SEEING LIKE A MARKET

Event Contracts and Market Topology

WWW.LAURIS.XYZ

March 2026

Abstract

Institutions hedge binary event risk with derivatives that price the entire volatility path when only the endpoint matters. A fund hedging a Fed rate decision through options pays for variance exposure it does not need, rents dealer balance sheet it could bypass, and absorbs replication frictions that add cost without transferring risk. Event contracts pay $1 if an outcome occurs and $0 otherwise, avoiding these costs entirely.

This paper introduces the *Vega Wedge*, measuring the structural cost gap between replicating a binary payoff through options and purchasing the equivalent probability in a prediction market. The Wedge decomposes into variance risk premium, dealer balance sheet cost, and replication friction. The *Liquidity Seesaw* identifies when prediction market execution costs fall below these structural savings at institutional scale.

We test the framework across eighty-seven event-contracts in eleven categories spanning five asset classes and the variance risk premium gradient. High-VRP categories cross the threshold first. The ordering holds without exception across high-reliability categories. Smile analysis shows prediction markets produce the same distributional features that derivatives encode as volatility surfaces, through aggregation alone, without dealer apparatus. The same features transmit at a fraction of the cost.

These are not competing tools for the same task. Prediction markets and derivatives represent distinct forms of order for pricing contingent claims. The distinction follows Hayek (1973)'s typology of spontaneous and constructed order, applied here for the first time to financial market microstructure. The Vega Wedge measures what that distinction costs. The Liquidity Seesaw predicts where it breaks.

Keywords: Event contracts, prediction markets, variance risk premium, market microstructure, Hayek, spontaneous order

JEL Classification: G13, G14, D47, D82

I. Introduction

I.1. The Problem of Discrete Risk

Institutional risk frequently resolves in discrete states. Will the Federal Reserve cut rates at the September meeting? Will headline CPI print above the 3% threshold? Will the antitrust authority approve the merger? Will the yield curve invert before year-end?

These questions have binary answers: yes or no. The uncertainty is real, affecting portfolio values, funding costs, and strategic decisions, but the resolution is digital. One state or the other.

The financial system makes institutions hedge these risks with continuous instruments. A pension fund worried about a Fed surprise buys Fed Funds options and pays for the whole volatility path when it only cares about the endpoint. A corporate treasurer hedging merger approval ends up in exotic structures with path dependencies that have nothing to do with the actual outcome. A macro fund with a CPI view takes on Greeks it does not want and cannot hedge out.

This mismatch creates a structural efficiency leak. Institutions pay for optionality they do not need, rent dealer balance sheets they could bypass, and take on model risk for dynamics that will not affect their payoff. The leak is built into how derivatives markets were designed.

Variance Risk Premium. Options prices systematically exceed fair value for variance exposure. This is compensation for bearing volatility uncertainty, not mispricing. But for binary outcome hedgers, volatility uncertainty is irrelevant. They do not care whether implied vol was 20% or 40%. They care whether the Fed cut or held. VRP functions as a tax on these hedgers.

Dealer Balance Sheet. The volatility surface does not exist in nature. Dealers construct and maintain it by warehousing risk, managing inventory, providing liquidity. This requires capital, and post-crisis regulation made capital expensive. Hedgers rent dealer balance sheet whether they need intermediation or not.

Replication Friction. Synthesizing a binary payoff from vanilla options requires multiple legs, each carrying particular transaction costs. Strike discreteness creates approximation error. Margin requirements may exceed the actual exposure, especially for more structured products, which means replication has costs beyond the risk being transferred.

Event contracts are binary instruments: \$1 if the outcome occurs, \$0 otherwise. They bypass all three sources of leak by pricing the endpoint directly. The price emerges from aggregation. No path, no Greeks, no vol-seller dealer intermediation beyond exchange matching. VRP does not apply because there is no path uncertainty. Balance sheet rental does not apply because there is no intermediation.

Prediction markets have reached institutional scale. Monthly volume exceeds ten billion dollars. Goldman Sachs has opened an Event Desk. Intercontinental Exchange invested two billion dollars in Kalshi at an eight billion dollar valuation. The regulatory path has cleared. Market makers from traditional finance provide liquidity on prediction market venues. The capital has arrived. The theory has not caught up.

I.2. The Question

If event contracts are structurally cheaper for binary risk, why have institutions not adopted them?

If we focus purely on the market microstructure side, the answer, as with many things in markets, is liquidity. Not ignorance, not inertia. Liquidity.

Prediction markets have historically been thin. In 2024, spreads ran 0.5–0.7% of notional, but market impact for institutional positions added 10–13% more. Depth on FOMC contracts sat in single-digit millions while Treasury futures traded in billions. A $3 million position – our assumed entry threshold for institutional hedging – would have incurred 13% total execution cost. The structural savings existed but could not be captured at size.

A fund seeking to hedge $10 million of FOMC event risk faced a stark choice. In options markets, the trade executes cleanly: tight spreads, deep books, minimal impact. The cost is structural (VRP, dealer margin, replication friction) but execution is certain. In prediction markets, the trade might not execute at all. It would move the market against itself. The savings were real but inaccessible.

This creates what we term the *Liquidity Seesaw* (Table 1; Figure 2 formalizes the crossing condition):

Table 1. The Liquidity Seesaw

	Legacy Markets	**Event Markets**
Structural Cost	High (VRP, dealer margin, balance sheet)	Low (no Greeks, peer-to-peer, direct pricing)
Execution Cost	Low (deep liquidity, tight spreads)	High (wide spreads, thin books, severe impact)

The structural cost functions as a tax; the execution cost functions as switching friction. Following Klemperer (1987), capital migrates when the tax exceeds the friction. Incumbents harvest locked-in users via structural costs; entrants compete by reducing friction.

The Seesaw is not fixed. It tips when event market liquidity scales enough that execution costs fall below structural savings. By late 2025, execution costs for $3 million positions had fallen to 2–3%, a 78% reduction from 2024 levels. The gap narrowed from 12 percentage points to under 2. But the trajectory varies by category. The structural wedge is large where variance risk premium is high and small where derivatives are liquid and well-arbitraged. Which categories cross first, why, and at what scale? This paper tests the ordering across the full gradient – eleven categories, 87 event-contracts, five asset classes.

I.3. The Contribution

This paper makes four contributions.

First, a theoretical claim. Prediction markets and derivatives markets are different forms of order for pricing contingent claims. The volatility surface is *Taxis*: constructed, maintained by dealers who bear costs for its upkeep. The event order book is *Cosmos*: emergent, arising from dispersed participation without central coordination. The distinction follows Hayek (1973), applied to law, urban planning, and institutional design – but never to financial market microstructure. We argue the distinction is structural, not metaphorical, with measurable consequences for how contingent claims are priced.

The theoretical claim decomposes into three sub-claims:

1. **The Cost Wedge (Vega Wedge).** Constructed order imposes measurable structural costs (VRP, balance sheet, replication friction) that emergent order avoids. Section III formalizes; Section V tests.

2. **The Observability Wedge.** Constructed order, when dealer maintenance withdraws from public markets, ceases to emit observable prices while continuing to function inside relationships. Emergent order fails loudly; constructed order fails quietly. Section V documents this asymmetry.

3. **Content Equivalence.** Emergent order produces the same distributional features (fat tails, skew, regime dynamics) that constructed order encodes as volatility surfaces. The cost differential is apparatus rent, not a quality discount. Section V tests via smile analysis.

Second, a measurement framework. The *Vega Wedge* captures the total cost differential between replicating a binary payoff through options and purchasing it directly in an event market. It decomposes into structural components (variance risk premium, dealer balance sheet cost, replication friction) and microstructural components (spread, depth, market impact). The *Liquidity Seesaw* identifies when prediction market execution costs fall below structural savings, making event contracts cost-competitive at institutional scale. Together, these tools operationalize the ontological distinction as a measurable cost comparison.

Third, empirical application. We test the framework across eleven categories spanning the variance risk premium gradient, from central bank rate decisions to contested elections and cryptocurrency thresholds. The sample covers 87 event-contracts and 51 independent VRP observations across five asset classes. The gradient predicts which categories cross the Seesaw threshold first. The empirical work also reveals an asymmetry in how the two forms of order fail: constructed order can cease to emit observable prices while continuing to function inside dealer relationships. Smile analysis shows that prediction markets produce the same distributional features that derivatives encode as volatility surfaces – through aggregation alone, without dealer apparatus.

Fourth, a measurement apparatus. The framework is implemented as a reproducible pipeline that ingests from five institutional data sources across four venue classes, computes structural wedges and execution costs for arbitrary event categories, and produces every empirical result reported here through a single deterministic execution. New categories require a data connector and a configuration entry, not new methodology. What we have built, others can operate – and point at categories this paper does not cover.

I.4. The Theory of Prediction Markets

Prediction markets have been predominantly understood at two levels:

- **L1: Behavioral.** Prediction markets as speculative venues. People bet on elections, sports, crypto prices, and anything else that has a market. Foundational, and correct, but incomplete.

- **L2: Informational.** Prediction markets as truth production. Aggregated bets outperform polls, models, and experts. Prices encode dispersed knowledge via crowd-sourced epistemology. Also correct, but also incomplete. This framing has been remarkably productive, grounding decades of prediction market research and, as recently as last year, informing significant institutional capital deployment with explicit citation to Hayek (1945).

Neither framing is our contribution. We argue for a third level:

- **L3: Ontological.** Prediction markets as an emergent topology for coordinating state-contingent prices without constructed dealer intermediation. The distinction rests on Hayek (1973): Cosmos (emergent order) versus Taxis (constructed order). The volatility surface is built and maintained. The order book emerges from participation. Crucially, observability is constitutive of Cosmos: the price is public because publicity is the mechanism. Taxis can function while ceasing to emit public prices. The failure modes differ: Cosmos fails loudly, Taxis fails quietly. Section II develops this foundation. This ontological claim generates specific predictions. If two forms of order exist, and one imposes structural costs the other avoids, then above a liquidity threshold the efficient order should displace the inefficient one for claim types where both can function. Succession, not coexistence. Displacement, not competition.

