Master in International Finance

RESEARCH PAPER

Academic Year 2025 - 2026

Efficiently Inefficient: Polymarket's Pricing of Federal Reserve and Bitcoin Events

Valeriu Gîrlea
Velissarios Tsagkalidis

Under the supervision of Prof. Alina Roşu
President of the Jury: Prof. Daniel Schmidt

Keywords: Prediction Markets, Polymarket, Price Discovery, Law of One Price, Arbitrage

Public report

AI disclaimer

AI was used as a tool to assist with wording, language polishing, and the writing of analysis code in this Research Paper. The intellectual content, research design, arguments, and findings are those of the authors.

Contents

List of Figures

List of Tables

Abstract

This thesis examines whether the law of one price holds at the boundary between Polymarket, a prediction-market venue, and the established financial instruments written on the same events. Two case studies anchor the comparison, pairing Polymarket against CME Federal Funds futures and against CME Bitcoin options. For each pairing the thesis asks whether the two venues agree on the price of the same event, which of them discovers the common price first, and whether any spread between them survives as a profit once real frictions are paid.

1 Introduction

Forecasting future events has long been a central concern of economics and finance, where an accurate prediction carries both informational and monetary value. The traditional tools for the task, expert judgment and econometric modelling, each have well-known limitations (Wolfers & Zitzewitz, 2004; Arrow et al., 2008). Expert forecasts are difficult to aggregate and slow to revise, while structural models are only as responsive as the data and assumptions on which they rest, so both tend to adapt sluggishly to new information as it arrives.

Prediction markets offer a different mechanism for generating forecasts. A question is posed publicly together with its possible outcomes, and participants trade contracts that pay out according to which outcome occurs. Anyone who believes that the price attached to an outcome misstates its true likelihood can trade against it, and stands to profit if the eventual resolution proves them right. The price of each contract therefore aggregates the dispersed information and beliefs of many participants into a single, continuously updated number that can be read as the market's implied probability of the event.

Once a niche curiosity, these markets have moved firmly into the mainstream. Polymarket, a prediction-market protocol built on the Polygon blockchain, reports first-quarter trading volume above twenty-five billion dollars and single-day volume exceeding four hundred million (Arkham Intelligence, 2026), and has emerged as the leading platform of its kind. Nothing illustrates this rise more clearly than the 2024 United States presidential election, by a wide margin the largest event Polymarket has ever hosted and one that nothing since has matched. The presidential race alone reportedly drew more than 3.7 billion dollars in trading volume, and major news outlets cited its odds as real-time indicators of public sentiment throughout the campaign (CNBC, 2024).

This expansion has at the same time exposed a structural limitation of the prediction-market ecosystem as a whole. Across a heterogeneous landscape of operator-run platforms and decentralized protocols, there is no shared, machine-verifiable notion of event identity. Traditional financial assets are pinned to globally unique identifiers such as ISINs, futures contract codes, or token addresses, but a prediction market specifies its contingent claim through a platform-specific natural-language description together with its own resolution rules, oracle sources, and cutoff times (Polymarket, n.d.). Determining whether two markets promise the same payoff is therefore neither trivial nor reliably automatable, which is precisely what complicates comparison across venues (Shepstone, 2026).

This thesis turns that comparison problem into an opportunity by anchoring it to events whose identity is unambiguous. A prediction-market contract and a traditional-finance instrument can reference the very same underlying event, and when they do, financial theory holds that they should carry the same price. We examine two such pairings on Polymarket. The first is the market "Fed decision in {month}?", which prices the Federal Reserve's interest-rate decision at a given FOMC meeting and whose natural counterpart is the CME 30-day Federal Funds futures contract. The second is the market on whether Bitcoin closes above a specified price level on a given date, whose counterpart can be reconstructed from CME Bitcoin options. Because each event resolves against a precise, settlement-grade reference rather than a free-text description, the two sides can be compared cleanly, which sidesteps the event-identity problem that otherwise obstructs cross-venue work.

For each pairing we ask three questions. Do the two venues agree on the price of the same event? When they disagree and then reconverge, which venue moves first, that is, where is new information incorporated and where is it merely followed? And does the convergence of the two prices leave room for an exploitable trading opportunity once real transaction costs are paid? Together these questions speak to how efficiently a fast-growing, retail-driven prediction market prices risk relative to the mature instruments of finance, and to whether the informational content of prediction-market prices can be harvested or only observed. The next chapter positions this inquiry within the literature on state pricing, prediction-market efficiency, and price discovery, before the empirical chapters develop each case study in turn.

2 Literature Review

2.1 Prediction-market contracts as Arrow-Debreu securities

The natural starting point for reading a prediction-market price is the most elementary object in asset pricing. In the time-state preference framework of Arrow (1964) and Debreu (1959), the primitive security is an elementary claim that pays one unit if a particular state of the world obtains at a given date and nothing in every other state. A Polymarket outcome contract is exactly this object. It pays one dollar if its event resolves true and zero otherwise, so its traded price is, by construction, the price today of one dollar delivered in that single state. Such a price is a state price, the cost of insuring one unit of consumption against the realization of one state.

Two consequences follow from this identification. First, a state price is mechanically a probability up to discounting. Write π_s for the price of the claim that pays in state s, let r be the riskless rate over the contract's horizon, and let $x = (x_s)$ be any payoff vector across states. The no-arbitrage value of x is then $p(x) = \sum_s \pi_s x_s = (1+r)^{-1}\,\mathbb{E}^Q[x]$, where Q is the risk-neutral measure under which $\pi_s = (1+r)^{-1}Q(s)$. At the short horizons and near-zero discounting of a typical event contract, $\pi_s \approx Q(s)$, so the quoted price is, to first order, the risk-neutral probability of the state. Reading a Polymarket price as an implied probability is therefore not a convenient heuristic. It is the definition of the object being traded. Second, when a market lists an exhaustive and mutually exclusive set of outcomes with quoted prices p_i, those prices must sum to the value of a certain dollar, $\sum_i p_i = (1+r)^{-1} \approx 1$, which is precisely the unit-sum identity that the platform's prices are designed to satisfy. The Arrow–Debreu framing thus delivers, from first principles, both the probabilistic reading of a single contract and the adding-up constraint across a market.

2.2 No-arbitrage and the law of one price

The modern statement of the law of one price under no-arbitrage is due to Harrison and Kreps (1979). They consider agents trading claims x in a space X at prices set by a linear functional π on a state space $(\Omega, \mathcal{F}, P)$. Their central result (Theorem 1) is that a price system is viable, meaning consistent with some agent's optimum and hence free of arbitrage, "if and only if there exists a continuous and strictly positive extension of π to all of X." That extension is a system of state prices, equivalently an equivalent martingale measure under which discounted prices are martingales. Arbitrage-free pricing and the existence of state prices are therefore one condition seen from two sides.

The same logic, applied across venues rather than within one, is what binds a prediction market to the rest of finance. Ross (1978), whose argument rests on the premise that "all arbitrage opportunities will be exhausted in markets with open access," shows that once such a positive linear operator exists, "quite complicated return streams [can] be valued by simply using existing market valuations." Bansal and Viswanathan (1993) reach the same place from no-arbitrage, deriving a single "low-dimensional nonnegative [...] pricing kernel" that "prices both primitive and derivative securities."

The implication for this thesis is that if a Polymarket contract and some established financial instrument both pay off on the realization of the same event, no-arbitrage requires one state price to value both, whoever happens to trade on each venue. The familiar objection that one number is a subjective crowd belief while the other is a risk-neutral market price dissolves, as both are prices of traded claims, hence both are state prices rather than survey responses, and the law of one price applies to them in full. This is the proposition the thesis ultimately puts to the test, and it is what makes the comparison a question about a single price observed in two places rather than a comparison of two different things.

2.3 Prediction-market literature

A small empirical literature studies these venues directly, and read against the state-pricing framework above it shows both what is settled and what is not. Wolfers and Zitzewitz (2004) survey the evidence, from the Iowa Electronic Markets to within-firm forecasting exchanges, that markets in event-contingent claims aggregate dispersed information into prices that behave as reasonably well-calibrated probabilities and often outperform alternative forecasts such as large polling organizations, though the same authors caution that calibration deteriorates for low-probability events. The essence of that work is to ask whether prediction-market prices are accurate. It largely does not ask whether those prices obey the cross-market pricing discipline of Section 2.2.

Three recent studies push the question toward prices and frictions, and together they mark the boundary this thesis sets out to cross. Saguillo et al. (2025) measure arbitrage within Polymarket itself, both intra-market rebalancing that exploits violations of the $\sum_i p_i = 1$ identity and combinatorial arbitrage across logically related markets, and estimate roughly forty million dollars of realized profit. The summing constraint they trade is the same partition identity that Section 2.1 derives from state pricing. Gebele and Matthes (2026) move one level out, across prediction-market platforms. Aligning more than one hundred thousand events over ten venues, they find about six percent concurrently listed and "persistent execution-aware price deviations of 2–4% on average," which they attribute to semantic non-fungibility, the absence of a shared, machine-verifiable event identity, so that "arbitrage positions cannot be netted across venues and must be held until resolution," producing law-of-one-price violations "driven by structural frictions rather than informational disagreement." Tsang and Yang (2026) examine Polymarket's microstructure during the 2024 U.S. presidential election and show its pricing frictions collapsing as the market matured, with the half-life of arbitrage deviations falling "from several hours [...] to well under a minute" and Kyle's λ, a measure of price impact, dropping from 0.53 to 0.01. The common thread is decisive for positioning the thesis. Every one of these studies remains inside the prediction-market ecosystem, whether on a single platform or among competing prediction-market platforms.

2.4 Price discovery and the limits of arbitrage

Recognizing that two venues must share one state price raises a question the static law of one price cannot answer, namely which venue moves first when the two prices diverge and then reconverge, and why the gap does not close instantly. Naming the leader takes more than comparing how much each price moves. Hasbrouck (1995) supplies the standard instrument. Modelling closely linked prices as noisy views of a single unobservable random-walk efficient price within a vector error-correction system, he defines a market's information share as "the proportional contribution of that market's innovations to the innovation in the common efficient price." The venue that holds its ground while the other adjusts to close the gap is the price-discovery leader, and the venue that chases is the follower.

Shleifer and Vishny (1997) show that real arbitrage is carried out by specialized, capital-constrained agents deploying other people's money, so a mispricing can survive whenever closing it is costly, slow, or risky. Pedersen (2015) names the resulting equilibrium “efficiently inefficient”, meaning prices are left inefficient enough to reward those equipped to trade against them, yet efficient enough to deter everyone who is not. A cross-venue pricing gap can therefore be real and directional and still persist, gated not by disagreement about the event but by the frictions a participant must cross to act on it. The obstruction is structural rather than informational, which is the same barrier Gebele and Matthes document at the cross-platform margin.

2.5 Positioning and the untested crossing

Taken together, these literatures are individually mature but leave one question open. No-arbitrage theory (Harrison and Kreps, Ross, and Bansal and Viswanathan) establishes that a single state price must govern a Polymarket claim and any instrument written on the same event. The empirical prediction-market work (Wolfers and Zitzewitz, Saguillo et al., Gebele and Matthes, and Tsang and Yang) characterizes the accuracy, microstructure, and arbitrage of these venues, but always within the prediction-market universe, on one platform or among competing prediction-market platforms. What none of it does is ask whether the law of one price actually holds across the boundary between a prediction market and the mature instruments of traditional finance that reference the very same events. That crossing is where this thesis is situated. It takes Polymarket prices as the state prices they are and confronts them with the prices of established financial instruments on the same events, asking three questions the existing work tends to run together, namely whether the two venues agree, which of them discovers the common price first, and whether any gap between them survives as a profit once real frictions are paid. The equivalences invoked above are well-established. What has not been done is to measure them on this frontier, where convergence has long been asserted but never tested.

3 Polymarket versus Fed Funds Futures

3.1 Framework & Data

3.1.1 ZQ Fed Funds futures pricing

Two venues price the same event, a Federal Reserve rate decision, but in different units, ZQ as a futures price and Polymarket as a set of outcome probabilities. Comparing them requires first fixing how each instrument is constructed, which is the task of this chapter and the foundation for the spread Δ defined at its close. The CME 30-Day Federal Funds futures contract (ZQ) settles to $100 - \bar{r}$, where $\bar{r}$ denotes the arithmetic average of the daily effective federal funds rate (EFFR) over all calendar days of the delivery month. The minimum price increment is 0.005, corresponding to half a basis point per tick. Pricing is path-dependent within the contract month. A rate change at an intra-month FOMC decision enters $\bar{r}$ only through the days from the meeting forward, while the days preceding the meeting retain the prevailing rate. The realized price is therefore a weighted combination of the pre-FOMC rate r_{pre} and the post-FOMC rate r_{post}, with weights determined by the share of the month falling on either side of the meeting date.

