studies that treat Polymarket prices as probabilities, or volume as attention, therefore risk reading a microstructure difference as a difference in forecast quality or public interest.

References

1. Aggarwal, R., Conroy, P.: Price discovery in initial public offerings and the role of the lead underwriter. Journal of Finance **55**(6), 2903–2922 (2000). https://doi.org/10.1111/0022-1082.00312
2. Amihud, Y., Mendelson, H.: Dealership market: Market-making with inventory. Journal of Financial Economics **8**(1), 31–53 (1980). https://doi.org/10.1016/0304-405X(80)90020-3
3. Bessembinder, H.: Issues in assessing trade execution costs. Journal of Financial Markets **6**(3), 233–257 (2003). https://doi.org/10.1016/S1386-4181(02)00064-2
4. Biais, B., Hillion, P., Spatt, C.: Price discovery and learning during the preopening period in the paris bourse. Journal of Political Economy **107**(6), 1218–1248 (1999). https://doi.org/10.1086/250095
5. Demsetz, H.: The cost of transacting. Quarterly Journal of Economics **82**(1), 33–53 (1968). https://doi.org/10.2307/1882244

6. Dome: Dome API: Polymarket order book data. Developer infrastructure service (2026), order-book snapshot coverage 14 October 2025 – 28 April 2026
7. Fleming, M.J., Remolona, E.M.: Price formation and liquidity in the U.S. treasury market: The response to public information. Journal of Finance **54**(5), 1901–1915 (1999). https://doi.org/10.1111/0022-1082.00172
8. Glosten, L.R., Milgrom, P.R.: Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. Journal of Financial Economics **14**(1), 71–100 (1985). https://doi.org/10.1016/0304-405X(85)90044-3
9. Goldstein, M.A., Kavajecz, K.A.: Eighths, sixteenths, and market depth: Changes in tick size and liquidity provision on the NYSE. Journal of Financial Economics **56**(1), 125–149 (2000). https://doi.org/10.1016/S0304-405X(99)00061-6
10. Ho, T., Stoll, H.R.: Optimal dealer pricing under transactions and return uncertainty. Journal of Financial Economics **9**(1), 47–73 (1981). https://doi.org/10.1016/0304-405X(81)90020-9
11. Huang, R.D., Stoll, H.R.: Dealer versus auction markets: A paired comparison of execution costs on NASDAQ and the NYSE. Journal of Financial Economics **41**(3), 313–357 (1996). https://doi.org/10.1016/0304-405X(95)00867-E
12. Huang, R.D., Stoll, H.R.: The components of the bid-ask spread: A general approach. Review of Financial Studies **10**(4), 995–1034 (1997). https://doi.org/10.1093/rfs/10.4.995
13. Jameson, M., Wilhelm, W.: Market making in the options markets and the costs of discrete hedge rebalancing. Journal of Finance **47**(2), 765–779 (1992). https://doi.org/10.1111/j.1540-6261.1992.tb04409.x
14. Kyle, A.S.: Continuous auctions and insider trading. Econometrica **53**(6), 1315–1335 (1985). https://doi.org/10.2307/1913210
15. Lee, C.M.C.: Market integration and price execution for NYSE-listed securities. Journal of Finance **48**(3), 1009–1038 (1993). https://doi.org/10.1111/j.1540-6261.1993.tb04028.x
16. Madhavan, A.: Market microstructure: A survey. Journal of Financial Markets **3**(3), 205–258 (2000). https://doi.org/10.1016/S1386-4181(00)00007-0
17. Madhavan, A., Panchapagesan, V.: Price discovery in auction markets: A look inside the black box. Review of Financial Studies **13**(3), 627–658 (2000). https://doi.org/10.1093/rfs/13.3.627
18. Manski, C.F.: Interpreting the predictions of prediction markets. Economics Letters **91**(3), 425–429 (2006). https://doi.org/10.1016/j.econlet.2006.01.004
19. Mitts, J., Ofir, M.: From iran to Taylor Swift: Informed trading in prediction markets (2026), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=6426778, working paper, SSRN 6426778
20. Polymarket: Fees. Polymarket Documentation (2026), https://docs.polymarket.com/trading/fees, accessed 18 May 2026
21. Polymarket: Gamma API. Polymarket Documentation (2026), https://docs.polymarket.com/api-reference/gamma, accessed 18 May 2026
22. Polymarket: Liquidity rewards. Polymarket Documentation (2026), https://docs.polymarket.com/marketmakers/liquidity-rewards, accessed 17 May 2026
23. Polymarket: Migrating to CLOB V2. Polymarket Documentation (2026), https://docs.polymarket.com/v2-migration, accessed 18 May 2026
24. Polymarket: Negative risk markets. Polymarket Documentation (2026), https://docs.polymarket.com/advanced/neg-risk, accessed 18 May 2026
25. Polymarket: Prices and order book. Polymarket Documentation (2026), https://docs.polymarket.com/concepts/prices-orderbook, accessed 18 May 2026