The claim is not that prediction markets are "better" in some generic sense. The claim is that they represent a distinct mode of price discovery with different cost structures, and that for a specific class of risks – binary, observable, time-bounded – those cost structures favor emergent order once liquidity crosses threshold.

I.5. Scope and Limitations

The claim is narrow.

Claim class. We consider only digitizable risks: events that resolve to binary or multinomial outcomes, with unambiguous settlement, within bounded time horizons, and material institutional exposure. Fed decisions, CPI prints, merger approvals, yield curve inversions.

Participant class. We consider institutional hedgers: non-bank market participants such as pension funds, endowments, hedge funds, corporates, and sovereign wealth funds. We exclude banks (whose regulatory capital treatment differs) and dealers (whose business model depends on the apparatus we claim is contestable).

Threshold condition. We do not claim event contracts are cheaper today for institutional positions in all categories. We claim they are structurally cheaper, and that execution costs are converging toward the point where the structural advantage can be captured at scale. For some categories the threshold has been crossed. For others it has not. The variance risk premium gradient predicts the ordering.

We do not claim event contracts will replace derivatives markets. We claim they will absorb the binary-outcome hedging function from derivatives markets, for claim classes where the structural wedge exceeds execution friction, leaving derivatives to serve continuous-risk hedging, path-dependent exposure management, and relationship-intensive bespoke transfers.

Venue versus structure. This paper is not about operationalization. These are claims about market topology. Practical implementation concerns are out of scope. We do not speculate on mechanism design for new infrastructure or products. The intent is to provide theory and measurement apparatus for researchers, practitioners, and operators – a potential origin point for a Lakatosian progressive research programme.

The structural cost comparison abstracts from venue-specific implementation. Prediction market platforms differ substantially in resolution architecture. Polymarket relies on UMA oracle resolution, which carries smart contract and oracle risks that matter for institutional adoption. Kalshi employs legally standardized resolution with determination rules hard-coded into public contract documents, operating under CFTC authorization with traditional legal enforcement. The structural wedge applies to the form of order, not to any particular venue. Kalshi demonstrates that emergent order can operate within traditional legal infrastructure. The platform-specific risks are separable from the instrument economics.

The structural cost advantage is necessary but not sufficient for institutional migration. Platform risk, regulatory integration, fiduciary defensibility, operational onboarding: these barriers exist and this paper does not measure them. What we measure is the necessary condition. Whether and how institutions operationalize that advantage is a different inquiry, requiring different methods and different data.

I.6. Roadmap

Section II develops the theoretical apparatus: the displacement claim, the ontological distinction between order types, the segmentation hypothesis linking variance risk premium to threshold ordering, and falsification conditions. Section III formalizes the Vega Wedge, decomposing the structural cost differential and specifying measurement methodology for each component. Section IV presents data sources and measurement across eleven categories spanning the VRP gradient, from FOMC decisions and CPI prints through commodity thresholds, merger approvals, and cryptocurrency touch markets. Section V reports findings: threshold comparison across categories, the Observability Wedge, smile analysis, and gradient validation. Section VI discusses limitations, welfare implications, and the contribution in its strongest form.

II. Thesis

II.1. The Displacement Claim

Claim. For digitizable risks where variance risk premium is positive, event contracts impose lower structural costs than options replication for binary payoffs. The cost advantage is structural but inaccessible until prediction market execution costs fall below the savings. Once they do, hedging flow diverts.

The claim has three components:

1. **Digitizable risks** are a distinct category with specific characteristics
2. **Displacement** means something precise: diversion of hedging flow, not mere coexistence
3. **The threshold** is identifiable, measurable, and varies across the variance risk premium gradient

Each requires precision.

II.2. Defining Digitizable Risks

A risk is digitizable if it satisfies four conditions:

Binary or multinomial outcome. The resolution is discrete. Yes or no. Candidate A, B, or C. Above or below threshold. Not continuous price discovery but state realization.

Observable settlement. The outcome can be determined without judgment. Fed announces rate. CPI publishes. Merger closes or terminates. No ambiguity, no interpretation, no "in the eye of the beholder."

Time-bounded. Resolution occurs at a known or bounded time. FOMC meeting dates are published years in advance. CPI releases monthly on schedule. Merger deadlines are contractually specified. The uncertainty is what, not when.

Material exposure. Institutions actually hold this risk. It affects portfolio values, funding costs, business outcomes. Not a curiosity but a genuine hedging need.

Examples:

- Fed rate decisions (binary per meeting, multinomial across path)
- CPI threshold crossings (above/below target)
- Election outcomes (binary or multinomial)
- Merger approvals (binary)
- Regulatory decisions (binary)
- Asset price thresholds (touch markets, index levels)

Non-examples:

- Subjective assessments (“will the economy be strong?”)
- Open-ended timelines (not time-bounded)
- Risks no one actually hedges (no material exposure)

II.3. The Institutional Hedger

The relevant actor is the institutional hedger: non-bank market participants with genuine exposure to digitizable risks who currently manage that exposure through derivatives.

This includes:

- **Pension funds** with rate exposure around FOMC
- **Hedge funds** with event-driven strategies
- **Corporates** with M&A contingencies
- **Asset managers** with CPI-linked mandates
- **Sovereign wealth funds** with policy-sensitive allocations

This excludes:

- **Banks**, whose regulatory capital treatment creates different incentives
- **Dealers**, whose business model depends on the constructed apparatus
- **Retail speculators**, who lack the position sizes for institutional relevance

The threshold question is: at what liquidity level can these institutions execute hedges in prediction markets at total cost below their current derivatives cost?

II.4. What Displacement Means

Displacement does not mean elimination. Derivatives markets will persist. They serve functions beyond binary event pricing: continuous risk transfer, path-dependent hedging, cross-asset exposure management, relationship-intensive bespoke transactions.

Displacement means diversion of flow. Hedging demand that currently expresses through options (buying variance protection to hedge binary outcomes) migrates to event contracts where total cost is lower.

The dynamic is reflexive:

- Flow diverts to prediction markets
- Prediction market liquidity increases
- Execution costs fall further

- More flow diverts

Simultaneously:

- Options flow decreases
- Dealer incentive to maintain volatility surfaces for that claim class decreases
- Dealer maintenance cost per trade rises, incentivizing further withdrawal, degrading the apparatus, until the market ceases to function for that claim class
- The prediction market absorbs the additional flow and gets thicker

II.5. The Segmentation Hypothesis

Not all event categories are equally contestable.

The structural wedge varies with variance risk premium. High-VRP events offer larger structural savings because the tax embedded in options pricing is higher. Low-VRP events offer smaller savings. VRP becomes the segmentation variable: it determines which categories prediction markets can contest and when.

The prediction: high-VRP categories cross threshold first. The structural wedge is larger, so less liquidity improvement is required to tip the Seesaw. Low-VRP categories require deeper prediction market liquidity before the structural advantage can be captured at institutional scale.

This paper examines eleven categories spanning the gradient: central bank rate decisions (FOMC, ECB, BOJ), macroeconomic releases (CPI, GDP), cryptocurrency thresholds (BTC), commodity thresholds (gold, silver, retail gas), equity index levels, and contested elections. The categories were selected to populate the gradient from low-VRP (telegraphed rate decisions with narrow implied-realized spreads) through high-VRP (contested asset price thresholds and political outcomes with wide variance risk premia). The gradient generates a specific ordering prediction for these eleven categories. Section V tests it.

The gradient extends beyond what this paper measures. Merger approvals, regulatory decisions, and yield curve inversions are digitizable risks that occupy positions on the same VRP axis. The segmentation logic applies to any category satisfying the conditions of Section II.2. The eleven categories examined here are the empirical test of a prediction that spans the full space of digitizable risks.

The segmentation logic follows (Klemperer, 1995): where the tax is high and switching friction is falling, migration accelerates. Where the tax is low, friction dominates longer. The gradient predicts the ordering.

II.6. Falsification Condition

The ontological claim is falsifiable.

The framework generates two independent falsification conditions.

The ontological distinction is falsified if the structural cost differences do not exist. If prediction markets recreate VRP-equivalent costs through informed trader extraction or adverse selection pricing, the distinction collapses to description. If derivatives achieve near-zero structural costs for binary payoffs through disintermediation, the cost asymmetry vanishes. The distinction holds as long as the cost structures genuinely differ by kind.

The displacement prediction is falsified independently. If both orders persist at scale for the same claim class, if institutional flow does not migrate above sustained threshold crossing, that is coexistence at equilibrium. The prediction fails. But the ontological distinction may still hold. Cosmos and Taxis can genuinely differ in cost structure while coexistence persists through regulatory lock-in, fiduciary convention, or relationship value. Coexistence falsifies the prediction. It does not falsify the distinction.

The segmentation hypothesis adds a second falsification condition: if low-VRP categories cross threshold before high-VRP categories, the gradient mechanism fails. The ordering prediction is testable across the full gradient independently of whether displacement occurs for any individual category. Section V reports the result for the eleven categories examined, with elections consistent at lower measurement reliability. The condition remains live for categories not yet tested: the gradient predicts where mergers and regulatory decisions should fall relative to categories already measured.

II.7. The Hayekian Foundation

The displacement claim rests on a deeper observation about market structure. The volatility surface and the event order book are different approaches to price discovery.