3.1.2 Polymarket FOMC contracts

For each scheduled FOMC meeting, Polymarket lists an event of the form "Fed decision in {month}?" resolved through a set of mutually exclusive binary outcome contracts, a 50+ bp cut, a 25 bp cut, no change, and a 25+ bp hike. Each outcome trades on a continuous order book between 0 and 1 and is interpreted as the market-implied probability p_i that the Federal Open Market Committee delivers outcome i at the meeting in question. Since the outcomes partition the decision space at the granularity at which Polymarket lists them, the prices are constrained to satisfy $\sum_i p_i = 1$.

The Polymarket-implied expected rate change at the meeting, in basis points, is the probability-weighted sum

$$E[\Delta r] = \sum_i p_i \cdot b_i,$$

where $b_i \in \{-50, -25, 0, +25, +50\}$ is the basis-point move associated with outcome i. This expectation forms the input to the implied-ZQ construction that enables the comparison of the two venues, discussed in the next chapter.

3.1.3 Mapping Polymarket to an implied ZQ price

Translating Polymarket probabilities into a ZQ-comparable price requires a path-dependent reconstruction of the average effective rate over the contract month. Let r_{pre} be the prevailing EFFR entering the contract month and let $r_{\text{post}} = r_{\text{pre}} + E[\Delta r]/10{,}000$ be the rate that obtains after the FOMC decision, with $E[\Delta r]$ in basis points as defined in 3.1.2. Denote by D_{pre} and D_{post} the number of calendar days in the contract month falling before and on or after the meeting, and let $D_{\text{total}} = D_{\text{pre}} + D_{\text{post}}$. The Polymarket-implied ZQ price is then

$$\text{ZQ}_{\text{implied}} = 100 - \frac{D_{\text{pre}} \cdot r_{\text{pre}} + D_{\text{post}} \cdot r_{\text{post}}}{D_{\text{total}}},$$

which mirrors the contract's own settlement convention from 3.1.1 and recovers the realized ZQ price exactly under the counterfactual that the FOMC delivers the Polymarket-implied expected move and EFFR is otherwise constant.

The spread between the two venues is the central object of the analysis. Define

$$\Delta = \text{ZQ}_{\text{implied}} - \text{ZQ}_{\text{actual}}$$

expressed in basis points by multiplying by 100 (so that one ZQ tick of 0.005 corresponds to 0.5 bps of Δ). A positive Δ indicates that Polymarket prices a more dovish path than ZQ, and a negative Δ the converse. All cross-market results in the chapters that follow are framed in terms of Δ, with Polymarket-implied and observed ZQ entering only as constituent legs.

3.2 Accuracy and Convergence Dynamics

3.2.1 Sample composition

Section 3.1 defined the spread, this section measures how it behaves across the panel. The merged panel pairs minute-resolution ZQ futures prices with the synchronous Polymarket-implied ZQ price for each scheduled meeting. The futures leg is sourced from Databento, which stores the CME Globex data, and the Polymarket leg from the historical pricing available through the Polymarket API. It covers fourteen FOMC meetings from December 2024 through July 2026 and 1,366,836 matched minute-resolution observations, of roughly 2.39 million total panel rows. The difference is the minutes when the futures venue is closed and no matched spread exists. Days-to-meeting (DTM) range from 1 to 189 (Table 1). Figure Figure 1 plots the two legs and Δ over a single contract's life and previews what this chapter is about. The legs track closely, Δ is wide far from the meeting and collapses to near zero at it, and it steps down as each intervening FOMC resolves. Per-meeting Δ statistics show two patterns. Dispersion grows with contract horizon, and the unconditional mean of Δ varies in sign across the panel, dovish ($\Delta > 0$) on the most recent 2026 contracts and hawkish or near-zero on several 2025 contracts.

3.2.2 Term structure of $|\Delta|$

How far apart the two venues sit should depend on how much of the rate path is still unresolved, so the chapter's first question is how the spread scales with the horizon to the meeting. The minute-resolution panel is aggregated to (meeting, DTM) cells, retaining the cell mean of $|\Delta|$ and its observation count $n_{m,d}$. Cells with zero $|\Delta|$ or non-positive DTM are dropped as undefined under the log specification. The regression is

$$\log |\Delta|_{m,d} = a + \beta \cdot \log d + \sum_k \gamma_k \cdot \mathbf{1}[m = k] + e_{m,d},$$

where $|\Delta|_{m,d}$ is the cell mean spread for meeting m at days-to-meeting d and the meeting fixed effects γ_k absorb level differences across contracts. Standard errors are HAC with a ten-day lag, accommodating residual autocorrelation in the daily aggregates.

The estimated slope is $\beta = 0.87$, with a 95% confidence interval of $[0.73, 1.01]$. The specification fits well for one regressor and a set of fixed effects, with $R^2 = 0.50$ across 1,400 (meeting $\times$ DTM) cells. What matters is where that exponent sits. It is close to unit (linear-in-time) scaling and well above the $\beta = 0.5$ that would obtain if the spread accumulated as the standard deviation of an i.i.d. diffusive process. Only the December 2024, December 2025, and March 2026 fixed effects are individually distinguishable from the baseline, so the scaling is a feature of the panel as a whole rather than of one contract. In levels this reads as $|\Delta| \approx d^{0.87}$, as dispersion widens roughly in proportion to the horizon.

3.2.3 Sign bias

The term structure measures how wide the spread is. Its sign asks a separate question, namely which venue prices the more dovish path and whether that leaning is reliable. A one-sample t-test of Δ against zero, run separately for each meeting, rejects a zero mean at $p < 0.001$ for thirteen of the fourteen contracts, with January 2026 the lone exception (Table 1). The per-meeting lean is therefore real, but its direction is not. The 2026 contracts price a reliably more dovish path than the futures, averaging between +1.09 and +2.07 bps, while several 2025 contracts run the other way, with September and October 2025 averaging close to −0.9 bps. Because Δ is Polymarket-implied ZQ minus observed ZQ, this says Polymarket leans dovish on the recent contracts and hawkish across parts of 2025, a per-contract bias that is statistically reliable but whose sign the panel itself does not pin down.

3.2.4 The prior FOMC as the engine of the spread

The term structure of 3.2.2 establishes that the spread widens with the horizon to the meeting but does not say why. The reason is that the horizon is not empty. Between any observation and the contract's settlement sits a stack of as-yet-unresolved FOMC meetings, and each one carries its own cross-venue disagreement about what the Committee will do. As each intervening meeting resolves, the disagreement it carried collapses and the spread steps down. A long horizon is wide not through abstract diffusion but because it still contains many unresolved meetings, so the scaling of $|\Delta|$ with days-to-meeting is the accumulation of these not-yet-collapsed disagreements.

A before/after test makes the mechanism explicit. Splitting each meeting's observation window at the nearest preceding FOMC, which we call the prior meeting, $|\Delta|$ is on average roughly five times larger before that prior meeting resolves than after it. The effect is strongest on July 2025, where the spread is 12.1× wider beforehand, and it clusters near sevenfold on the three 2026 contracts (6.4 to 7.2×). Only May 2025, at 0.8×, runs against the pattern. Every one of these differences is significant at $p < 0.001$ (Table 1).

The same mechanism is visible in event time on a single impulse. The 18 March 2026 FOMC decision repriced the April, June, and July 2026 contracts at once, and the post-event spread scales cleanly with contract horizon, peaking at 1.5 bps on April, 5.4

bps on June, and 7.9 bps on July (Figure 2). The cross-sectional ordering of the panel thus reappears from a single FOMC shock rather than from averaging, and the farther a contract sits from its own settlement, the more unresolved meetings the shock leaves in front of it and the wider its spread opens.

What opens at the announcement then closes slowly. On every contract the spread expands almost instantaneously when the decision lands and decays over the following days. An exponential fit to that post-resolution decay is estimable on ten meetings. The half-lives lie between 9 and 26 days, with a median near 17, and the April 2026 contract decays the fastest at 9.2 days (Table 1). Two to three weeks is therefore the representative horizon over which Polymarket digests a resolved meeting. The contrast with the futures venue is stark. ZQ absorbs the same decision within minutes, as the path-dependent settlement formula of 3.1.1 maps the realized rate into the contract price the moment the post-FOMC days enter the monthly average. The spread is thus the joint product of two adjustment speeds operating on radically different timescales, with Polymarket the slower-adjusting leg, a reading that Section 3.3 confirms directly through the price-discovery decomposition.

3.3 Price discovery between Polymarket and ZQ

3.3.1 Why naive measures mislead

One venue absorbs information in minutes and the other over weeks, and showing which of them leads requires identifying the leader directly, which is the work of this section. Two natural shortcuts suggest themselves, and both point the wrong way.

The first is to read leadership from how much each price moves. The standard deviation of day-over-day changes in Polymarket-implied ZQ exceeds that of observed ZQ in ten of the fourteen meetings, with a median ratio $\sigma(\Delta \text{ZQ})/\sigma(\Delta \text{Poly})$ of 0.55. Polymarket's continuous probability revisions and one-cent quantization generate more daily price variance than the tightly quoted futures, not less, so here the more variable venue is the less informative one. Raw variance does not establish leadership and, if anything, flags the noisier venue.

The second is to classify the days on which Δ itself moves most and ask which leg drove each one. Let $|\Delta\Delta_t| = |\Delta\Delta\text{Poly}_t - \Delta\Delta\text{ZQ}_t|$ be the size of the day's move in Δ. Each day in the top decile of $|\Delta\Delta_t|$ is attributed to whichever leg dominates that move. A day is ZQ-driven if $|\Delta\text{ZQ}| > 2|\Delta\text{Poly}|$, Poly-driven if the reverse, and joint otherwise. Across the 140 top-decile days the split is 41% Poly-driven, 37% ZQ-driven, and 22% joint. Neither venue dominates the tail of large Δ moves. Raw variance and leg attribution both fail to name a leader, which is exactly why the rest of the chapter turns to the Hasbrouck information-share decomposition.

3.3.2 Hasbrouck information shares

Neither raw variance nor leg attribution names a leader, so the question calls for an estimator designed to answer it. The Hasbrouck (1995) information share, introduced in the Literature Review, is estimated from a vector error-correction model that asks which venue's price does the work of closing the spread when the two legs drift apart. The venue that moves to close the spread is the follower, and the venue that holds its ground while the other adjusts is the price-discovery leader. Computing this share cleanly on the present data takes care on three points, each a feature of the data rather than a refinement of the theory.

A vector error-correction model normally has to estimate the long-run relationship that ties two prices together. Here that relationship is known in advance. Because both legs are views of the same rate decision, no-arbitrage requires them to move one-for-one, so their long-run difference is simply the spread Δ and there is nothing to estimate. We impose this one-for-one link rather than fitting it, which makes Δ itself the error-correction term the model watches, the quantity each leg is pulled back toward when the two drift apart. The choice is both economically correct and practically necessary. Left to be estimated freely, the long-run relationship is unstable on a sample as short as a single contract's daily history, whereas imposing it reduces the model to two ordinary regressions of each leg's change on the lagged spread, which stay well-behaved (Table 2).

Polymarket also trades thinly. On fifteen-minute bars only 7 to 10% carry a genuine update on both legs, so sampling on a fixed clock feeds the model long stretches in which nothing happened, and those empty increments bias the decomposition. We therefore sample in event time, keeping a bar only when one of the legs actually moves. Stripping out the dead intervals has a second payoff. The share then comes out as a single sharp number rather than a wide range.

Finally, the whole construction rests on the spread being mean-reverting, so we test that assumption directly. An augmented Dickey-Fuller test on Δ is run for each meeting, and a confidence interval is attached to every share by a moving-block bootstrap of the event-time estimate (Table 2, Figure 3). All shares are estimated over ZQ-open hours only. The roughly thirty percent of time when the futures venue is closed, overnight and at weekends, is necessarily excluded.

3.3.3 The leading venue is ZQ

The decomposition is decisive and one-sided. In event time the Polymarket information share sits below 0.05 for most meetings across 2025 and 2026, reading 0.008 on December 2025 and June 2026, 0.031 on October 2025, and 0.039 on April 2026. For nine of the thirteen meetings whose spread is stationary, the bootstrap confidence interval lies entirely below 0.5, which makes ZQ the leader at conventional significance (Table 2, Figure 3).

The remaining four, March, May, June, and July 2025, carry wide intervals that straddle 0.5 and cannot be called in either direction. March 2025 is typical of the four. Its point estimate of 0.04 sits firmly on the ZQ-led side, but the 95% interval $[0.001, 0.76]$ straddles 0.5 and so cannot rule out the opposite reading.

December 2024 falls outside this count because its spread is not stationary, so no information share is defined for it. July 2026 is only borderline stationary and is read with more caution than the rest.

A robustness check concerns the approach to the meeting. As the decision becomes near-certain the spread's variance collapses, because ZQ's price freezes once the outcome is fully priced. A regime split at the prior FOMC isolates the affected window. The pre-decision window holds the bulk of the observations and the genuine two-sided price discovery, and it is ZQ-led on every meeting, matching the whole-window estimate. The only sub-windows that read Polymarket-led are exactly those in which the spread has stopped mean-reverting (Table 2). A near-meeting reading would therefore overstate Polymarket's role, and the full-window decomposition used here avoids it. The conclusion is unambiguous and survives every check. ZQ is the price-discovery venue, and Polymarket follows.