26. Polymarket: Tick size API reference. Polymarket Documentation (2026), https://docs.polymarket.com/api-reference/market-data/get-tick-size, accessed 18 May 2026
27. Polymarket: Trading overview. Polymarket Documentation (2026), https://docs.polymarket.com/trading/overview, accessed 18 May 2026
28. Saguillo, O., Ghafouri, V., Kiffer, L., Suarez-Tangil, G.: Unravelling the probabilistic forest: Arbitrage in prediction markets (2025), arXiv preprint arXiv:2508.03474
29. Sirolly, A., Ma, H., Kanoria, Y., Sethi, R.: Network-based detection of wash trading (2025), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5714122, working paper, SSRN 5714122
30. Tsang, K.P., Yang, Z.: The anatomy of Polymarket: Evidence from the 2024 presidential election (2026), arXiv preprint arXiv:2603.03136
31. Wang, Z., Chao, L., Bao, Y., Cheng, L., Liao, J., Li, Y.: Polymarket data: Complete data infrastructure for Polymarket. Dataset, Shanghai Innovation Institute, HuggingFace (2026), https://huggingface.co/datasets/SII-WANGZJ/Polymarket_data
32. Wolfers, J., Zitzewitz, E.: Prediction markets. Journal of Economic Perspectives **18**(2), 107–126 (2004). https://doi.org/10.1257/0895330041371388

A Plots and Tables

Table 8. First-trade execution ratio by category. % ≈ 1.0 is the share of first trades at the posted best quote; % < 0.85 executes substantially inside the spread.

Category	N	Median	P25	P75	% ≈ 1.0	% < 0.85
Monetary policy	192	1.00	0.909	1.00	77.6%	21.9%
Economy	305	1.00	0.969	1.00	80.0%	18.4%
Geopolitics	1,811	1.00	0.967	1.00	81.2%	17.5%
Politics	734	1.00	0.956	1.00	80.4%	18.1%
Elections	1,076	1.00	0.961	1.00	83.0%	15.6%
Sports	45,838	1.00	1.000	1.00	82.6%	15.9%
Crypto	881	1.00	0.870	1.00	75.7%	23.2%
Other	2,498	1.00	0.953	1.00	81.6%	17.6%

Table 9. Median effective half-spread paid (¢) by time bucket and category (full sample). MonPol = monetary policy; Geo = geopolitics.

Bin	MonPol	Economy	Geo	Politics	Elections	Sports	Crypto	Other
$<$1m	14.50	12.00	16.00	13.50	13.50	1.50	12.50	14.00
1–6m	6.00	3.15	5.00	5.25	4.00	1.00	4.50	6.50
6–15m	7.50	3.50	6.00	3.50	3.00	1.00	2.50	4.50
15–30m	2.50	6.25	4.00	2.00	2.50	1.00	1.50	4.00
30–60m	1.50	1.50	2.50	1.15	1.50	1.50	1.50	2.00
1–3h	1.00	1.50	1.00	0.80	2.00	1.00	1.00	1.00
3–6h	0.95	1.00	0.90	0.50	1.00	1.00	1.00	1.00
6–12h	0.50	1.00	0.50	0.50	1.00	1.00	0.65	1.00
12–24h	0.50	0.85	0.50	0.50	0.50	1.00	0.55	0.90