Hayek (1945) argued that prices aggregate dispersed knowledge no central authority could possess. The price system is a "telecommunications system" that transmits information from those who have it to those who need it without requiring explicit communication of the underlying knowledge.

In *Law, Legislation and Liberty* Hayek (1973), Hayek distinguished two types of order:

Taxis (constructed order). Order created by deliberate arrangement according to a plan. Organizations, hierarchies, designed systems. Taxis requires a designer, resources for maintenance, ongoing coordination.

Cosmos (spontaneous order). Order emerging from individuals following rules without awareness of the overall pattern their actions create. Markets, language, common law. Cosmos requires only participants and rules. No designer, no central coordinator.

The distinction is ontological, not structural. A Taxis price exists because an apparatus produces it. A Cosmos price exists because participants converge on it. When the apparatus withdraws from public markets, the Taxis price ceases to exist as observable phenomenon – a matter of ontology, not degradation. When participants disperse, the Cosmos price becomes noisy but remains observable. The mode of existence differs, not just the mechanism of discovery.

The price must be observable for Hayekian coordination to function. A private price, produced inside a bilateral relationship, may transfer risk between parties but does not aggregate dispersed knowledge or coordinate action among strangers. This is constitutive of the Cosmos/Taxis distinction.

Cosmos produces observable prices by construction. The order book is public because publicity *is* the mechanism: dispersed orders aggregate into visible depth, executed trades print to the tape, the price exists as artifact of the matching process. You can see where the market is, even if you cannot see why. Observability is not a design choice. It is what makes an order book an order book.

Taxis need not produce observable prices. When maintenance is active and public-facing, the volatility surface emits quotes, depth, continuous price formation. But when flow thins below the threshold where public quoting is profitable, Taxis prices retreat into the apparatus. Pricing continues inside dealer relationships; clients with access receive quotes; trades clear. The price ceases to exist as public object. Invisible to those outside.

This asymmetry is testable. Where we find instruments that exist legally but not as observable markets, we find Taxis at its boundary. Section V.3 documents one such case: Fed Funds options, listed and clearable, but with no public market. The absence is the predicted failure mode of constructed order when maintenance withdraws from public instantiation.

Hayek (1973) developed this distinction for legal and institutional orders. We apply it to market microstructure: the same ontological distinction characterizes how contingent claims can be priced. A narrower domain than Hayek's, but the structural logic transfers. Does pricing order require a constructing apparatus, or does it emerge from participants following simple rules? Bowles (2017) apply Hayek's framework to financial markets, supporting this extension.

Hayek anticipated coexistence across ordering types. Our displacement prediction extends his distinction into a domain where the cost asymmetry is measurable and the functional overlap is complete. Coexistence persists where Cosmos and Taxis serve different functions. It should not persist where they serve the same function at different cost.

The volatility surface is Taxis. It is constructed by dealers who:

- Run models to interpolate and extrapolate prices across strikes and tenors
- Commit capital to warehouse inventory and provide liquidity
- Coordinate with other dealers through interdealer markets
- Maintain infrastructure for margin, clearing, and settlement

Glosten (1994) showed that equilibrium price schedules emerge from order flow without dealer intervention. The electronic open limit order book aggregates dispersed information through the interaction of patient and impatient traders, producing prices that reflect available information without a designated market maker. This is how Cosmos generates prices.

Zhu (2014) extended this to venue competition. Informed traders self-select into lit venues because their correlated orders create execution risk in dark pools. Uninformed traders get better execution in dark pools and migrate there. Informed flow concentrates on the transparent venue, improving price discovery. The mechanism requires that the lit venue exist and emit prices. Dark pools are parasitic on lit markets; when the lit venue disappears, price discovery stops.

Emergence is cheap. The structural costs we measure in options do not exist in the event market. No VRP because there is no variance to bear, only states. No balance sheet rental because there is no intermediary, only peer-to-peer matching.

But Cosmos requires density. Sparse participation produces noise rather than signal: prices move on small trades, spreads exceed information content, the market becomes unreliable.

Fed Funds options illustrate the ontological distinction concretely. The contract exists, but the apparatus is maintained privately, on dealer balance sheets, for bilateral negotiation. The CME listing provides clearing; the pricing happens in voice. The result is a market that functions for those inside the apparatus but fails to instantiate the objects (quotes, depth, continuous price formation) that make something observable as a market. This is Taxis at its boundary: not absent, but illegible from outside. Cosmos fails differently: thin participation produces wide spreads and shallow depth, but the order book remains visible. Cosmos fails loudly, on-screen. Taxis fails quietly, by not printing anything you can analyze.

The Vega Wedge measures the cost difference between these orders. The Seesaw measures when Cosmos achieves sufficient density to serve as venue.

When the Seesaw tips, Cosmos eats Taxis.

II.8. On "Never Necessary"

If spontaneous order could always price binary claims, why did constructed order emerge? The objection rests on a conflation. Dealer intermediation solved coordination problems, not pricing problems.

For binary, observable, time-bounded claims, the constructed pricing apparatus solved coordination problems, not pricing problems. A claim about what the apparatus was *for*, not a counterfactual about what could have existed earlier.

Dealer intermediation was not functionless. It provided real functions:

- **Counterparty reliability.** Before full collateralization was technically feasible, counterparty risk required trusted intermediaries
- **Regulatory arbitrage.** Before regulatory clarity on event contracts, derivatives provided a permissible wrapper for contingent claims
- **Balance sheet capacity.** Before peer-to-peer matching at scale, someone had to warehouse risk across time

Genuine coordination problems. But coordination problems, not pricing problems. The apparatus solved how to *transact* on contingent claims, not how to *price* them.

Technology has dissolved these coordination problems:

- **Full collateralization** eliminates counterparty risk
- **Regulatory clarity** eliminates arbitrage need

- **Peer-to-peer matching** eliminates warehousing requirement

Technology has dissolved these coordination problems. Whether the apparatus was also necessary for bootstrapping liquidity in thin markets is an open question; FOMC prediction markets did not achieve institutional depth until 2024. What the data show is that once Cosmos achieves sufficient depth, pricing operates without constructed order. The apparatus may have been necessary for bootstrapping. It is no longer necessary for pricing.

Prediction markets for elections, regulatory decisions, and platform-specific outcomes prove it: price discovery emerged from participation alone, without apparatus. The scaffolding is now unnecessary.

Smile analysis on BTC (Section V.4) provides direct evidence: the distributional features that derivatives encode as volatility surfaces (fat tails, skew, regime dynamics) emerge from aggregation alone, without dealer apparatus.

The displacement pattern has precedent. Jain (2005) documented electronic trading displacing floor trading across 120 exchanges, with electronic venues capturing primary price discovery within 2–5 years of achieving competitive liquidity. The mechanism is consistent: lower structural costs, initial liquidity disadvantage, threshold crossing, rapid migration.

Cantillon and Yin (2007) examined the Bund futures migration from LIFFE to DTB. DTB won through newcomers, not switchers. Access deregulation under the EU Investment Services Directive, combined with electronic trading, lowered barriers for firms that previously used brokers or did not participate. The market expanded, liquidity concentrated, the incumbent collapsed. The coordination problem solved itself through expansion.

A note on precedent scope. Both Jain and Cantillon-Yin document venue migration: the same instrument moving between trading venues. Our claim concerns instrument substitution: different instruments serving the same hedging function. The mechanism (threshold crossing, liquidity concentration, incumbent withdrawal) is analogous, but the objects being displaced differ. Venue migration precedents inform the dynamics; they do not prove instrument substitution will follow the same pattern.

Prediction markets follow the same pattern. Crypto rails, regulatory clarity, and zero-barrier account creation expand the participant base. Retail speculators, crypto-native funds, and event-driven traders who would never have accessed dealer-intermediated derivatives now participate directly. The market expands first. Then it tips.

III. Theoretical Framework: The Vega Wedge

III.1. Payoff Equivalence

A binary event contract pays \$1 if outcome O occurs and \$0 otherwise (Figure 1). Let p denote the market-implied probability of O. The contract trades at price p, yielding expected value $p \times \$1 + (1-p) \times \$0 = p$. A buyer's maximum loss is p (if O does not occur); maximum gain is $(1-p)$ (if O occurs). The payoff is discontinuous at the outcome boundary.

The same payoff can be replicated using vanilla options. Consider an underlying F (e.g., Fed Funds futures) where outcome O corresponds to F exceeding threshold K at expiration. The digital payoff is replicated by a vertical call spread:

- Long 1 call at strike K
- Short 1 call at strike $K + \varepsilon$

As $\varepsilon \to 0$, the spread payoff converges to:

- \$1 $\times$ (contract multiplier) if $F > K$ at expiration
- \$0 if $F \leq K$ at expiration

The convergence is exact in the limit. The vertical spread is a textbook replication of a digital option (Carr, 1998). The equivalence is definitional, not approximate. Yet the instruments are not priced equivalently. The options replication embeds costs absent from the event contract. The Vega Wedge captures this differential.

III.2. Decomposing the Structural Wedge

The structural Wedge comprises three components, each representing a cost embedded in options but absent from event contracts.