3.3.4 Synthesis: the role of each market in price discovery

The decomposition overturns the variance-based intuition. ZQ is the price-discovery venue across the panel. It impounds the common efficient price first, and Polymarket is a near-pure follower whose information share is small and, across the cross-section of meetings, has fallen over 2025 into 2026, with a Spearman correlation of -0.42 between the share and the meeting date. That decline is genuine rather than mechanical, since the share is essentially uncorrelated with the day-count ratio $D_{\text{total}}/D_{\text{post}}$ that governs how a given price move maps into rate space (Table 2). The earlier readings missed this. Raw price variability favored Polymarket because that venue quotes more noisily, and the leg-by-leg attribution split the large moves almost evenly, so each in isolation would have named the wrong leader. Only the event-time decomposition, once the one-for-one restriction, the bootstrap, and the frequency and regime checks are in place, recovers the correct ordering. The descriptive findings of Section 3.2 fall into line with it. The spread widens roughly in proportion to days-to-meeting and decays after each FOMC on a multi-day half-life, which is precisely the signature of one venue absorbing information within minutes while the other digests it over weeks.

3.4 The convergence trade

3.4.1 What the information share implies for a trader

The Hasbrouck result is not merely descriptive. If ZQ leads and Polymarket follows, then the spread Δ should close predominantly through Polymarket moving toward ZQ, and the basis on a single contract should itself mean-revert. It does. The hourly AR(1) coefficient on Δ has a median of 0.88, and an augmented Dickey-Fuller test rejects a unit root in twelve of the fourteen meetings, consistent with the per-meeting stationarity of 3.3.2. One execution fact shapes what follows. Because ZQ leads and is liquid, with a sub-basis-point tick and spread, only the Polymarket leg need be crossed to harvest the convergence. How costly that crossing is, and whether the edge survives it, is the subject of the final chapter.

3.4.2 The predictive asymmetry

The finding that ZQ leads and Polymarket follows carries a directional implication, and it can be tested directly. For each meeting the one-hour-ahead return of each leg is regressed on the demeaned spread Δ. If ZQ leads, the spread should predict Polymarket's subsequent return but not ZQ's, and it should do so with a negative loading, since the spread closes by Polymarket moving toward ZQ (Figure 4). The Polymarket loading is negative and significant on every one of the fourteen meetings, with Newey-West t-statistics between -2.4 and -8.5. The pooled estimate is -0.26 ($t = -5.7$), which implies that roughly 26% of any spread closes through Polymarket within the hour. The ZQ loading is an order of magnitude smaller and statistically flat, pooling to $+0.02$. Polymarket converges to ZQ, and ZQ does not revert to Polymarket. This sign asymmetry restates the price-discovery result as out-of-sample predictability, and it is the cleanest single verification of the cross-venue convergence this section sets out to establish. The venue the information-share decomposition labels the follower is precisely the venue whose returns the spread predicts.

3.4.3 Construction

If the spread predicts its own closing, the natural next step is to trade it. On five-minute bars restricted to ZQ-open hours, the spread is standardized by a trailing one-day rolling mean and standard deviation. A position is opened against the spread when the z-score exceeds 1.5 in absolute value, shorting the spread when Polymarket is rich to ZQ and going long when it is cheap, and closed when the z-score reverts below 0.5 or a one-day holding cap is reached. All statistics are trailing, so the rule is walk-forward with no look-ahead. Profit and loss is the change in the spread in basis points, decomposed into a Polymarket-leg and a ZQ-leg contribution to verify that the convergence is Polymarket-driven.

3.4.4 Gross results

Pooled across the thirteen valid meetings, the strategy takes 4,337 trades and earns an average gross payoff of +0.71 bps per round trip. The edge is overwhelming on the usual tests, with a 94% win rate and a t-statistic of 21.8 (Figure 4). The result is not driven by any single contract. Every meeting is individually profitable gross of cost. Fifty-eight percent of the aggregate profit accrues on the Polymarket leg, and the share is higher still on the meetings the information-share analysis identifies as most strongly ZQ-led. This confirms that the spread closes mainly through Polymarket moving, exactly as the discovery result predicts. The residual ZQ-leg contribution is the unhedged drift of the anchor during the holding period rather than ZQ reverting to Polymarket. The payoff distribution is strongly right-skewed. A 94% win rate of small gains coexists with a thin tail of large convergence profits and a bounded left tail, so the economic content of the alpha sits in the rare large dislocations that the strategy is designed to fade rather than in routine bar-to-bar noise.

3.4.5 Frequency and robustness

Average gross payoff rises from 0.59 bps at fifteen-minute bars to 0.71 at five minutes and 1.15 at one minute, with the t-statistic reaching 31 and the Polymarket-leg share rising to 78%. Gating entries and exits to bars with a genuine (non-forward-filled) Polymarket update leaves the result essentially unchanged, so the edge is not an artifact of trading on stale prices. Raising the entry threshold concentrates the trade in larger spreads and lifts the per-trade pickup, with the break-even cost rising from 0.57 bps at a 1.0σ trigger to 1.34 bps at 3.0σ.

3.5 Cost sensitivity and synthesis

3.5.1 Net P&L and the data constraint

Section 3.4 establishes a gross edge, whether it survives as a net edge turns on cost, and on one cost in particular. The break-even round-trip cost equals the per-trade gross payoff by construction, at 0.71 bps at the baseline trigger and rising to 1.34 bps when only the largest spreads are traded. Two costs bound the net result. The ZQ leg is negligible, with a sub-basis-point tick and spread and low-touch execution, and is not the binding constraint. Commercially available futures data confirms that the futures venue can be traded second-by-second if required. The binding cost is the Polymarket spread, and it enters with a multiplier. Because the FOMC event resolves through several binary outcome contracts rather than a single instrument, replicating the ZQ-implied expected rate is not one trade but a basket position across the two or three contested legs, so the realized round-trip cost is a corresponding multiple of one leg's half-spread. Going to finer sampling does not relax this. The gross edge rises at one-minute bars, but the trade count rises faster, so each additional unit of gross alpha is bought with an additional spread crossing, and the net economics deteriorate rather than improve as the strategy is run faster.

Whether the break-even clears the true cost cannot be settled from the data available, because Polymarket's historical order book is not recoverable. The prices-history endpoint degrades to twelve-hour granularity once a market resolves. The order-book-history endpoint is discontinued. The two routes to an actual number are therefore to reconstruct an effective spread from on-chain exchange fills, which are uncapped and timestamped, or, more cleanly, to record the live bid-ask for future meetings, which a continuous logger can do prospectively. What can be said from the data in hand is that Polymarket's effective tick, measured as the smallest genuine price increment, is small in price-basis-point terms, roughly 0.01 to 0.05 bps, compressed by the same probability-to-price mapping that governs the implied-ZQ construction of 3.1.3. A break-even of 0.71 to 1.34 bps per leg is therefore not mechanically out of reach, but it sits in a plausibly marginal, sign-uncertain zone once the two-to-three-leg multiplier is applied. The honest statement the data supports is a gross-alpha-and-break-even result, not a net-profitability claim.

3.5.2 Closing claim

The cross-platform basis between Polymarket and ZQ on FOMC contracts hosts a real and robust information lead. ZQ discovers the rate path first, Polymarket follows, and the follower's convergence is statistically overwhelming (t-statistics above 20), directly verified by the predictive asymmetry of 3.4.2, and it strengthens rather than decays as sampling frequency rises. Whether the lead is net-harvestable, however, is gated by frictions that this study can bound but not fully observe, namely a multi-leg Polymarket spread crossed on each of several binary outcome contracts.

This is the substantive observation that frames the section within the prediction-market efficiency literature. That literature has largely asked whether prediction-market prices are accurate (whether they forecast outcomes as well as or better than alternative aggregators), and on that test Polymarket performs well, with its expected-rate forecast error falling monotonically to roughly 0.4 bps in the final hour before a decision. The result here is orthogonal to accuracy and concerns the dynamics of information arrival. A venue can be well-calibrated to the eventual outcome and still systematically lag another venue in incorporating news, and the lag is what a price-discovery analysis isolates. Forecast accuracy and price-discovery leadership are distinct properties, and Polymarket exhibits the first while ceding the second to the futures market.

The market is, in the precise sense the evidence supports, efficiently inefficient. A measurable and directionally exploitable pricing lead exists, but the microstructural friction that would have to be crossed to arbitrage it away, namely wide, multi-leg, retail prediction-market spreads, is the most plausible reason it persists. The obstruction is not informational but structural. This is a sharper reading than a conventional efficiency test would license. A no-free-lunch test applied to the retail participant the venue is designed to serve would classify the market as efficient, because the lead does not clear that participant's costs. A venue can host genuine, leading pricing information that its own user base cannot harvest net of the spreads it must cross, and the Polymarket-versus-ZQ basis on FOMC contracts over the December 2024 to July 2026 panel is, on the evidence assembled here, an instance of exactly this case. Settling whether the lead is harvestable by a participant who can post passively into the Polymarket book rather than cross it, the one execution profile under which the break-even estimates might be cleared, requires the live bid-ask record that this study could not reconstruct retrospectively, and is the natural extension of the work.

4 Polymarket versus CME Bitcoin Options

The Fed case study established that a Polymarket contract and a mature financial instrument written on the same event can be compared cleanly when both resolve against a settlement-grade reference. This chapter applies the same programme to a second pairing. Polymarket lists daily markets of the form *"Bitcoin above K on {date}?"*, and the Chicago Mercantile Exchange lists options on Bitcoin futures whose strikes bracket those same thresholds and whose settlement is fixed by the same reference rate at the same instant. The two venues therefore price the same contingent claim, the event that Bitcoin closes above a stated level on a stated day, and no-arbitrage requires them to assign it one price. As in the Fed chapter we ask whether they agree, which venue discovers the price first, and whether any spread between them survives as a profit once real frictions are paid.

4.1 Framework & Data

4.1.1 The Polymarket Bitcoin contract as a digital claim

The Polymarket contract is, by construction, an Arrow-Debreu security in the sense of Section 2.1. The probability the platform attaches to the security ending in the required state is the price one pays today for a payoff of one dollar in that future state. Polymarket lists more than ten strikes for each resolution date, each trading as a separate market with its own quoted probability between zero and one. A participant who buys a "YES" share at 40 cents on the market *"Bitcoin above 80,000 on May 10?"* pays 40 cents for a contract that pays one dollar if Bitcoin settles at or above 80,000 and nothing otherwise, a net result of +60 cents or −40 cents. The 40-cent price is therefore the market-implied probability that Bitcoin ends above that strike. Read across strikes for a single date, the set of YES prices traces a discretised risk-neutral distribution of the settlement price, and each contract has time-t price equal to $e^{-r(T-t)} \cdot Q_{\mathrm{PM}}(S_T \geq K)$.

4.1.2 The CME counterpart: $N(d_2)$ and the call-spread replication

If the Polymarket contract were priced correctly, its quoted probability would equal the risk-neutral probability of the underlying ending in the money. The Black-Scholes (1973) formula,

$$c = S_0\, N(d_1) - K\, e^{-rT}\, N(d_2),$$

contains within it the object the comparison requires. The term $N(d_2)$ is exactly the risk-neutral probability that the underlying finishes in the money at expiry, $Q(S_T \geq K)$. The original Black and Scholes (1973) derivation proceeds via a hedging argument rather than an explicit change of measure, but the risk-neutral interpretation is established by Cox and Ross (1976). They show that European option prices admit a representation as discounted expectations under a probability measure constructed so that the discounted underlying is a martingale, and in that representation $N(d_2)$ appears as the no-arbitrage probability of exercise. Breeden and Litzenberger (1978) make the link explicit, showing in their equation (6) that the price of a digital security paying one dollar if $S_T \geq Y$ is $G(Y, T) = B(T) \cdot N[d_2(X = Y)]$, the discount bond multiplied by the risk-neutral exceedance probability. The same risk-neutral reading of $N(d_2)$ is developed in discrete time by Rubinstein (1976) and Cox, Ross and Rubinstein (1979), extended by Merton (1973), and treated as standard in Hull (2021), Björk (2009) and Shreve (2004). A Polymarket binary on the event $\{S_T \geq K\}$ thus has, by construction, the same time-t price as the corresponding CME digital, and the law of one price requires the two to coincide.

What turns this identity into something measurable is a tradable security replicating the digital payoff, since $N(d_2)$ is a quantity inside a formula rather than an instrument one can buy. A tight call spread does this. Buy a call at a lower strike K_1 and sell a call at a higher strike K_2. As the two strikes are taken infinitesimally close, the spread payoff converges to the step function and replicates the Polymarket "YES" payoff. A tight put spread replicates the "NO" share. The width-normalised call spread $[C(K_1) - C(K_2)]/(K_2 - K_1)$ is the finite-difference estimator of $-\partial C/\partial K$, which is the digital price, and in the limit it equals $e^{-r(T-t)}\, N(d_2)$. Under Arrow-Debreu and the law of one price, the call spread and the Polymarket "YES" share value the same future state of the world and must therefore trade at the same price. The empirical question of the rest of the chapter is whether the spread between the two probabilities is reliably non-zero, and if so whether a strategy can profit from it.