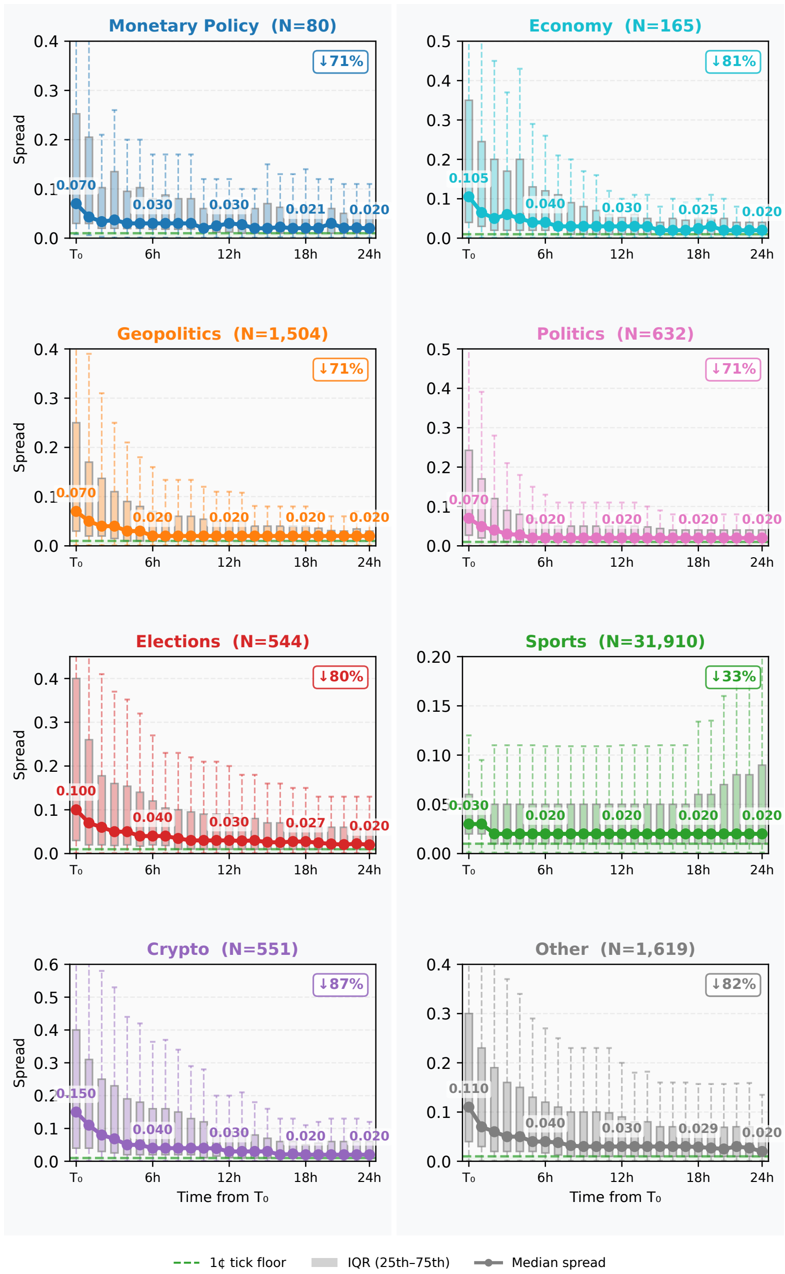

Fig. 4. Bid–ask convergence from T_0 by category (Type A markets, baseline threshold). Conventions as in Fig. 1.

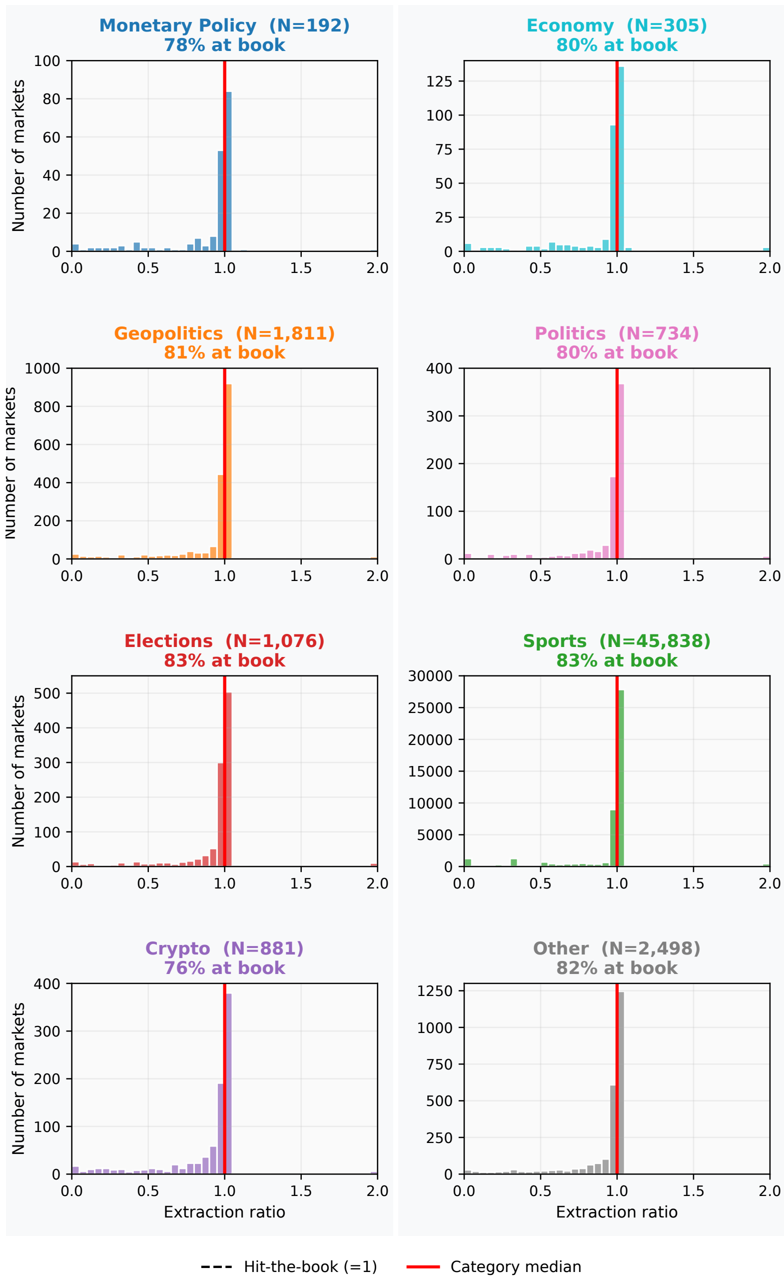

Fig. 5. First-trade execution ratio at T_0 by category. The dashed line marks the hit-the-book threshold at 1.00 and the solid line the median.