Component 1: Variance Risk Premium (VRP)

Options prices embed compensation for variance uncertainty. Define:

- IV: Implied variance (squared implied volatility from option prices)
- RV: Realized variance (actual squared returns over the holding period)
- VRP = IV − RV

Carr and Wu (2009) developed model-free methodology for measuring VRP using synthetic variance swaps. Their approach constructs a portfolio of out-of-the-money options that replicates variance exposure:

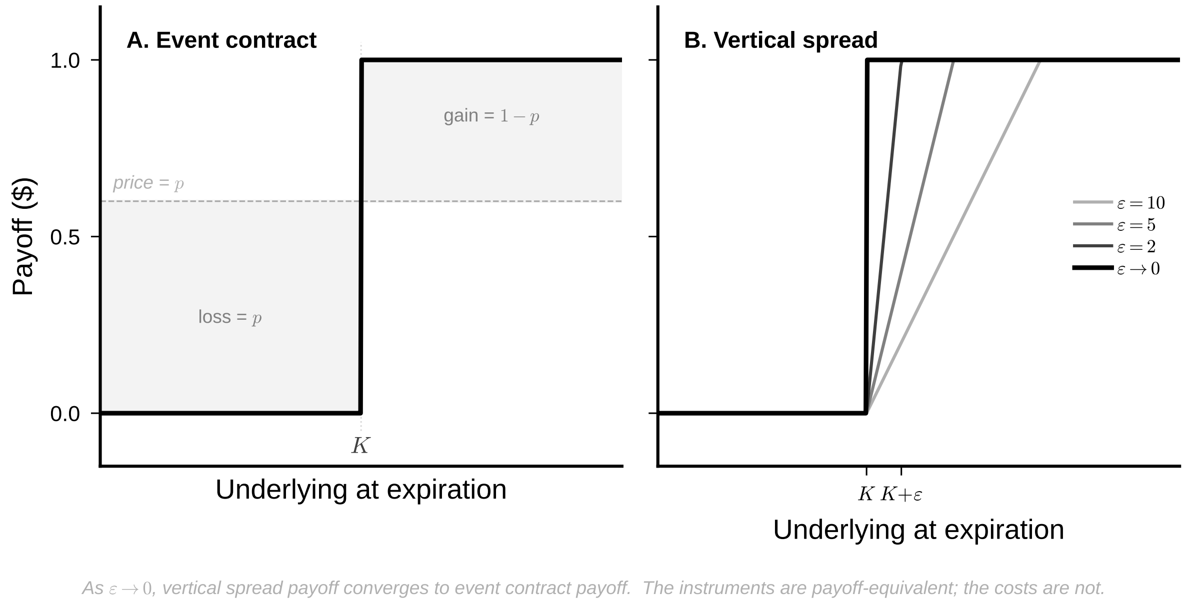

Figure 1. Payoff equivalence between event contracts and vertical spreads. Panel A: event contract step-function payoff ($1 above threshold K, $0 below). Panel B: vertical call spread convergence as $\varepsilon \to 0$, showing ε values of 10, 5, 2, and the limit. As the spread width narrows, the vertical spread payoff converges to the event contract payoff. The instruments are payoff-equivalent; the cost differential is the Vega Wedge.

$$IV = \frac{2}{T}\sum_i \frac{\Delta K_i}{K_i^2} e^{rT} O(K_i) \tag{1}$$

Where $O(K)$ is the out-of-the-money option price at strike K, and the sum spans available strikes weighted by inverse strike squared.

The empirical finding, replicated across markets and time periods: VRP is persistently positive.[1] Implied variance systematically exceeds realized variance. Investors pay a premium, the variance risk premium, to hedge variance exposure.

The magnitude is large. Bollerslev et al. (2009) found VRP explains over 15% of quarterly S&P 500 return variation. The premium is a systematic feature of options markets, not a transient anomaly.

Three explanations dominate the literature:

1. **Jump risk.** Investors fear discontinuous price movements that realized variance understates. Bollerslev et al. (2011) decomposed VRP into diffusive and jump components, finding the jump component dominates return predictability.

2. **Volatility-of-volatility.** Uncertainty about future volatility itself commands a premium. When volatility is volatile, variance hedges are more valuable.

[1] Carr and Wu (2009) define variance risk premium as RV − IV, measuring variance seller returns. We use IV − RV, measuring the cost to variance buyers. The empirical magnitude is identical; the sign reflects perspective.

3. **Model uncertainty.** Drechsler and Yaron (2011) showed that ambiguity about the correct probability distribution, or more formally, Knightian uncertainty, drives much of VRP. Investors pay to avoid model risk.

The VRP dialectic. VRP is simultaneously legitimate risk compensation and, for binary hedgers, a forced purchase of unwanted service. Vol-sellers genuinely bear uncertainty; the premium they extract is not rent in the pejorative sense. But binary outcome hedgers do not need variance protection – they need state protection. VRP compensates for a service they did not request and cannot unbundle from the instrument. Fair compensation from the vol-seller's perspective. Unnecessary cost from the binary hedger's. Both true. The question is whether an alternative instrument can serve the binary hedger without imposing variance risk transfer.

For binary outcome hedgers, VRP is thus a structural tax: not because vol-sellers are overcharging, but because the instrument bundles variance exposure the binary hedger does not want. Consider a fund hedging the probability of a Fed rate cut. The fund cares about one thing: does the Fed cut or not? It does not care about the path of Fed Funds futures between now and the meeting. It does not bear variance risk, only state risk.

Yet to hedge via options, the fund must purchase instruments that price the entire volatility path. The VRP component of that price is compensation for variance risk the fund does not bear. It pays for optionality it does not use.

The event contract strips this away. It prices the terminal state directly. No variance to bear because there is no path, only outcome. The VRP embedded in options replication does not exist in event contract pricing.

Quantifying VRP for Binary Payoffs

Translating VRP from variance space to binary contract space requires adjustment. Variance swap VRP reflects total variance exposure; the VRP relevant to binary replication reflects only the variance component of the vertical spread.

We measure VRP directly rather than borrowing literature estimates. The principle is consistent across categories: implied volatility from the relevant options market minus realized volatility from the underlying. The instruments differ. FOMC uses a novel decomposition of SOFR derivatives – Three-Month SOFR options (SR3) for implied volatility, Fed Funds futures (ZQ) for the clean policy signal – because Fed Funds options lack observable market activity (see Section IV.3 for why this absence is itself a finding about market structure). BTC uses Deribit's implied volatility index (DVOL) against realized spot returns. Other categories use the instruments natural to their markets. Category-specific methodology and decompositions are detailed in Section IV.

VRP is not a constant. It varies with genuine uncertainty across event categories, functioning as the segmentation variable identified in Section II.5. The gradient spans the eleven categories in our sample and extends beyond them.

A note on measurement precision: VRP levels are sensitive to annualization assumptions. We compute realized volatility over short windows centered on event announcements and annualize; implied volatility derives from options with varying tenors to expiry. Absolute VRP levels should be held loosely. The relative ordering across categories – the gradient – is robust to these assumptions.

Component 2: Dealer Balance Sheet Cost

The volatility surface is maintained by dealers who perform four functions:

1. **Warehousing.** Dealers hold inventory, taking the other side of client trades when no natural counterparty exists
2. **Interpolation.** Dealers quote prices at strikes and tenors where no trades have occurred, constructing the continuous surface from sparse observations
3. **Hedging.** Dealers manage the Greeks of their inventory through dynamic hedging in underlying and related instruments
4. **Intermediation.** Dealers connect buyers and sellers across time, bearing timing mismatch risk

This requires capital. Basel III, the Supplementary Leverage Ratio (SLR), and clearing mandates have made it expensive.

He and Krishnamurthy (2013) showed that when dealer capital constraints bind, risk premia rise nonlinearly. The marginal cost of balance sheet increases precisely when hedging demand is highest: during stress periods when clients most need protection.

Fleckenstein and Longstaff (2020) measured this directly. Comparing cleared versus uncleared derivatives, they estimated regulatory costs account for approximately 32 basis points of funding basis in interest rate markets. This is pure balance sheet cost: the price of regulatory capital consumed by dealer inventory.

For the SLR specifically, the cost is acute because netting across counterparties is prohibited for leverage ratio calculation. Gross exposures count against capital even when net exposure is zero. A dealer perfectly hedged on a risk-neutral basis still consumes capital for leverage purposes.

These costs embed in option prices. When a fund buys variance protection, it rents dealer balance sheet. The dealer must:

- Hold capital against the position (regulatory requirement)
- Fund the margin (opportunity cost)
- Manage the Greeks (operational cost)
- Bear the inventory risk (risk premium)

These costs pass through. The fund pays for intermediation whether it needs it or not.

Event contracts bypass this entirely. In a peer-to-peer event market:

- No warehousing: Trades match directly between participants
- No interpolation: Prices exist only where orders exist
- No hedging: Participants hold binary exposure, not Greeks
- No intermediation: The exchange matches and settles, nothing more

Counterparty risk is eliminated through full collateralization. Both sides of an event contract post their maximum loss upfront. If the outcome is 60% probability, the Yes buyer posts $0.60 and the No buyer posts $0.40. The sum is $1.00, the maximum payout. There is no credit exposure, no margin calls, no variation margin. The exchange holds the full payout amount from trade inception.

This eliminates the balance sheet function. No dealer capital consumed, no intermediation cost incurred. The balance sheet tax embedded in options pricing does not exist.

The per-event balance sheet cost is small in absolute terms – basis points scaled to days – but non-zero and structurally unavoidable in options markets while structurally absent in event markets. We estimate balance sheet cost using the Fleckenstein and Longstaff (2020) funding basis, scaled to event window duration. The methodology and estimates are detailed in Section IV.

Component 3: Replication Friction

Synthesizing a binary payoff from vanilla options incurs mechanical costs beyond VRP and balance sheet:

Transaction costs. The vertical spread requires two legs: long the lower strike call, short the higher strike call. Each leg incurs bid-ask spread. The total transaction cost is twice the single-leg spread.

Strike discreteness. Options trade at discrete strikes. The ideal replication requires $\varepsilon \to 0$, but available strikes force ε to remain positive. This creates approximation error: the spread payoff is not exactly binary but has a linear region between strikes.

Margin differential. Options positions require initial margin and daily variation margin. The margin requirement may exceed the economic exposure, especially for spread positions where margin is often calculated on gross rather than net exposure. Event contracts require collateral equal to maximum loss, economically efficient but potentially different from options margin.