Bringing the two venues into the same units requires bracketing every Polymarket

market by the nearest pair of CME call strikes that straddle the hurdle. For each Polymarket market identified by a resolution date and a hurdle price K, the bracketing pair of CME call strikes $K_L < K < K_U$ on the matching expiry is located, and the CME-implied probability at the hurdle itself is the width-normalised call spread with a midpoint correction,

$$\text{CME\_prob} = \frac{C(K_L) - C(K_U)}{K_U - K_L} + \frac{m - K}{K_U - K_L},$$

clipped to $[0, 1]$, where $m = (K_L + K_U)/2$ is the bracket midpoint. The first term is the raw width-normalised call spread, and the second linearly relocates the digital from the bracket midpoint to the hurdle, eliminating the leading bias the raw form carries when K does not sit exactly between K_L and K_U. Because the call-spread digital has a small slope across the strike range rather than the step shape of the Polymarket payoff, a trade taken against a real spread can still lose if Bitcoin settles inside the bracket, even in the absence of any actual spread at the hurdle itself (Figure 5). With this convention in place, the central object of the chapter is the cross-venue spread, the analogue of the Fed chapter's Δ, expressed in percentage points as

$$\Delta = \text{Poly\_YES} - \text{CME\_prob},$$

so that a positive Δ indicates Polymarket prices the "YES" outcome richer than the CME-implied digital. All cross-market results in the sections that follow are framed in terms of Δ, with the two underlying probabilities entering only as constituent legs.

4.1.3 Real data: the CME daily-resolution option chain

Because Polymarket operates daily-resolution markets across a range of strikes, executing the matching trades requires options with daily expirations. The CME lists weekly options on the Bitcoin futures contract expiring on each business day from Monday through Friday, alongside monthly contracts settling on the last Friday of each month and three deferred Friday weeklies. Together these products constitute a daily-resolution option chain consistent with the settlement schedule of the Polymarket markets. Each CME contract settles to the CME CF Bitcoin Reference Rate (BRR), fixed at 16:00 London time, and the Polymarket market for the same date resolves against the same reference. That fixing does not fall at a constant UTC hour. Under British Summer Time it lands at 15:00 UTC and under Greenwich Mean Time at 16:00 UTC, so the settlement instant shifts by an hour between summer and winter. When there is a maturity mismatch, the CME leg is closed at the option's official settlement price and the Polymarket leg is marked at its last quote at or before the BRR timestamp. This matched settlement removes any horizon or theta mismatch.

4.2 Accuracy and Convergence Dynamics

4.2.1 Sample composition

The merged panel pairs minute-resolution CME quotes with the synchronous Polymarket history for every daily Bitcoin market. It covers 1,332 daily *"Bitcoin above K"* markets, spanning September 2024 through May 2026, with distance-to-expiry running from the settlement minute out to roughly 188 hours. From this panel a trade is constructed for every market in which the spread clears the entry threshold of section 4.4.1, yielding 1,303 settlement-exit trades. The CME leg is sourced from Databento's GLBX.MDP3 feed (the bbo-1m schema for quotes and the statistics schema for settlements) and the Polymarket leg from the Gamma and CLOB historical endpoints.

4.2.2 The spread is reliably different from zero

Section 4.1 defined Δ, and this section measures how reliably it differs from zero. A per-market one-sample t-test of the signed Δ rejects a zero mean for 1,256 of 1,332 markets with a pooled signed mean of +4.14 percentage points. In all but a handful of markets the two venues sit on reliably different probabilities, and on average Polymarket sits slightly rich. Because the signed Δ can be traded from either side, a rich Polymarket share is shorted and a cheap one bought. The sign itself therefore carries no economic information, and the tradeable quantity is the magnitude $|\Delta|$. Restricting the test to markets with only a small midpoint correction in the sense of section 4.1.2 leaves the picture unchanged (1,240 of 1,297 markets significant, pooled mean +4.34 points), so the spread is not an artefact of the bracketing reconstruction.

4.2.3 Term structure of $|\Delta|$

How far apart the two venues sit should depend on how much of the move to settlement is still unresolved. Aggregating the panel into (market, distance-to-expiry) cells and regressing $\log|\Delta|$ on log distance-to-expiry with market fixed effects and HAC errors yields a slope of +0.47 on the full sample (95% CI [0.45, 0.49], $R^2 = 0.65$, 71,238 cells). The spread widens with horizon at almost exactly the +0.5 rate that a diffusive i.i.d. accumulation would produce, and well short of the near-linear +0.87 of the Fed contracts. Restricting to tight brackets lowers the slope to +0.36, and restricting to near-the-money markets flips it negative (-0.17). The Bitcoin term structure is therefore not the clean linear-in-time accumulation seen in the Fed chapter, which is expected, since the Bitcoin horizon contains no stack of discrete, individually repriced prior events to collapse one by one. Averaged across markets, the spread collapses to zero as maturity approaches (Figure 6).

4.2.4 Drivers of the spread and its sign

A pooled regression of $\log |\Delta|$ on candidate drivers (HAC errors) identifies what widens the spread. Moneyness enters with a large negative coefficient (-3.37, $z = -283$), so the spread is by far the widest near the money, where the digital probability is most sensitive to the underlying. Distance-to-expiry enters positively ($+0.34$, $z = +144$), consistent with the term structure of section 4.2.3. Polymarket activity, measured as the share of distinct quotes in a market, enters strongly positively ($+5.11$, $z = +88$), meaning the spread is widest when the Polymarket book is active and repricing. As for direction, although the per-market test of section 4.2.2 rejects a zero mean for 96% of markets, the mean Δ is positive in only 806 of 1,313 markets, about 61%. Polymarket leans rich more often than not, but not so reliably that direction can be treated as known. As in the Fed chapter, the per-market lean is statistically real while its sign is, for the panel as a whole, only weakly determined, and the spread itself is fundamentally a magnitude phenomenon.

4.3 Price discovery between Polymarket and CME

The Fed chapter could name a price-discovery leader cleanly. ZQ moved first and Polymarket followed. This section runs the same two diagnostics on the BTC pair, shows why each is contaminated by the call-spread reconstruction, and states plainly that the leadership question is, on the current construction, unresolved.

The standard estimator is the Hasbrouck (1995) information share, computed here from a vector error-correction model on the two legs with the no-arbitrage one-for-one long-run restriction imposed. The decomposition is estimable on 1,273 markets and, taken at face value, says Polymarket leads. The Polymarket information share averages 0.68 with a median of 0.87, and Polymarket is the leading leg in 929 of 1,273 markets, the opposite of the Fed verdict.

The estimate, however, inherits a feature of how the CME side is built. The CME-implied probability is not a single tradable price. It is a width-normalised digital reconstructed from a long-short call spread, and the bracket geometry feeds into every price change the model observes on that leg. When the underlying drifts across a fixed bracket the reconstructed probability slides along the slope of the call price between K_L and K_U rather than tracking a single tradable instrument, and that slide enters the vector error-correction model as a CME-side price move the venue did not actually make. Three panel-wide correlations identify the share as a measure of this reconstruction rather than of price discovery. The Polymarket information share correlates $+0.69$ with the inverse call-spread width $1/W$, $+0.47$ with the amplitude of the CME-side probability across the panel, and $+0.38$ with the CME calibration error against the realised outcome. Each of those three quantities is a property of the call-spread reconstruction rather than of either venue's information flow, so a genuine information share should be independent of all three. The bootstrap confidence interval also fails to put either venue cleanly on one side of 0.5, so the estimate is wide as well as biased.

An estimator suspected of contamination calls for an independent check, and the natural one is the predictive-asymmetry test the Fed chapter ran in 3.4.2. The follower is the leg whose one-hour-ahead return loads negatively on the demeaned Δ, since the spread closes by the follower moving toward the leader. On the Bitcoin panel the Polymarket leg loads -0.04 (Newey-West $t = -30$) and the CME leg loads $+0.19$ ($t = +58$). By the sign rule the negative-loading leg, Polymarket, is the follower, which would make CME the leader and reverse the Hasbrouck verdict above. The two diagnostics therefore reach opposite conclusions on the same panel, and the large positive CME loading is itself implausible for a genuine leader.

A disagreement of this kind between two estimators is what reconstruction noise produces. The Hasbrouck share absorbs the geometric slide of the bracket into the CME leg's apparent moves and reads it as CME-side discovery contributed to the spread, while the predictive-asymmetry regression reads the same slide as a CME-side trend that loads positively on the spread. Neither estimator delivers a trustworthy verdict on the BTC pair, and leadership is therefore left open. The rest of the chapter argues from the spread itself rather than from the venue that closes it.

4.4 The overpricing trade

4.4.1 Construction

The Fed chapter identified Polymarket as the follower and built its trade around that fact, fading the spread inside a contract's life and letting Polymarket close it. The Bitcoin pair offers no such anchor. Section 4.3 leaves leadership open, so the trade cannot be aimed at the venue that will close the spread. The chapter's other findings are nonetheless enough to build a trade. The spread is reliably non-zero on 1,256 of 1,332 markets, and both legs are forced to a single number at the BRR fixing. A position taken against the spread and held to settlement therefore depends only on the spread being non-zero at entry, since the convergence at settlement is mechanical.

For each daily market the book is scanned and a position is opened at the first minute the spread is larger than 2.5 percentage points in absolute value, provided the Polymarket mid sits in the interval $(0.005, 0.995)$. The narrow band keeps the trade off the boundary values 0 and 1, at which the Polymarket price is not actually tradeable and a naive read would produce spurious profits. When Polymarket is rich to the CME digital, the NO share is bought and the call spread bought, and when Polymarket is cheap the opposite position is taken. The position is held to the CME settlement timestamp, at which both legs resolve onto the same realised outcome. The realised PnL is the sum of a Polymarket-leg and a CME-leg contribution, marked at entry on the bid-ask midpoint of each leg and again at the settlement fixing. Applied across the panel, the rule generates 1,303 trades, held to settlement over a mean horizon of 130 hours.

4.4.2 Gross results

Pooled across the 1,303 trades the strategy earns a total gross profit of \$177.70, an average of \$0.136 per trade at a per-trade Sharpe of 0.39, with a 79.3% win rate (Table 3). Because trades cluster within expiry days, the naive trade-level t-statistic of +14.1 overstates significance, and clustering by the 221 distinct expiry days still gives $t = +10.7$. The profit beats 100% of random-side placebo assignments and rises monotonically with the size of the entry spread (Spearman $\rho = +0.55$), the dose-response pattern expected if the spread is the source of the edge.

The most natural alternative reading of this headline is that the strategy is a disguised directional Bitcoin bet. The entry Δ correlates only -0.14 with the realised outcome (Figure 7), and a direction-robust test sharpens the point. On the 529 short-Polymarket trades where Bitcoin moved against the position the strategy still earns a profit (Figure 8). The per-trade PnL plotted against the entry Δ forms a well-defined V-shape, which is what a side-flipping rule must produce. Short trades earn more on more negative spreads and long trades earn more on more positive ones (Figure 9).

4.4.3 Anatomy of the edge

The profit is concentrated along three dimensions, and the location of each is as informative as the headline total. The first is the side of the trade. The 674 trades that short an overpriced Polymarket YES against the call spread earn +\$205.95 (mean \$0.306), while the 629 trades on the opposite side, buying a cheap Polymarket "YES", lose \$28.26 between them. The strategy is, in substance, a short-overpricing trade. The second is the size of the entry spread. Trades opened on a spread above 15 points account for \$159.5 of the \$177.7 total, roughly 90% of the profit, and the four buckets below 15 points contribute the rest. The third is the rank of the trade. The top 5% of trades earn 32% of the profit, and the top quartile earns slightly more than the entire total, so the median trade is close to break-even and the strategy lives in its tail.

The economic content of the edge concentrates further inside one identifiable subset of markets. In the band where Polymarket prices the YES outcome near a coin-flip, with a YES price between 0.4 and 0.5, the platform assigns an average probability of 49% to outcomes that realise only 19% of the time, a 30-point overpricing that the CME digital does not share. This bucket of 391 trades converts the mis-calibration into a mean \$0.286 per trade. The pattern is consistent with retail Bitcoin sentiment running consistently bullish relative to the CME-implied probability path, but isolating that channel from the data in hand is beyond the scope of this paper.