Operational complexity. Managing a vertical spread requires monitoring both legs, rolling positions as expiration approaches, adjusting strikes if the underlying moves, and handling exercise/assignment. Event contracts are operationally simple: single position, automatic settlement, no management required.

Pin risk. When the underlying settles near strike prices at expiration, vertical spreads face assignment uncertainty. The hedger may end up with unintended positions requiring liquidation. This risk is elevated for tight spreads attempting to approximate digital payoffs. Event contracts have no pin risk; the outcome is binary by construction.

We estimate replication friction from observed bid-ask spreads and execution data in the relevant derivatives markets. As with balance sheet cost, the magnitude is small relative to VRP but structurally unavoidable. Estimates are reported in Section IV alongside VRP decompositions for each category.

III.3. The Total Structural Wedge

The structural Wedge is the sum of the three components:

$$W_{structural} = VRP_{binary} + C_{balance\ sheet} + F_{replication} \tag{2}$$

Where:

- VRP_{binary} = Variance risk premium tax on binary payoff replication
- $C_{balance\ sheet}$ = Dealer balance sheet cost embedded in options
- $F_{replication}$ = Transaction, approximation, and operational friction

Across all categories examined, VRP dominates the structural wedge, typically accounting for over 90% of the total. Balance sheet cost and replication friction are non-trivial but small relative contributions. This means the gradient identified in Section II.5 – high-VRP categories offer larger structural wedges – operates almost entirely through VRP variation. The wedge decomposition for each category is reported in Section IV.

The structural wedge varies across the VRP gradient. Low-VRP categories offer narrow wedges that prediction markets must overcome with very low execution costs. High-VRP categories offer wide wedges where even moderate prediction market liquidity suffices. The Seesaw tips first where the wedge is widest.

III.4. The Microstructural Penalty

The structural wedge is half the comparison. Event contracts may be structurally cheaper, but they must be accessed through markets that impose their own costs.

The microstructural penalty comprises:

Bid-ask spread. The cost of immediacy. Crossing the spread to execute incurs direct cost proportional to position size.

Market depth. The available liquidity at various price levels. Thin depth means large orders move prices against themselves.

Market impact. The permanent price effect of trading. Beyond temporary spread crossing, large trades may move the equilibrium price.

$$M_{event} = S_{spread} + I_{depth} + P_{impact} \tag{3}$$

Where:

- S_{spread} = Bid-ask spread cost
- I_{depth} = Depth-adjusted execution cost
- P_{impact} = Permanent price impact

This decomposition follows (Madhavan, 2000).

Measurement Approach

For spread: Mean absolute deviation of executed prices from contemporaneous mid-prices.

For depth: We construct effective depth curves from trade data. Without historical order book snapshots, we infer resting liquidity from the relationship between trade size and price deviation. This is a second-best approach; direct order book observation would be preferred.

For impact: We estimate using a simplified Kyle (1985) framework. The Kyle lambda measures price impact per unit of order flow:

$$\Delta P = \lambda \cdot Q \tag{4}$$

Where Q is order quantity and λ captures market depth. We calibrate λ from observed price movements around large trades, acknowledging this captures both temporary and permanent impact without clean separation.

Measurement Limitations

The microstructural analysis relies on trade data, not order book data. This creates several limitations:

1. **Depth is inferred, not observed.** We see what traded, not what was available to trade. Resting liquidity that was never hit remains invisible.
2. **Impact conflates components.** True market impact has temporary (spread/inventory) and permanent (information) components. Trade data alone cannot separate them cleanly.
3. **Selection bias.** Executed trades may not represent the full liquidity picture. Large institutional orders may be worked over time or executed OTC, never appearing in our data.

These limitations bias our execution cost estimates in unknown directions. The trajectory (costs falling dramatically over the sample) is likely robust, but the precise levels at any point carry uncertainty.

Extension: The Full Wedge Calculation

The framework as presented assumes derivative microstructure costs are negligible: that options markets are liquid enough that execution cost resides primarily on the PM side. This holds for FOMC (liquid SOFR derivatives) and BTC (liquid Deribit market). For categories where derivative markets are illiquid or require imperfect proxies, the comparison expands: PM execution cost versus structural wedge plus derivative execution cost plus hedge leakage from proxy imperfection. The full inequality is:

$$M_{PM} < W_{structural} + M_{derivative} + H_{leakage} \tag{5}$$

The primary categories in our sample – FOMC and BTC – have liquid derivative benchmarks where derivative-side execution cost is negligible. Other categories in the eleven-category dataset involve thinner derivative markets or proxy instruments, where derivative execution cost and hedge leakage become material. Section IV documents the derivative liquidity conditions for each category. The full inequality applies where derivative microstructure is non-trivial.

III.5. The Seesaw Condition

The Liquidity Seesaw tips when:

$$W_{structural} > M_{event} \tag{6}$$

When the structural wedge exceeds the microstructural penalty, event contracts become cheaper on a total-cost basis. Hedging flow should divert.

The structural wedge functions as a tax; the microstructural penalty as switching friction. The Seesaw tips when the tax exceeds the friction. Reflexive dynamics operate on both arms: prediction market depth responds to institutional flow, dealer maintenance intensity responds to flow loss. A general equilibrium model of these coupled dynamics is a natural extension that the empirical findings parameterize but that exceeds this paper's scope.

The condition is size-dependent. Small positions face proportionally smaller impact; large positions face proportionally larger impact. The threshold position size where the seesaw tips:

$$Q^* = \frac{W_{structural} - S_{spread}}{\lambda} \tag{7}$$

Where λ is the Kyle impact parameter. Positions below Q^* favor event contracts (if structural wedge exceeds spread); positions above Q^* may favor options (if impact exceeds structural advantage).

Threshold Estimation Approach

The Kyle-based theoretical approach proved difficult to calibrate with available data. Trade-by-trade price impact showed high variance and unclear relationship to position size, likely reflecting the conflation of liquidity-driven and information-driven trades.

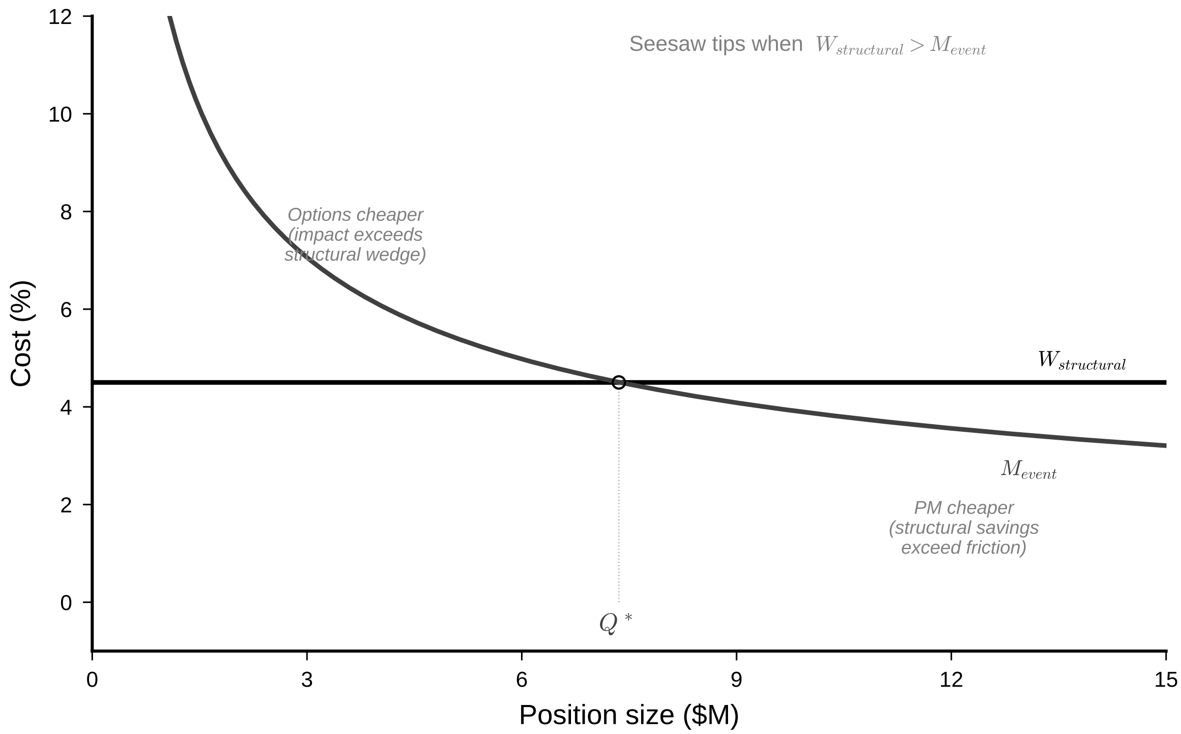

Figure 2. The Liquidity Seesaw. $W_{structural}$ (horizontal) represents the structural cost of options replication, invariant to position size. M_{event} (decreasing) represents PM execution cost, declining as position size falls relative to available depth. The crossing point Q^* identifies the position size at which event contracts become cheaper on a total-cost basis. High-VRP categories (wider $W_{structural}$) cross at larger position sizes; low-VRP categories require deeper PM liquidity to cross.

We adopt a simpler approach: construct effective depth curves from cumulative trade data and directly estimate execution cost at reference position sizes ($1M, $3M, $5M, $10M).

For each event:

1. Order trades by absolute price deviation from mid
2. Construct cumulative volume curve at each deviation threshold
3. Interpolate to find the deviation (execution cost) at reference position sizes
4. Compare total execution cost to structural wedge

This approach is data-intensive but relies on assumptions about trade flow reflecting resting liquidity. The Q^* derivation above motivates the size-dependent comparison. The empirical implementation evaluates the Seesaw condition directly at reference position sizes rather than estimating λ, because trade data without order book observation do not support clean separation of temporary and permanent impact components.

III.6. Threshold Status

The Seesaw condition is testable. For each event category, we measure the structural wedge and estimate execution cost at reference position sizes. The comparison yields a verdict: PM wins, PM loses, or marginal. Section V reports threshold status for each category across the VRP gradient.

IV. Data

IV.1. Category Selection

For event hedging via derivatives, the variance risk premium dominates structural costs: the compensation volatility sellers demand for bearing uncertainty around discrete catalysts. VRP determines category position on the Liquidity Seesaw. High-VRP categories offer larger structural wedges; low-VRP categories offer smaller ones. The segmentation hypothesis predicts that high-VRP categories cross the displacement threshold first.

Testing this requires categories spanning the VRP spectrum. Two polar cases confirm a direction; they cannot validate a gradient. Ordering is a claim about the full sequence, not the endpoints. Intermediate categories – where the threshold verdict is not predetermined by design – are where the hypothesis faces genuine falsification risk.

We analyze eighty-seven event-contracts across eleven categories spanning five asset classes. The selection criterion is joint: a liquid derivative instrument must exist for VRP measurement and an active prediction market must exist for execution cost comparison. Categories where only one side exists are excluded. Romania's 2025 presidential election generated $616M in prediction market volume but has no liquid country-ETF options chain for VRP extraction. Agricultural commodities have deep futures options markets but negligible prediction market coverage. The joint requirement constrains the testable universe but ensures every event-contract permits the full Seesaw comparison.

Cat.	Asset	Deriv.	Tier	Pred. VRP	PM Venue	Evts	VRP
FOMC	Rates	SR3 + ZQ	Tier 2	Low	Polymarket	12	6
BTC	Crypto	Deribit	Tier 1	High	Polymarket	20	6
ECB	Rates	6E FX options	Tier 3†	Moderate	Polymarket	3	3
BOJ	Rates	6J FX options	Tier 3†	Moderate-High	Polymarket	5	2
CPI	Macro	ZN Treasury	Tier 3	Low-Moderate	Kalshi	4	4
GDP	Macro	ZN Treasury	Tier 3	Variable	Kalshi	3	3
Gold	Commodities	GC options	Tier 1	High	Polymarket	3	3
Silver	Commodities	SI options	Tier 1	Negative*	Polymarket	3	1
Retail Gas	Commodities	RB options	Tier 3‡	High	Kalshi	4	4
Equity Indices	Equities	ES + NQ	Tier 1	Moderate	Kalshi	13	2
Elections	Political Risk	Country ETF options	Tier 3+§	High (predicted 3–5%)	Polymarket	17	17
Total						**87**	**51**

†FX proxy; measured VRP reflects full exchange-rate volatility, not rate-decision-specific component. Rate-specific VRP likely substantially lower. ‡Quality warnings on all retail gas events; stale option quotes at relevant tenors. §Country equity ETF proxy; heavier contamination than Tier 3. Route B standardization (SVEP) used as primary metric. See §IV.7. *Negative over measurement window; structural sign requires longer observation.

The "Independent VRP Obs." column distinguishes VRP degrees of freedom from PM liquidity observations. FOMC generates twelve event-contracts across six meetings, but VRP is computed once per meeting from the SR3/ZQ decomposition – the twelve contracts share six VRP values. Similarly, equity indices produce thirteen PM observations from two VRP measurements (one per index). Elections produce seventeen PM observations from approximately seventeen independent VRP measurements (one per event-period, each anchored on a distinct country ETF and resolution date). The eighty-seven event-contracts provide eighty-seven independent PM friction observations but approximately fifty-one independent VRP observations. This distinction matters for interpreting the gradient: the VRP spectrum is measured with $N \approx 51$, not $N = 87$.

The categories span the VRP gradient by design. FOMC represents the hard case: telegraphed decisions, deep derivative markets, minimal variance risk premium. BTC represents the favorable case: elevated and persistent volatility, substantial risk premium embedded in Deribit options. Between these poles, nine categories occupy intermediate positions. CPI and equity indices sit in the low-to-moderate range. International central banks (ECB, BOJ) occupy the moderate range but with proxy basis risk that inflates measurement noise. Gold and retail gas exhibit high VRP. Elections occupy the predicted high-VRP range with the heaviest proxy contamination in the sample. Silver serves as a structural control: negative VRP over our measurement window means the structural wedge does not exist, and the framework should correctly predict that prediction markets cannot win regardless of liquidity.

Not all categories carry equal evidentiary weight. The quality of the VRP-to-threshold comparison depends on two independent dimensions: the fidelity of the VRP measurement and the depth of the PM market.

Category	**Deriv. Quality**	**PM Quality**	**Reliability**
FOMC	Liquid (SR3/ZQ validated)	Deep ($96M–$365M per contract)	**High**
BTC	Liquid (DVOL, direct)	Variable ($237K–$33M per contract)	**High** for liquid contracts
Equity Indices	Deep (ES/NQ, direct)	Moderate ($1M–$4.3M)	**Moderate** (1–2 testable)
Gold	Liquid ATM (GC, direct)	Moderate ($1–$3.2M)	**Moderate** (1 clean event)
Silver	Liquid (SI, direct)	Moderate ($2.6–$2.9M)	**High** (structural control)
CPI	Strong proxy (ZN)	Thin ($64K–$2.3M)	**Low**
ECB	Contaminated proxy (6E)	Thin ($0.8–$1.1M)	**Low**
BOJ	Contaminated proxy (6J)	Thin ($1.5–$1.9M)	**Low**
GDP	Weak proxy (ZN)	Negligible ($145–$360K)	**Low**
Retail Gas	Stale quotes (RB)	Thin ($1.3–$2.2M)	**Low** (quality warnings)
Elections	Contaminated proxy (equity ETFs)	Variable ($295K–$115M)	**Low-Moderate** (8 headline-clean)

The gradient analysis in §IV.8 and the threshold tests in §V weight categories accordingly. FOMC, BTC, equity indices, gold, and silver provide the primary evidence – five categories with eighteen independent VRP observations and fifty-one event-contracts. CPI via Treasury options provides a defensible secondary observation. Elections provide a gradient-consistent result at lower measurement reliability, with eight headline-clean event-periods out of seventeen tested (§IV.7). ECB, BOJ, GDP, and retail gas are reported for completeness and gradient population but should not be weighted as equivalent to the direct-measurement categories. Their threshold verdicts are robust – proxy contamination overstates VRP, so PM-loses verdicts are conservative – but their gradient positions carry wider uncertainty bands than the table suggests.

Three measurement approaches are required, forming a provenance hierarchy that reflects decreasing fidelity.

Tier 1: Direct measurement. The derivative instrument prices the event risk without an intermediary correlation step. BTC via Deribit DVOL, equity indices via ES/NQ options, gold via GC options, silver via SI options. VRP is the spread between the derivative's implied volatility and realized volatility of the same underlying over the event window.

Tier 2: Validated decomposition. The SR3/ZQ decomposition addresses the FOMC measurement challenge developed in §IV.3, where the native instrument (Fed Funds options) does not emit observable prices. The methodology follows the approach used by the Federal Reserve Bank of Atlanta's Market Probability Tracker and has been validated by institutional practitioners. VRP is extracted from a proxy pair whose joint behavior around the event of interest is well characterized.

Tier 3: One-hop proxy. The nearest liquid derivative absorbs the event but also absorbs unrelated dynamics. ECB via EUR/USD options, BOJ via JPY/USD options, CPI and GDP via Treasury options, retail gas via RBOB gasoline options. The basis risk in each proxy chain is explicit and one-directional: proxy VRP overstates event-specific VRP because the proxy embeds volatility from non-event sources. PM-loses verdicts are conservative; gradient positions may be inflated.

All seventy event-contracts in the original ten categories carry the `vrp_adj` provenance label, indicating adjusted VRP computed through the full pipeline rather than raw or fallback estimates. No events were excluded from the pipeline; the seventy reported are the complete set of events satisfying the joint requirement across the sample period. Seventeen additional event-contracts cover contested elections across twelve countries, tested through the Route B methodology described in §IV.7.

IV.2. Data Sources

The pipeline ingests data from four sources for the original ten categories, producing a combined dataset of 2,889,424 rows across seventy event-contracts. Seventeen additional election event-contracts are sourced through Polymarket trade data and OPRA equity ETF options chains across ten ETF proxies (§IV.7).

Prediction Market Data. Polymarket provides trade-level data for 47 events via the Goldsky subgraph. Kalshi provides trade-level data for 23 events via the Kalshi API. Both sources record timestamp, price, size, and side at the individual trade level. Combined prediction market data spans November 2024 to February 2026 across 1.67 million executed trades.

Derivatives Data. Databento provides futures and options data for ten CME and CBOT symbols spanning rates (SR3, ZQ), foreign exchange (6E, 6J), equity indices (ES, NQ), commodities (GC, SI, RB), and treasuries (ZN). Deribit provides the DVOL index for BTC, a model-free implied volatility measure analogous to VIX. Time span: September 2024 to February 2026.

Onboarding Validation. All seventy events in the original ten categories pass pipeline validation. The validation gate requires: (1) non-missing VRP computation from derivative data spanning the event window, (2) prediction market trade data present with at least one executed trade, (3) structural wedge

Source	Rows	Events Covered	Date Range
Polymarket (Goldsky)	1,625,309	47	Nov 2024 – Feb 2026
Kalshi	49,473	23	Jan 2024 – Feb 2026
Databento (10 symbols)	1,213,860	50	Sep 2024 – Feb 2026
Deribit DVOL	782	20	Nov 2024 – Feb 2026
Total	**2,889,424**	**70**	**Jan 2024 – Feb 2026**

computable from VRP plus the balance sheet and friction estimate described in §IV.2.1, (4) derivative data coverage spanning the full event window. No candidate events were excluded pre-pipeline; the seventy are the complete set satisfying the joint requirement across the sample period. The seventeen election event-contracts pass a separate validation gate documented in §IV.7, including event isolation and Route B standardization checks. For both prediction market venues, we observe executed trades, not the order book. Depth and market impact are inferred from trade flow rather than observed from resting liquidity.