4.4.4 Locked edge versus basis

Decomposing each trade clarifies the route of the gains. Two components combine to produce its realised PnL. The locked component is the value of the spread at the instant the position is opened, the entry Δ captured the moment both legs are marked, and it is mechanically independent of anything that happens afterwards because both legs are held in their entry quantities until the BRR fixing. The basis component is everything that moves the two legs differently between entry and settlement. It is the same finite-width residual that contaminated the price-discovery diagnostics of section 4.3, the slide of the CME-implied probability along the slope of the call price as the underlying drifts between K_L and K_U. It behaves empirically like a roughly symmetric lottery around zero, neither systematically helping nor hurting the trade.

The total profit of \$177.70 splits into a locked component of \$186.56 and a basis component of $-$\$8.86. The locked component is the steady, repeatable part. Stripped of basis it returns \$0.143 per trade with a t-statistic of +30 (Figure 10). The basis component, by contrast, is slightly negative on average and yet accounts for 70% of per-trade profit variance. The composition by trade rank makes the split concrete. The top 5% of trades earn a mean \$0.875, roughly two-thirds of which is basis. The bottom 5% lose a mean \$0.589 almost entirely through basis. The middle 90% earn a mean \$0.136 that is essentially all locked edge and almost no basis. The economically meaningful, harvestable claim of the chapter is therefore the locked entry edge of roughly \$0.14 per trade. The basis is risk to be borne rather than alpha, and both the largest gains and the largest losses are basis realisations, concentrated in the 11% of trades on which the two legs still differ on the eventual outcome at the moment of exit.

4.4.5 Trade distribution and analysis

The distribution of the spread is non-Gaussian, and its shape carries the chapter's central mechanism. Measured at entry, $|\Delta|$ is strongly peaked. A tall mode sits within a few points of zero, where the two venues broadly agree, while a long right tail and a distinct secondary cluster of large positive spreads near forty to fifty points pull the mean to +9.6 points against a median of only +2.5. The entry spread is therefore bimodal and one-sided rather than a symmetric scatter around a common price, with most markets close to agreement, a minority severely far apart, and almost all of the severely-apart cases having Polymarket rich. Measured at exit, the same spread looks entirely different. The residual basis that survives to settlement collapses onto a narrow mass centred essentially on zero, with a mean of $-$\$0.007 per trade and roughly symmetric side-lobes near ± 0.5 (Figure 11).

The shape of the PnL distribution is, moreover, stable across the sample. Splitting the per-trade PnL into the year 2025 and the year 2026 and overlaying the two histograms shows them very nearly coincident. Both carry the same tall spike of small positive outcomes just above zero, the same secondary mode of larger gains near +0.45, and the same modest left lobe of losers near −0.45. The two years differ only marginally in level, with a mean of +$0.128 per trade in 2025 against +$0.147 in 2026 and win rates of 78.7% and 80.0%, and not at all in shape. The cumulative profit curve tells the same story in time. It is close to flat for the first ten months of the sample and then turns sharply upward from the middle of 2025, climbing in a steady, near-monotone line to roughly $178 by early 2026, with no single jump carrying the result. The early flat stretch reflects the thin pre-September-2025 markets noted in section 4.2.1, which simply offer little to trade.

4.5 Cost sensitivity and synthesis

4.5.1 Gross and net profit: sizing, execution, and the cost constraint

The profit reported in section 4.4 is computed on standardised conventions, and how far it sits from a net figure depends on costs the present data does not fully measure. Each trade is delta-neutral to Bitcoin by construction. The long CME call spread and the short Polymarket YES share both pay off on the single event $\{\mathrm{BTC} \geq K\}$, so taking them in equal and opposite dollar nominal cancels the directional exposure and leaves only the cross-venue spread. The profit and loss reported in section 4.4 is normalised to one dollar of notional per leg and scales proportionally.

Gross profit and loss is marked on the bid-ask midpoint of each leg, at entry and at the settlement fixing. Using the midpoint rather than an executable price omits the half-spread a taker would actually pay, but on both venues that half-spread is narrow, a few basis points of notional on the liquid CME options and on the Polymarket book for these daily markets, and therefore second-order against a per-trade edge that section 4.4 measures in the tens of basis points. The mid-price convention shifts the level of the result by a small amount and changes neither its sign nor its statistical strength.

A fully costed net figure is not reported. The dominant cost is genuinely hard to pin down, and on the Polymarket leg is not reliably a cost at all. Polymarket charges no commission to a maker. A participant who posts a resting limit order pays nothing when it is filled and is in addition eligible for the platform's daily liquidity rewards, so a passively executed Polymarket leg carries a cost that is zero or, if the order rests close to the mid, negative. Only a taker crossing the book pays anything, and even then the fee on these slow daily markets is small. For most of the in-sample period Polymarket levied no trading fee whatsoever, with modest taker fees introduced to crypto markets only in early 2026. The Polymarket leg, executed with limit orders, therefore does not erode the edge and may add to it, which is why the unhedged short-Polymarket position of section 4.4 already clears costs comfortably.

The CME leg is the genuine cost. It carries a real options bid-ask a hedger must cross, and the executable quote was not yet extracted at the time of writing. The chapter therefore frames its conclusion against a break-even round-trip cost equal to the per-trade payoff of section 4.4.2, roughly \$0.14 of locked edge per dollar of notional. Whether the CME options' true bid-ask fits inside that budget is the one quantity that decides net viability, and measuring it from the bid-ask data already in hand is the work flagged in section 4.5.2.

4.5.2 Limitations and required further work

Three limitations bound this chapter. The first is the cost of the two trading legs, neither of which the present data fully measures. The second is the two-sided distortion the call-spread reconstruction introduces into a finite-width digital. The third is the still-open leadership question of section 4.3. None reflects a defect in the strategy itself. Each maps to a concrete piece of remaining work, and until that work is done the chapter's headline is a gross-and-break-even finding rather than a net one.

The first and most consequential limitation is the cost of the CME leg. The backtest marks both legs at the bid-ask midpoint, which omits the half-spread a hedger must actually cross, and on the CME side that omission is not negligible. The remedy is to rebuild the call spread at executable bid-ask prices from the bbo-1m feed, which would replace the haircut with a measured round-trip cost.

The cost of the Polymarket leg is harder to bound for a structural rather than a methodological reason. As in the Fed chapter, the Polymarket prices-history endpoint returns only a midpoint, so the venue's true bid-ask cannot be recovered after the fact for any market in the sample. The chapter states this constraint explicitly rather than concealing it behind a representative fee. The prospective remedy is a live logger recording the Polymarket bid and ask for forward Bitcoin markets as they trade, which would allow a genuine execution cost, net of the maker rebate a passive participant may instead collect, to be measured for the next cohort of markets even though it cannot be reconstructed for this one.

The second limitation is the reconstruction bias the call-spread digital carries when the bracketing strikes are far apart. Restricting the backtest to bracket widths of at most 200 points, and independently validating the wide-bracket digital against the equivalent CME-implied probability recovered by inverting Black-Scholes on the bracketing strikes for $N(d_2)$ directly, is the work that decides how much of the headline edge is real and how much is reconstruction noise. The $N(d_2)$ inversion would also serve as an independent check that the call-spread construction of section 4.1.2 agrees with the underlying theoretical object.

The third limitation is the price-discovery question left open by section 4.3. The same reconstruction width that biases the headline biases the Hasbrouck information share, and the diagnostic correlations recorded there make that explicit. Two remedies can move the verdict. The first is to restrict the information-share estimate to wide-bracket markets where the reconstruction noise is largest in absolute terms but most consistent across markets. The second is to port the event-time sampling of the Fed chapter, in which a bar is retained only when one of the legs actually moves, suppressing the empty-bar bias that drags the share toward 0.5 on coarse clocks. Both would let section 4.3 reach a verdict on the BTC pair.

4.5.3 Closing claim

Across more than eight million matched minutes and 1,332 daily markets, Polymarket and the CME differ persistently on the probability that Bitcoin closes above a stated level. The signed spread is reliably non-zero in 96% of markets, and the magnitude is the tradeable object. A settlement-exit strategy that fades the spread earns a large and statistically robust gross edge of $177.70 over 1,303 trades at a 79.3% win rate, a day-clustered t-statistic above ten, a bootstrap interval clear of zero, and a direction-robust overpricing floor that holds, concentrated on the side where Polymarket overprices the "YES" outcome. Whether that edge clears the executable cost of the CME leg and the bracket-width reconstruction bias of section 4.5.2 is the open question this chapter passes to the discussion.

5 Discussion

The two case studies of this thesis come from sharply different liquidity regimes on the same Polymarket platform, and the contrast organises everything that follows. On the Federal Reserve markets, where the next FOMC decision is priced as a basket of binary outcome contracts, the venue draws cumulative open interest and per-contract trading volume in the millions of dollars, with tight quotes and continuous order flow through the life of each contract. On the daily Bitcoin markets, where the contract resolves on whether Bitcoin closes above a stated price on a stated day, the same venue draws only a few thousand dollars of trading volume on a typical strike, and a book that absorbs only a small fraction of any serious order before it moves. Between the two settings sits roughly three orders of magnitude in volume, and at the upper extreme of the same venue, in the 2024 United States presidential market documented by Tsang and Yang (2026), the cross-venue pricing frictions briefly collapse to under a minute and Kyle's λ on the Polymarket book falls from 0.53 to 0.01. Pedersen (2015) named the resulting equilibrium "efficiently inefficient," and on Polymarket the equilibrium is not a single point. It is a gradient. The spread between the platform and the matched traditional-finance instrument persists at the level the prevailing book depth and the cost of crossing the venue can sustain, and the two case studies in this thesis observe the gradient at two interior settings on it. The Federal Reserve pairing sits closer to the high-volume end, with a Polymarket book deep enough that the cross-venue spread can be measured cleanly and the price-discovery question can be answered. The Bitcoin pairing sits closer to the low-volume end, with a book thin enough that the spread is wider, more persistent, and reconstructed against a derivatives leg whose own measurement noise contaminates the comparison.

The three questions Section 2.5 left open can now be revisited in turn against the evidence of the two empirical chapters. The first is whether the two venues agree on the price of the same event. They do not. On the Federal Reserve pairing the per-meeting spread is reliably different from zero on thirteen of fourteen contracts at $p < 0.001$, and on the Bitcoin pairing on 1,256 of 1,332 daily markets at the same threshold. The two findings have the same shape. Per pairing the lean is statistically real, but its direction is unstable across the panel, so the panel-level mean carries no information a trader can use. What is reliably non-zero in both cases is the magnitude $|\Delta|$, and that is the object the rest of the thesis treats as tradeable. Read against the Arrow-Debreu framing of Section 2.1, a Polymarket contract is by construction the price today of one dollar in a single future state, and the law of one price requires that the matched derivatives instrument carry the same price for that same state. On the two pairings of this thesis the requirement fails at every plausible significance threshold, at the volume scales the venue sustains on these contracts.

The second question is which venue discovers the common price first when the two diverge. On the Federal Reserve pairing the verdict is one-sided. The Polymarket information share sits at or below 0.05 on every meeting from January 2025 onwards, the bootstrap interval lies entirely below 0.5 on nine of thirteen stationary meetings, and the result survives every robustness check the chapter applies. ZQ moves first and Polymarket follows. On the Bitcoin pairing the same procedure cannot reach a verdict, because the CME-implied digital is reconstructed from a call spread rather than carried by a single tradable instrument, and the resulting noise contaminates the two diagnostics the chapter brings to bear in opposite directions. The asymmetry between the two cases is itself the interesting finding, and it is methodological rather than substantive. The Hasbrouck information share is a reliable price-discovery diagnostic only when both compared prices are themselves single tradable instruments. Where one leg is a basket reconstructed from derivatives, as the CME-implied digital here is, the share absorbs the basket's geometry and cannot be read as a measure of discovery. The price-discovery answer of this thesis therefore splits by construction. The Federal Reserve pair returns a clean ZQ-leads verdict. The Bitcoin pair returns a verdict on the diagnostic itself, namely that it cannot speak when one leg is synthetic, and the leadership question for that pair remains open pending an alternative measurement strategy.

The third question is whether any spread between the two venues survives as a profit once real frictions are paid. The case studies converge on the same shape of answer at sharply different scales. On the Federal Reserve pairing the convergence trade returns 0.71 to 1.34 basis points per round trip at a 94% win rate. On the Bitcoin pairing the settlement-exit strategy earns $177.70 over 1,303 trades at a 79.3% win rate, with a day-clustered t-statistic above ten and a locked entry edge of roughly $0.14 per dollar of notional. The Bitcoin gross is the larger and more striking of the two results, and it earns the weight of those numbers. It beats 100% of random-side placebos and rises monotonically with the entry spread, all on a venue where the reconstruction-side noise makes the comparison itself the harder of the two to make. The friction that nonetheless gates either trade from clearing net is the Polymarket spread, and the route by which it binds reflects the same volume gradient as the price-discovery answer above. On the Federal Reserve pairing the per-leg half-spread is small in price terms, but the cost compounds across the two or three contested binary outcomes that together replicate the ZQ-implied expected rate, so the round-trip cost is several times one leg's half-spread rather than equal to it. On the Bitcoin pairing the per-leg half-spread is itself wide, and the depth of the book limits how large the position can grow before the price moves materially against it. The structural reading is the same on both. The friction binding the spread on Polymarket is a property of the venue's retail microstructure at the volume scale these contracts attract, not a property of the information the venue aggregates. Three limitations bound this reading and are treated at length in section 4.5.2, namely that the Polymarket bid-ask cannot be reconstructed retrospectively so net profitability on both pairings is a break-even rather than a measured net, that the Bitcoin leadership question is left open by the reconstruction-leg contamination of section 4.3, and that both panels span under two years on a venue that grew roughly thirty-fold over the window.

Two framing papers of Section 2.4 sharpen the volume-gradient reading above. Shleifer and Vishny (1997) characterise arbitrage as the work of specialised agents deploying outside capital under a performance-based capital-allocation rule. Their canonical example, a Bund futures contract trading at different prices in London and Frankfurt, requires the arbitrageur's own capital and is exposed to interim margin calls if the spread widens before it converges. The Federal Reserve case study of this thesis is the modern analogue, and the agency mechanism Shleifer and Vishny identify defeats the marginal Polymarket-versus-CME participant for the same reason it defeats the Bund arbitrageur. Their headline result, that performance-based arbitrage is most ineffective when the spread is widest and the arbitrageur is most fully invested, is what the Bitcoin pairing observes in its low-volume regime and what the Federal Reserve pairing escapes only because its book is deep enough to keep the spread small. Pedersen (2015) closes the frame. A market is "efficiently inefficient" when, in his words, "competition among professional investors makes markets almost efficient, but the market remains so inefficient that they are compensated for their costs and risks."

The combined reading separates two properties of a prediction-market venue that the literature has tended to treat as one. Calibration to eventual outcomes, the central success documented by Wolfers and Zitzewitz (2004) and recurrent in the surveys that followed it, is a property of how a venue aggregates information about the world. Obedience to the cross-venue law of one price is a property of its arbitrage relationship to the rest of finance. The present results show the two are independent of each other. Polymarket prices the rate path and Bitcoin's trajectory accurately enough to forecast the eventual realisations, while failing the cross-venue pricing discipline on the same instruments at the volume scales observed here. This separation, more than any single number in the chapters above, is what the thesis adds to the prediction-market literature.

The contribution of this thesis is twofold. Empirically, the cross-boundary law of one price fails between Polymarket and the matched traditional-finance instrument on events that resolve against the same settlement-grade reference, and the size of the failure scales inversely with the volume the Polymarket contract attracts. Methodologically, the thesis offers a reproducible recipe for matching a prediction-market contract to its derivatives counterpart on the same settlement reference, and records the constraint that the Hasbrouck information share is reliable as a price-discovery instrument only when both compared prices are single tradable instruments.

6 Conclusion

This thesis began with three questions about the boundary between Polymarket and the matched traditional-finance instruments written on the same events. On the evidence of the chapters above, each is settled. On both pairings the law of one price fails reliably, and the tradeable object is the magnitude of the spread. The Federal Reserve pair identifies the futures market as the leader and Polymarket as the follower. The Bitcoin pair returns no verdict on the leadership question, only the constraint that the Hasbrouck instrument cannot speak when one of the compared legs is a basket reconstructed from derivatives. A strategy that fades the spread earns a statistically robust gross payoff on both pairings, and on the Bitcoin pair the headline is \$177.70 over 1,303 trades with a day-clustered t-statistic above ten. What that strategy does not do on either pairing is clear the cost of bridging the Polymarket leg at scale, and that cost is the binding constraint of the thesis.

That binding constraint is Polymarket's retail microstructure operating at the volume scales these contracts attract. On the Federal Reserve pair it appears as the multi-leg counter on a basket of contested binary outcomes. On the Bitcoin pair it appears as the depth limit of a single thin book. In both cases the friction is a property of the venue's order book at present scale, not of the information the venue aggregates. The market is, in the precise sense Pedersen (2015) names, efficiently inefficient. The spread is wide enough to be statistically real, and stable enough across two case studies spanning three orders of magnitude in Polymarket volume to be a regime rather than a coincidence.

On the empirical side, the natural next step is the prospective bid-ask logger flagged in section 4.5.2, on the next cohorts of FOMC and Bitcoin markets, which would convert the break-even claim of this thesis into a measured net. The deeper question is one the venue's own growth will resolve. Tsang and Yang (2026) documented Polymarket's cross-venue frictions collapsing to under a minute in the 2024 presidential market, at a volume scale that dwarfs the Federal Reserve and Bitcoin panels of this thesis. The next two years of Polymarket data will decide whether the venue carries those contracts up its own volume gradient and closes the spread, or stabilises them at the efficiently inefficient equilibrium documented here. Polymarket is, on the evidence assembled, an accurate forecaster of the world's events and an inefficient pricer of the financial claims written on them. Whether the venue ever assigns a single state price to both readings of the same future state is the question its next phase of growth will answer.

Annex

Table 1: Per-meeting spread summary. N counts matched minute observations and "Days" the tracking window. $\overline{\Delta}$ is the mean spread in basis points, $\%\,\Delta > 0$ the share of observations above zero, and t the one-sample test of $\overline{\Delta} = 0$. "Ratio" is mean $|\Delta|$ before the prior FOMC divided by mean $|\Delta|$ after, and "Half-life" is the post-prior-FOMC exponential decay half-life in days. n/a marks contracts with no prior-FOMC split (December 2024, July 2026) or no estimable decay.

Meeting	N	Days	$\overline{\Delta}$ (bps)	$\%\,\Delta > 0$	t	Ratio	Half-life (d)
December 2024	70,686	67	−12.10	46.1	−252.5	n/a	n/a
January 2025	63,451	69	+0.89	79.4	137.1	10.0	15.7
March 2025	64,717	77	+0.07	58.9	20.5	2.0	24.3
May 2025	93,188	89	−0.37	32.6	−68.2	0.8	16.5
June 2025	107,409	109	−0.54	43.0	−86.5	4.3	17.9
July 2025	110,381	112	−0.98	49.8	−108.2	12.1	18.3
September 2025	112,320	115	−0.92	23.1	−264.4	2.3	19.4
October 2025	113,251	113	−0.90	36.7	−155.1	3.0	n/a
December 2025	113,727	114	−0.35	51.2	−47.4	1.2	15.0
January 2026	107,939	114	0.00	53.3	−0.1	4.6	26.4
March 2026	110,685	121	+1.41	58.7	181.0	6.4	13.6
April 2026	132,528	142	+1.09	65.7	158.6	7.2	9.2
June 2026	123,776	127	+2.07	84.1	322.9	7.1	n/a
July 2026	42,778	43	+1.42	73.4	121.8	n/a	n/a

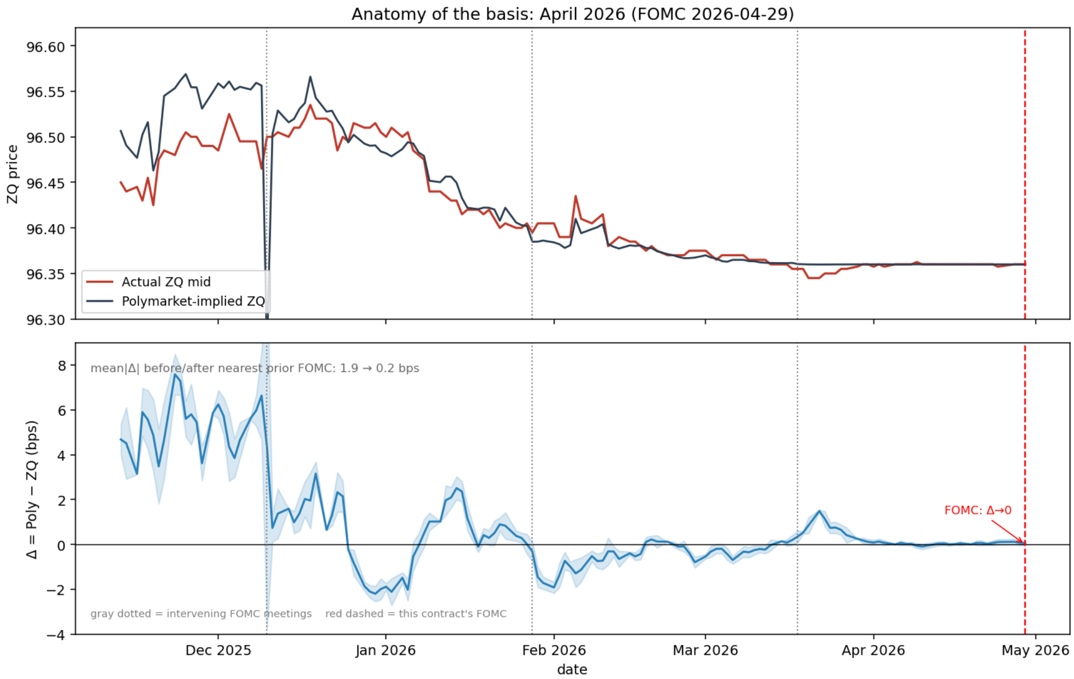

Figure 1: Anatomy of the basis. Polymarket-implied ZQ and observed ZQ over the life of the April 2026 contract (top), and the spread $\Delta = \mathrm{ZQ}_{\mathrm{implied}} - \mathrm{ZQ}_{\mathrm{actual}}$ in basis points (bottom).

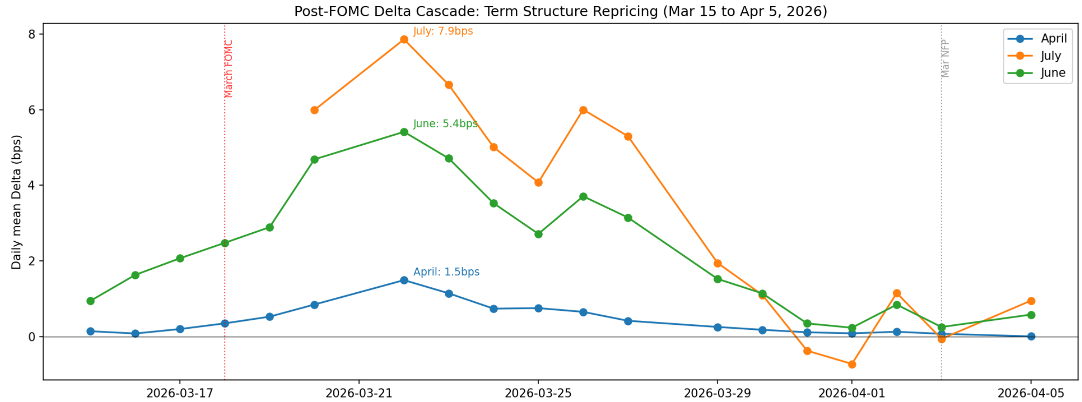

Figure 2: Post-FOMC Δ cascade. Daily mean Δ for the April, June, and July 2026 contracts around the 18 March 2026 FOMC decision.

Table 2: Per-meeting price discovery. IS_{poly} is the event-time Hasbrouck information share of Polymarket with its 95% moving-block-bootstrap interval. IS (1h) is the same share estimated on hourly bars, mechanically inflated toward 0.5 by within-bar contemporaneity. "pre-FOMC" and "post-FOMC" split each window at the prior meeting, and n is the number of event-time bars. All shares use ZQ-open hours only. December 2024 fails the stationarity test (ADF $p = 0.47$) and July 2026 is borderline stationary (ADF ≈ 0.09) with no post-FOMC window.

Meeting	IS_{poly}	95% CI	IS (1h)	pre-FOMC	post-FOMC	n
December 2024	0.826	[0.132, 0.999]	0.904	n/a	n/a	8,819
January 2025	0.036	[0.011, 0.400]	0.061	0.037	0.018	12,378
March 2025	0.037	[0.001, 0.757]	0.030	0.011	0.930	10,007
May 2025	0.200	[0.059, 0.563]	0.554	0.103	0.329	21,265
June 2025	0.288	[0.064, 0.576]	0.570	0.290	0.986	30,456
July 2025	0.407	[0.059, 0.707]	0.707	0.399	0.945	31,378
September 2025	0.111	[0.062, 0.213]	0.397	0.111	0.083	29,996
October 2025	0.031	[0.010, 0.076]	0.092	0.032	0.046	33,633
December 2025	0.008	[0.001, 0.028]	0.095	0.011	0.006	24,441
January 2026	0.008	[0.001, 0.062]	0.006	0.007	0.029	24,231
March 2026	0.022	[0.004, 0.096]	0.039	0.019	0.431	20,654
April 2026	0.039	[0.012, 0.139]	0.145	0.037	0.988	31,995
June 2026	0.008	[0.003, 0.019]	0.046	0.008	0.895	34,809
July 2026	0.040	[0.001, 0.136]	0.206	0.040	n/a	15,024

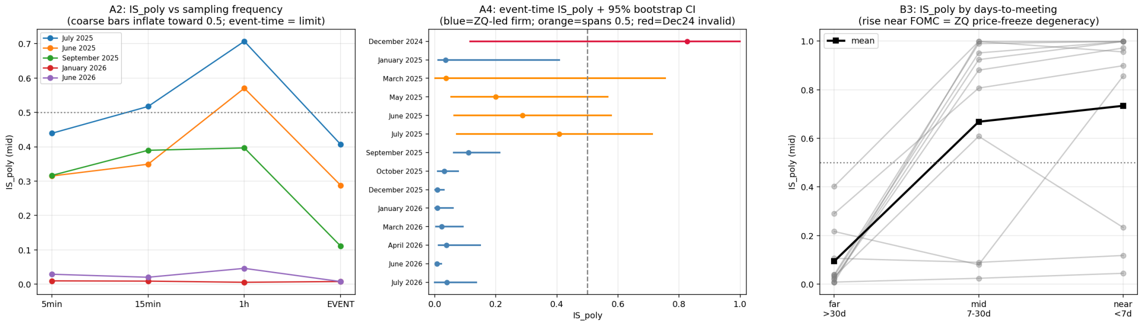

Figure 3: Robustness of the Polymarket information share. (Left) IS_{poly} falls as the sampling bar shortens, with event-time sampling as the limiting case. (Centre) Event-time IS_{poly} with 95% moving-block-bootstrap intervals. (Right) IS_{poly} by days-to-meeting.