IV.2.1. From VRP to Structural Wedge

The structural wedge – the total cost embedded in derivatives that prediction markets can potentially undercut – consists of three components per the §III.2 decomposition: variance risk premium, balance sheet cost, and replication friction. The pipeline computes:

$$W_{structural} = VRP_{adj} + C_{balance\ sheet} + F_{replication} \approx VRP_{adj} + 0.07\% \tag{8}$$

VRP_{adj} is the category-specific adjusted variance risk premium, computed as described in §IV.3–IV.6. The balance sheet and replication friction components are estimated at approximately 7 basis points. The balance sheet cost derives from the Fleckenstein and Longstaff (2020) funding basis – the spread between Treasury and derivatives-implied rates that reflects the cost of dealer balance sheet usage – scaled to event-window duration. The replication friction reflects the two-leg spread cost of constructing a binary payoff from vanilla options at institutional execution quality.

VRP accounts for approximately 93% of the total structural wedge across categories where VRP is positive. The 7bp estimate is conservative and applied uniformly: for rate derivatives where dealer balance sheets are optimized for the product, the true cost may be lower; for exotic or illiquid derivative categories, it may be higher. The gradient analysis uses the full structural wedge. The VRP column in the gradient table (Table IV.22) is reported because it is the dominant and variable component; the balance sheet and friction contribution shifts all categories upward by the same constant and does not affect the ordering.

IV.3. FOMC Data

Event Selection

We analyze twelve event-contracts across six FOMC meetings spanning November 2024 to January 2026. Events were selected based on: (1) availability of complete Polymarket trade data, (2) sufficient prediction market volume for microstructure analysis, and (3) variation in market uncertainty and policy outcomes to test the framework across conditions.

Unlike single-contract event categories, FOMC meetings generate multiple prediction market contracts per date. The January 2026 meeting has separate contracts for a 25 basis point cut, a 50 basis point cut, a 25 basis point hike, and no change. Each contract constitutes an independent observation for microstructure analysis. VRP is identical across contracts at the same meeting because it derives from the SR3/ZQ decomposition at that date. What varies is prediction market liquidity: the January 2026 cut-50bp contract attracted $365M in volume while the no-change contract attracted $134M. This separation of structural wedge (constant within meeting) and PM friction (variable across contracts) isolates the liquidity channel of the Seesaw while holding the wedge fixed – a natural experiment that no other category in our sample provides with equal clarity.

Meeting	Contracts	Decision	Uncertainty	Total PM Volume
Nov 2024	1	Cut 25bp	Moderate	$258M
Jan 2025	1	Hold	Low	$96M
Sep 2025	2	Hold	Moderate	$204M
Oct 2025	1	Hold	High	$164M
Dec 2025	3	Cut 25bp	Low-Moderate	$479M
Jan 2026	4	Hold	Low	$1,007M
Total	**12**			**$2,208M**

The sample varies on three dimensions. Event uncertainty: October 2025 featured genuine policy ambiguity; January 2026 was well telegraphed. Market development: the January 2025 meeting generated $96M in PM volume; twelve months later, January 2026 generated over $1 billion across four contracts. Contract multiplicity: later meetings support analysis of how PM liquidity distributes across outcomes at the same event, unavailable in single-contract categories.

The volume trajectory is itself a finding. One billion dollars in prediction market flow on a single FOMC meeting – on an offshore platform, with no institutional custody infrastructure – is not noise. It is demand for the instrument this paper describes. The plumbing has not arrived. The demand has.

The Measurement Challenge

The native instrument for hedging binary FOMC risk is the Fed Funds options contract (OZQ), listed on CME. If this market were liquid and transparent, we would measure VRP directly: implied volatility from Fed Funds options minus realized volatility from Fed Funds futures.

It is neither. By standard measures, it does not exist as an observable market.

Databento returns zero rows for Fed Funds options order book data across our entire sample period. The CME Interest Rate Liquidity Review omits the product entirely. CFTC Bank Participation Reports exclude it; fewer than five banks hold reportable positions. The physical options trading pit closed in May 2021. Industry sources describe volume as "tiny" and pricing as "by appointment."

The contract exists, is listed, and clears through CME. But trades occur bilaterally, via voice, between dealers who know each other. It does not instantiate the objects that make something observable as a market: continuous quotes, visible depth, public price formation.

The market continues to exist for participants while ceasing to exist for observers.

Taxis at its boundary. The apparatus is maintained privately; pricing happens inside relationships rather than on screens. The market functions for those inside – dealers hedge, institutions transact – but emits no public traces. The absence is structurally predictable. Constructed order requires maintenance. When flow thins below the threshold where maintenance is profitable, dealers retreat to bilateral negotiation.

§V develops this as the Observability Wedge. Here, we note only the methodological consequence: direct VRP measurement is impossible. We require proxies. The measurement challenge varies in degree across our eleven categories. FOMC is the extreme case, where the native instrument emits nothing observable at all. Elections face a different extreme: no election-specific derivative instrument exists at all, requiring country equity ETF proxies that absorb 65–90% non-election dynamics (§IV.7). The Tier 3 categories face a softer version of the same problem: the native instrument either does not exist or lacks the liquidity to support VRP extraction, requiring the use of a correlated derivative that absorbs the event but also absorbs other dynamics. The Tier 1 categories avoid this problem entirely. The gradient of measurement difficulty across categories is itself informative, and maps onto the gradient of institutional entrenchment that the Observability Wedge describes.

The SR3/ZQ Decomposition

We construct VRP from two observable instruments that bracket the unobservable target.

SR3: Three-Month SOFR Options. The Secured Overnight Financing Rate is the benchmark for short-term dollar funding. SR3 options are liquid, actively traded, and provide observable implied volatility. Around FOMC events, SOFR responds primarily to the policy decision – the same driver as Fed Funds. We follow the approach used by the Federal Reserve Bank of Atlanta's Market Probability Tracker for extracting policy expectations from SOFR-linked instruments.

ZQ: Fed Funds Futures. Fed Funds futures settle to the effective Fed Funds rate and reflect policy expectations directly. They do not contain the repo market dynamics embedded in SOFR. Realized volatility from ZQ isolates the policy component.

The Decomposition. The adjusted VRP estimate combines these instruments through a multi-step procedure:

$$VRP_{adj} \approx f(IV_{SR3}, RV_{ZQ}, \text{repo adjustment}) \tag{9}$$

where IV_{SR3} is implied volatility extracted from SR3 options via Black-76 inversion, and RV_{ZQ} is realized volatility from ZQ futures calculated from 5-minute midpoint returns over the $T-1$ to $T+1$ window. The pipeline applies a repo noise adjustment that accounts for the systematic component of the SR3-ZQ basis predictable from repo market conditions; consequently, VRP_{adj} does not equal the simple subtraction $IV_{SR3} - RV_{ZQ}$. The raw decomposition is also computed and stored; the adjustment is modest (mean magnitude approximately 10bp) and does not affect the ordering of any threshold verdict. The full adjustment procedure, including the Black-76 parameterization, is documented in the Methodology Annex.[2]

Isolating Repo Noise. SOFR embeds financing dynamics unrelated to monetary policy. Quarter-end effects, repo market stress, and collateral scarcity all affect SOFR independently of Fed decisions. Using SOFR realized volatility directly would contaminate the VRP estimate with noise the binary hedger does not bear. The decomposition isolates this noise:

$$\text{Repo Noise} = RV_{SR3} - RV_{ZQ} \tag{10}$$

The assumption underlying this approach is that ZQ realized volatility cleanly proxies the policy component of SOFR variance. This assumption is strongest around FOMC events specifically, when both instruments respond to the same Fed decision. Repo dynamics do not move on FOMC announcements; policy moves both rates. We claim validity only for event windows. For non-FOMC periods, the assumption is weaker and the decomposition is not applied.

Basis Risk. SR3 reflects 3-month SOFR expectations; ZQ reflects the monthly Fed Funds effective rate. Tenor mismatch and SOFR-Fed Funds spread dynamics introduce basis risk that the decomposition cannot fully eliminate. We estimate residual basis risk at 5 to 10 basis points, derived from historical SR3-ZQ spread volatility computed over FOMC event windows across the 2023–2025 period (twelve events outside our sample used for calibration). This is small relative to the VRP estimates but non-zero.

Full implementation details – Black-76 parameterization, strike selection rules, return calculation windows, annualization factors, tick conversion – are provided in the Methodology Annex.

VRP Results

Meeting	SR3 IV (%)	ZQ RV (%)	VRP_{raw} (%)	VRP_{adj} (%)
Nov 2024	0.618	0.324	0.295	0.445
Jan 2025	0.618	0.206	0.411	0.543
Sep 2025	0.887	0.409	0.478	0.697
Oct 2025	0.719	0.453	0.266	0.551
Dec 2025	0.711	0.209	0.502	0.417
Jan 2026	0.518	0.166	0.353	0.440

[2] ATM strike is defined as the listed strike nearest to the $T-1$ settlement price. For SR3Z4 with settlement at 95.48, the ATM strike is 95.50. No interpolation between strikes is applied; the nearest-strike convention is standard in short-term rate options and avoids introducing model-dependent smoothing at the IV extraction stage.