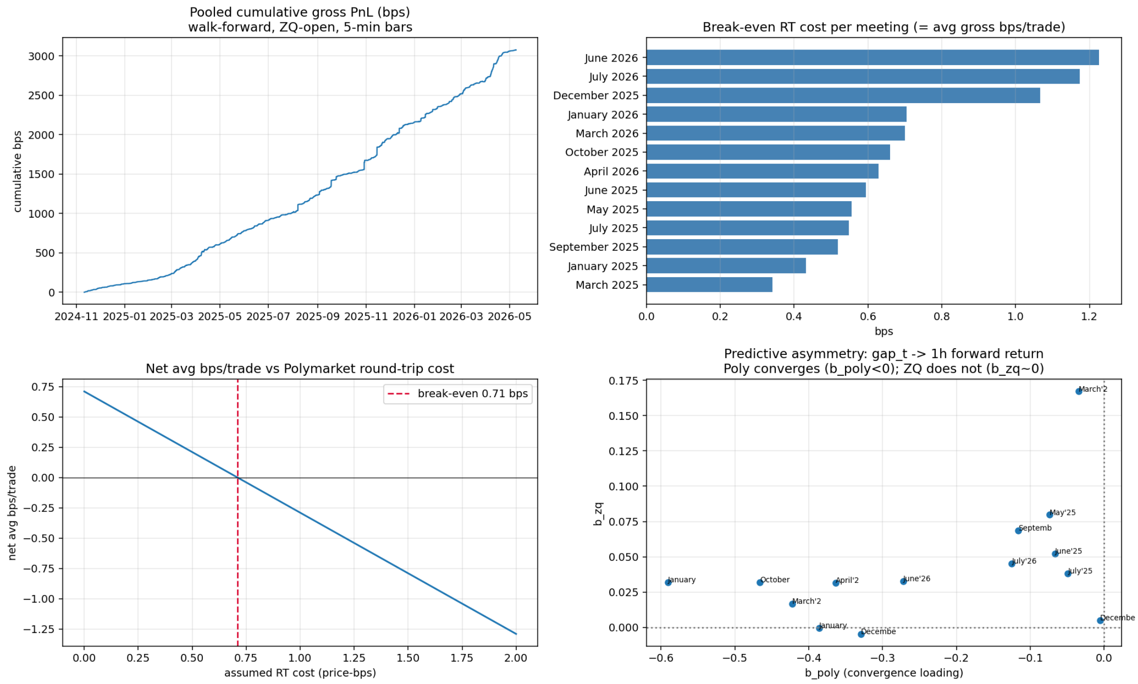

Figure 4: Convergence-trade summary. Pooled cumulative gross PnL (top left), break-even round-trip cost per meeting (top right), per-trade gross alpha against the assumed Polymarket round-trip cost (bottom left), and the predictive asymmetry of one-hour-ahead Δ loadings with Polymarket converging while ZQ does not (bottom right).

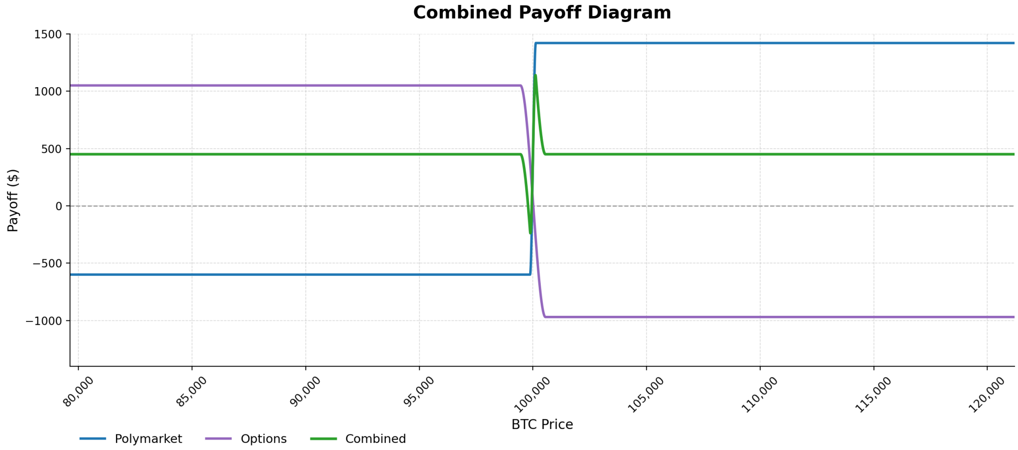

Figure 5: Combined payoff diagram. Expected payoff of the Polymarket leg, the CME call-spread leg, and the combined hedged position as a function of the Bitcoin settlement price.

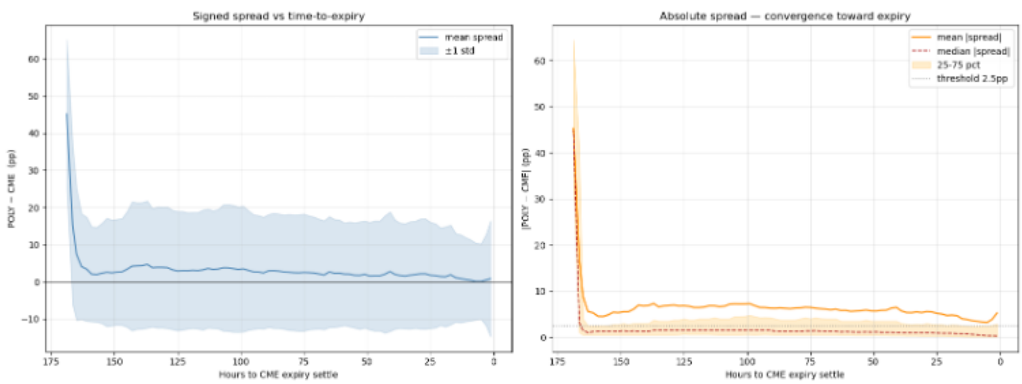

Figure 6: Term structure of the cross-venue spread. Signed Δ (left) and $|\Delta|$ (right) against hours to CME expiry.

Table 3: Gross settlement-exit results, in-sample (2025), out-of-sample (2026), and combined, for the 1,303-trade build. PnL is normalised to one dollar of notional per leg.

Period	**trades**	**total $**	**mean $/trade**	**win rate**
2025	727	+93.08	+0.128	78.7%
2026	576	+84.62	+0.147	80.0%
combined	1,303	+177.70	+0.136	79.3%

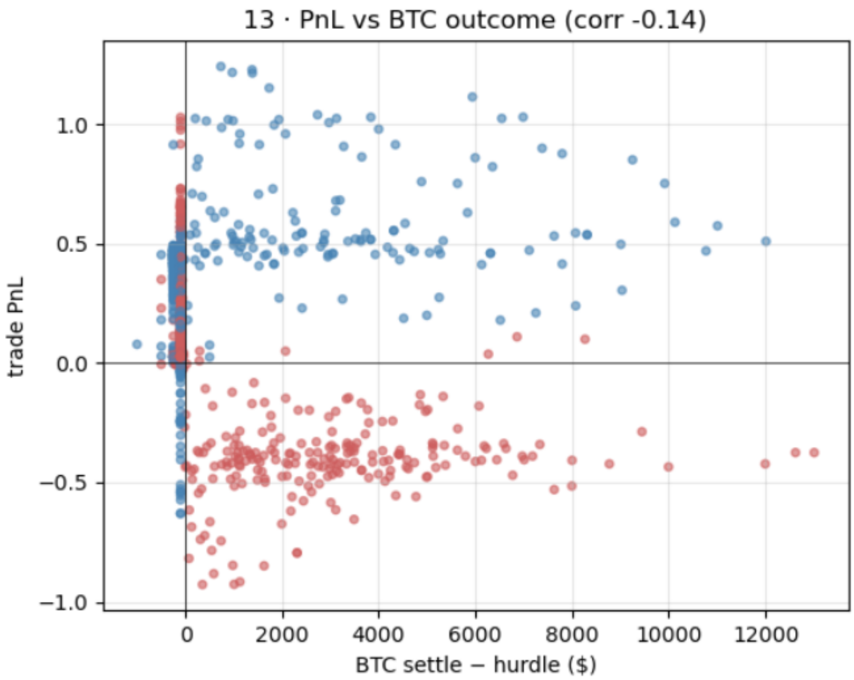

Figure 7: Trade PnL versus the Bitcoin outcome. Per-trade PnL against the signed distance of the Bitcoin settlement from the hurdle.

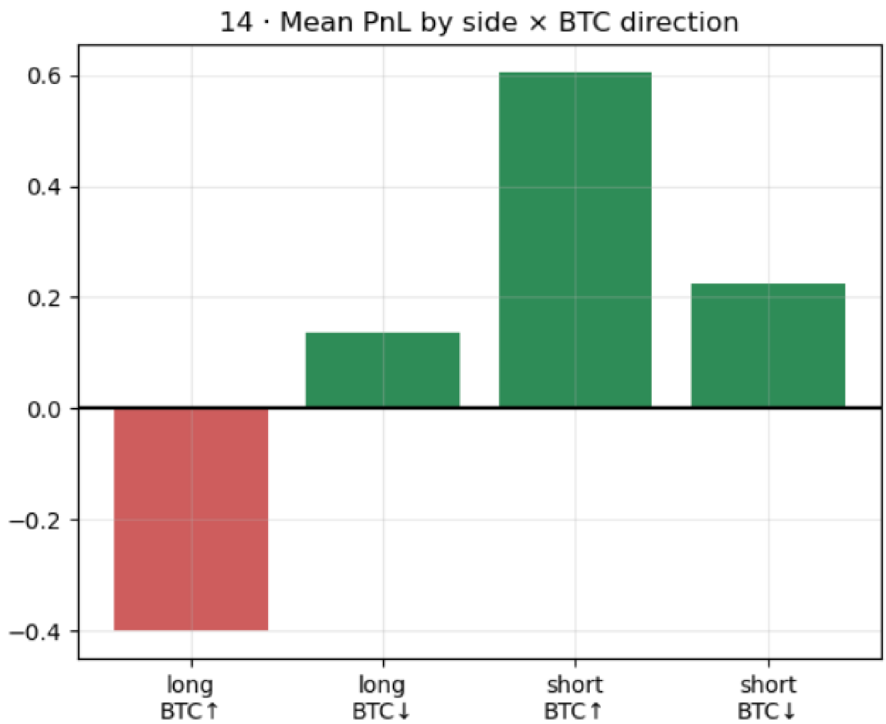

Figure 8: Mean PnL by side and Bitcoin direction. Average per-trade PnL split by trade side (long or short the Polymarket YES) and by whether Bitcoin moved with or against the position.

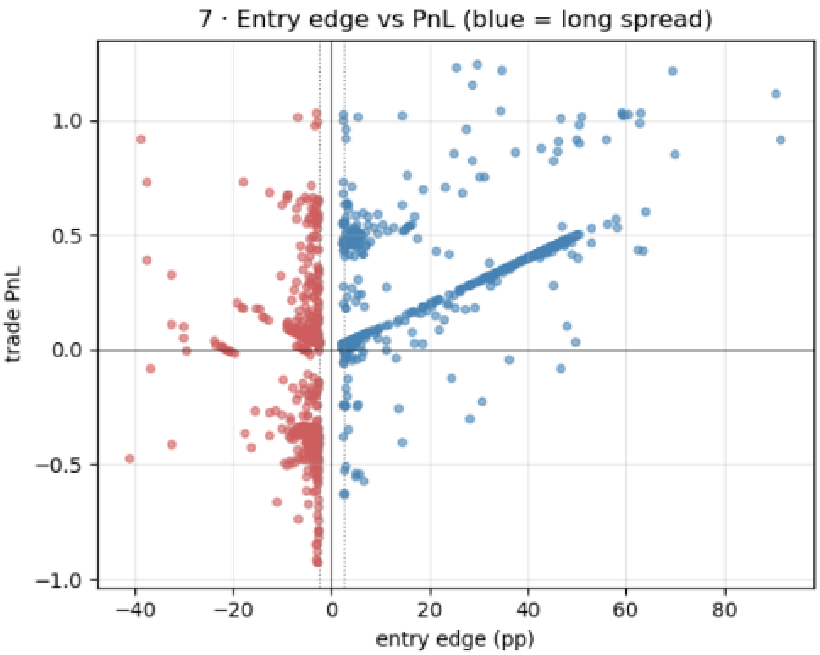

Figure 9: Entry Δ versus trade PnL. Realised per-trade PnL plotted against the entry Δ, with the V-shape arising from the side-flipping entry rule.

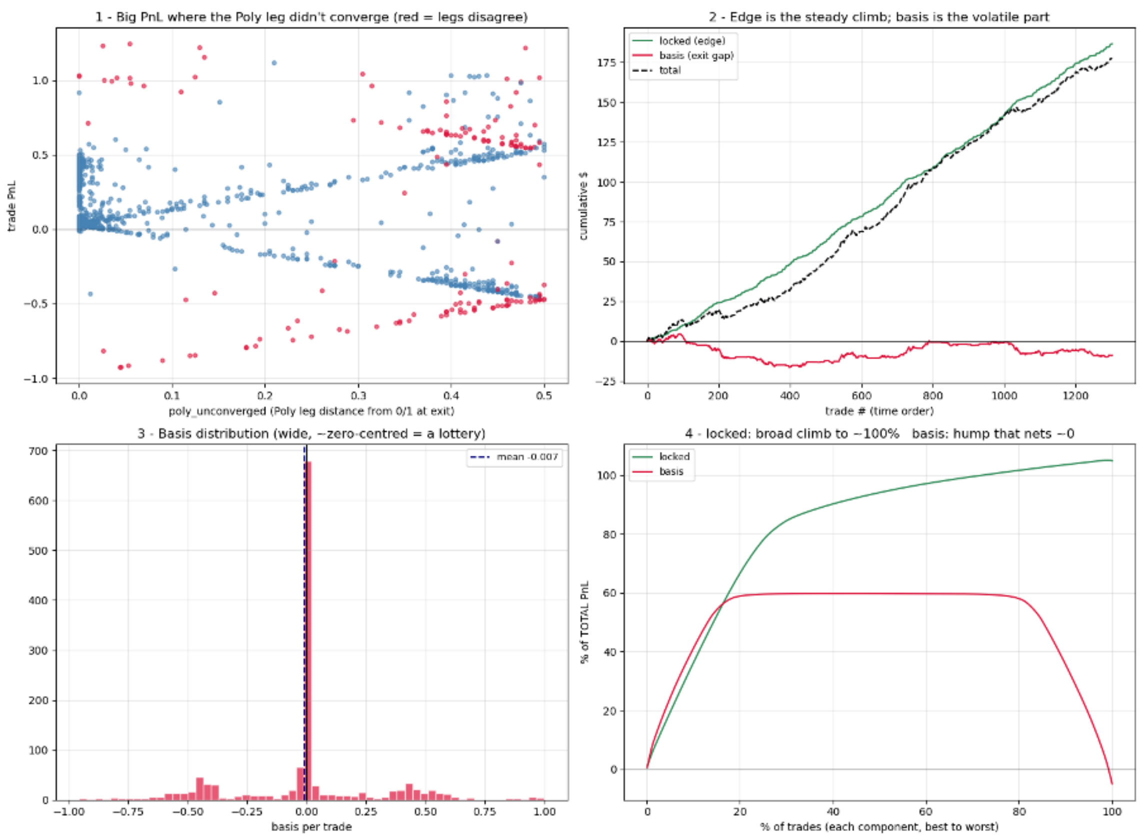

Figure 10: Locked edge versus basis. Decomposition of each trade into the locked entry edge captured the instant the position is opened, and the residual basis that moves between entry and settlement.

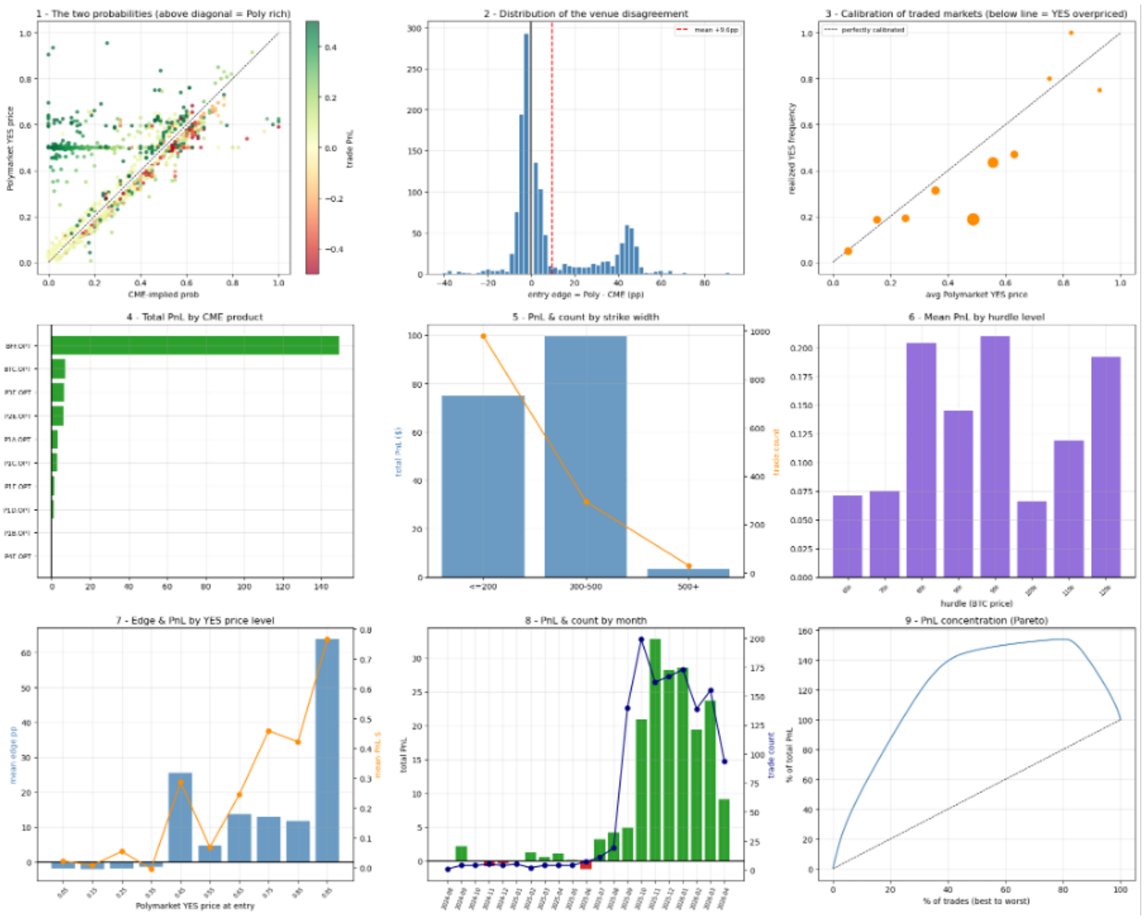

Figure 11: Bitcoin strategy diagnostics. Multi-panel trade-level diagnostics for the 1,303-trade settlement-exit build, showing the two probability series, the entry-spread distribution, calibration of traded markets, PnL by CME product and strike width, PnL by hurdle level, monthly PnL, and Pareto concentration of profit.

Table 4: Width distribution of the call-spread replication, per trade, across the 1,303 settlement-exit trades. W is the dollar width between the two bracketing CME strikes.

W **bucket (USD pts)**	**Trades**	**Share**	**Cumulative**	**Total $**	**Mean $/trade**
$\leq$ 200	979	75.1%	75.1%	$75.13	$0.077
300–500	294	22.6%	97.7%	$99.44	$0.338
500+	30	2.3%	100.0%	$3.13	$0.104
Total	**1,303**	**100.0%**		**$177.70**	**$0.136**

Bibliography

Arkham Intelligence. (2026, May 22). *Polymarket vs. Kalshi: How the world's two biggest prediction markets compare.* https://info.arkm.com/research/polymarket-vs-kalshi-how-the-worlds-two-biggest-prediction-markets-compare

Arrow, K. J. (1964). The role of securities in the optimal allocation of risk-bearing. *The Review of Economic Studies*, *31*(2), 91–96. https://doi.org/10.2307/2296188

Arrow, K. J., Forsythe, R., Gorham, M., Hahn, R., Hanson, R., Ledyard, J. O., Levmore, S., Litan, R., Milgrom, P., Nelson, F. D., Neumann, G. R., Ottaviani, M., Schelling, T. C., Shiller, R. J., Smith, V. L., Snowberg, E., Sunstein, C. R., Tetlock, P. C., Tetlock, P. E., . . . Zitzewitz, E. (2008). The promise of prediction markets. *Science*, *320*(5878), 877–878. https://doi.org/10.1126/science.1157679

Bansal, R., & Viswanathan, S. (1993). No arbitrage and arbitrage pricing: A new approach. *The Journal of Finance*, *48*(4), 1231–1262. https://doi.org/10.1111/j.1540-6261.1993.tb04754.x

Björk, T. (2009). *Arbitrage theory in continuous time* (3rd ed.). Oxford University Press.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, *81*(3), 637–654. https://doi.org/10.1086/260062

Breeden, D. T., & Litzenberger, R. H. (1978). Prices of state-contingent claims implicit in option prices. *Journal of Business*, *51*(4), 621–651.

CNBC. (2024, November 7). *Polymarket plans U.S. return after signaling Trump election win, founder says.* https://www.cnbc.com/2024/11/07/polymarket-plans-to-return-to-the-us-founder-says-.html

Cox, J. C., & Ross, S. A. (1976). The valuation of options for alternative stochastic processes. *Journal of Financial Economics*, *3*(1–2), 145–166.

Cox, J. C., Ross, S. A., & Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of Financial Economics*, *7*(3), 229–263.

Debreu, G. (1959). *Theory of value: An axiomatic analysis of economic equilibrium.* Yale University Press.

Gebele, J., & Matthes, F. (2026). Semantic non-fungibility and violations of the law of one price in prediction markets (arXiv:2601.01706). arXiv. https://doi.org/10.48550/arXiv.2601.01706

Harrison, J. M., & Kreps, D. M. (1979). Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, *20*(3), 381–408.

Hasbrouck, J. (1995). One security, many markets: Determining the contributions to price discovery. *The Journal of Finance*, *50*(4), 1175–1199. https://doi.org/10.1111/j.1540-6261.1995.tb04054.x

Hull, J. C. (2021). *Options, futures, and other derivatives* (11th ed.). Pearson.

Merton, R. C. (1973). Theory of rational option pricing. *Bell Journal of Economics and Management Science*, *4*(1), 141–183.

Pedersen, L. H. (2015). *Efficiently inefficient: How smart money invests and market prices are determined*. Princeton University Press.

Polymarket. (n.d.). *Resolution*. Retrieved June 1, 2026, from https://docs.polymarket.com/concepts/resolution

Ross, S. A. (1978). A simple approach to the valuation of risky streams. *The Journal of Business*, *51*(3), 453–475. https://doi.org/10.1086/296008

Rubinstein, M. (1976). The valuation of uncertain income streams and the pricing of options. *Bell Journal of Economics*, *7*(2), 407–425.

Saguillo, O., Ghafouri, V., Kiffer, L., & Suarez-Tangil, G. (2025). Unravelling the probabilistic forest: Arbitrage in prediction markets. In Z. Avarikioti & N. Christin (Eds.), *7th Conference on Advances in Financial Technologies (AFT 2025)* (Vol. 354, pp. 27:1–27:24). Schloss Dagstuhl – Leibniz-Zentrum für Informatik. https://doi.org/10.4230/LIPIcs.AFT.2025.27

Shepstone, C. (2026, April 20). How Kalshi and Polymarket settle event contracts (and disputes). *DeFi Rate*. https://defirate.com/prediction-markets/how-contracts-settle/

Shleifer, A., & Vishny, R. W. (1997). The limits of arbitrage. *The Journal of Finance*, *52*(1), 35–55. https://doi.org/10.1111/j.1540-6261.1997.tb03807.x

Shreve, S. E. (2004). *Stochastic calculus for finance II: Continuous-time models*. Springer.

Tsang, K. P., & Yang, Z. (2026). The anatomy of a blockchain prediction market: Polymarket in the 2024 U.S. presidential election (arXiv:2603.03136). arXiv. https://doi.org/10.48550/arXiv.2603.03136

Wolfers, J., & Zitzewitz, E. (2004). Prediction markets. *Journal of Economic Perspectives*, *18*(2), 107–126. https://doi.org/10.1257/0895330041371321