Statistic	Value
VRP_{adj} mean	0.516%
VRP_{adj} median	0.498%
VRP_{adj} range	0.417% – 0.697%
Independent VRP observations	6
Event-contracts	12

[3] [4]

The 0.52% mean VRP confirms FOMC as a low-VRP category. The equity index options literature estimates VRP at 2 to 4 percent; FOMC sits well below this range, consistent with the structural prediction: Fed communication reduces genuine uncertainty, derivatives are liquid and well-arbitraged, and the structural wedge available for prediction markets to capture is small. This is the hard case for displacement.

Within the sample, VRP varies with event uncertainty. September 2025 produced the highest VRP (0.697%): the policy path was genuinely contested. January 2026 produced among the lowest (0.440%): the hold was fully telegraphed weeks in advance. The gradient is visible even within a single category. Higher-uncertainty events show larger structural wedges, supporting the segmentation hypothesis at the event level, not just the category level.

Sensitivity. The Methodology Annex reports FOMC VRP under alternative specifications varying the RV window, annualization factor, return frequency, and repo noise adjustment. Under all specifications, FOMC mean VRP remains below 1.0% and the category's position as the low end of the gradient is invariant. The gradient ordering across all eleven categories is robust to specification choice.

IV.4. BTC Data

Event Selection

We analyze twenty event-contracts spanning November 2024 to January 2026. Events were selected based on: (1) BTC price approaching or crossing significant thresholds ($100K, $110K, $125K, $130K, $150K, $200K, $250K), (2) sufficient prediction market volume for microstructure analysis, and (3) availability of matched Deribit DVOL data for VRP calculation.

BTC touch markets are structured as "Will Bitcoin hit [price] by [date]?" Each time horizon offers contracts at multiple strike levels. The January 2026 expiry has separate contracts at $100K, $105K, $110K, $125K, and $150K. This multi-strike structure creates a synthetic options chain: implied

[3] The repo noise adjustment produces VRP_{adj} values that differ from the raw $IV_{SR3} - RV_{ZQ}$ subtraction. The adjustment procedure accounts for the systematic component of the SR3-ZQ basis predictable from repo market conditions. December 2025 illustrates the effect: VRP_{raw} of 0.502% is adjusted downward to 0.417% because elevated year-end financing dynamics inflated the SR3-ZQ basis beyond its policy-driven component. The full adjustment methodology is documented in the Methodology Annex.

[4] Mean and summary statistics computed from unrounded pipeline output. Meeting-level estimates are reported to three decimal places for precision; rounded presentation values (e.g., 0.52%) match the unrounded computation to within rounding tolerance.

probabilities at discrete strikes that together describe the market's expected distribution. It is unique among our categories. FOMC contracts are binary on a specific outcome, as are CPI, GDP, and most equity index contracts. BTC's structure permits analysis of how PM execution costs vary with moneyness – a dimension exploited in the smile analysis in §V.

As with FOMC, VRP is computed at the time-horizon level rather than the contract level. All January 2026 contracts share the same DVOL snapshot and spot realized volatility; the VRP is identical across strikes within a horizon. What differs is the PM execution cost: the $150K touch contract with $33M in volume has materially lower friction than the $125K contract with $2.2M. This separation of VRP (constant within horizon) and PM friction (variable across strikes) isolates the liquidity channel of the Seesaw while holding the structural wedge fixed – the same natural experiment that FOMC provides, replicated in a high-VRP category.

Horizon	**Contracts**	**Strike Range**	**DVOL (%)**	**Total PM Volume**
Nov 2024	1	$100K	55.2	$32M
Dec 2024	2	$100K–$110K	60.0	$10M
Annual 2025	4	$130K–$250K	43.0	$58M
Oct 2025	3	$130K–$200K	43.3	$29M
Nov 2025	5	$115K–$200K	48.8	$28M
Jan 2026	5	$100K–$150K	47.6	$62M
Total	**20**			**$219M**

The sample captures varied market conditions: post-election rally (November and December 2024), range-bound consolidation (October 2025), and elevated two-way flow (November 2025). This variation tests whether the Vega Wedge framework holds across volatility regimes, not just in favorable conditions.

The within-horizon variation in PM volume is large and analytically important. The January 2026 horizon illustrates: the $150K touch contract attracted $33M in PM volume (execution cost at $3M: 0.13%); the $125K touch contract attracted $2.2M (execution cost at $3M: 13.8%). Same VRP, same structural wedge, sixty-fold difference in PM friction. §V exploits this variation to test whether PM liquidity determines threshold status within categories where the wedge is held constant.

Horizon	**Strike**	**PM Volume**	**Exec Cost 3M (%)**	**Structural Wedge (%)**
Jan 2026	$100K	$13.3M	0.93	4.08
Jan 2026	$105K	$7.1M	0.86	4.08
Jan 2026	$110K	$4.9M	0.85	4.08
Jan 2026	$125K	$2.2M	13.83	4.08
Jan 2026	$150K	$32.8M	0.13	4.08
Nov 2025	$115K	$5.7M	0.99	4.29
Nov 2025	$120K	$2.9M	10.60	4.29
Nov 2025	$130K	$2.0M	15.12	4.29
Nov 2025	$150K	$2.1M	14.58	4.29
Nov 2025	$200K	$17.1M	0.21	4.29

The pattern is clear: contracts above approximately $5M in PM volume have execution costs well below the structural wedge. Contracts below $3M have execution costs that exceed the wedge by multiples. The threshold is not a function of VRP – all contracts within a horizon share the same wedge – but of PM liquidity. This is the Seesaw mechanism operating at the contract level: same structural arm, different friction arm, different verdict.

Derivatives Data and Methodology Comparison

The BTC VRP calculation is direct:

$$VRP = DVOL - RV_{spot} \tag{11}$$

where DVOL is the Deribit Volatility Index at the period start – a model-free 30-day implied volatility measure analogous to VIX, constructed via the Carr-Madan variance swap replication – and RV_{spot} is realized volatility from BTC spot returns over the corresponding window, computed from hourly returns.[5]

No proxy chain is required. The same underlying asset prices risk on both venues – Deribit for derivatives and Polymarket for prediction markets – and the comparison is unmediated by any basis.

Dimension	**FOMC**	**BTC**
Native instrument	Unobservable (Fed Funds options)	Observable (Deribit)
Proxy required	Yes (SR3/ZQ)	No
IV source	SR3 options via Black-76	DVOL (direct, model-free)
RV source	ZQ futures (5-min midpoints)	Spot returns (hourly)
Decomposition steps	7	2
Noise isolation required	Yes (repo dynamics)	No
Provenance tier	Tier 2 (validated decomposition)	Tier 1 (direct)

BTC permits the comparison the FOMC methodology was designed to approximate. The Vega Wedge is directly observable as the gap between what the prediction market charges and what Deribit embeds in VRP. The methodological asymmetry is analytically valuable. FOMC proves the apparatus works when direct measurement is impossible. BTC proves the magnitude when direct measurement is possible. Together they validate the framework from opposite ends of the measurement difficulty spectrum – and the seven-step and two-step methodologies produce results at exactly the gradient positions the theory predicts.

VRP Results

BTC VRP Summary: Mean +4.83%, median +4.10%, range −4.61% to +15.65%. Positive VRP in five of six horizons. Six independent VRP observations; twenty event-contracts.

[5] DVOL is a 30-day forward-looking measure. For horizons exceeding 30 days (Annual 2025), the DVOL snapshot at period start is compared against RV computed over the full calendar window of the contract. The comparison captures the VRP level at the horizon start date. For longer horizons, term-structure effects mean the 30-day measure may not perfectly represent implied vol for the full period. This is standard in the VRP literature (Carr and Wu, 2009) and does not affect the category-level gradient position, which is the primary claim.

Horizon	**DVOL (%)**	RV_{spot} **(%)**	**VRP (%)**
Nov 2024	55.2	59.8	−4.61
Dec 2024	60.0	44.4	+15.65
Annual 2025	43.0	34.3	+8.70
Oct 2025	43.3	42.3	+1.00
Nov 2025	48.8	44.6	+4.20
Jan 2026	47.6	43.6	+4.00

The 4.83% mean VRP is approximately an order of magnitude above FOMC (0.52%). This confirms BTC as a high-VRP category. The structural wedge available for prediction markets to capture is large and persistent: five of six horizons show positive VRP.

November 2024 is the single negative-VRP horizon at −4.61%: the post-election rally generated realized volatility that exceeded the market's implied expectation. VRP can be negative in individual periods while remaining positive in expectation. Vol-sellers lost money that month. The structural compensation for bearing crypto variance risk is a distributional property, not a period-by-period guarantee.

December 2024 shows extreme positive VRP at +15.65%: implied volatility priced a continuation of rally momentum that did not fully materialize. The range from −4.61% to +15.65% illustrates the volatility of the premium itself. Point estimates should be held loosely. The ordering – BTC high, FOMC low – is robust to any reasonable windowing or specification.

The gradient between FOMC and BTC spans two orders of magnitude in structural wedge. FOMC offers 0.42% to 0.70%. BTC offers −4.61% to +15.65%. The segmentation hypothesis predicts this difference determines threshold status. §V tests whether it does.

IV.5. Direct-Measurement Categories

Four categories beyond FOMC and BTC permit direct VRP measurement (Tier 1): the derivative instrument prices the same underlying risk that the prediction market contract references, without an intermediary correlation step. IV is extracted via Black-76 inversion from the relevant futures option chain using ATM strikes (nearest listed strike to settlement price at $T-1$; no interpolation between strikes). RV is computed from the underlying futures or spot returns over the event window. The structural wedge follows directly.

A general measurement limitation applies to out-of-the-money contracts. The pipeline extracts ATM implied volatility. For contracts near the money, ATM VRP is the appropriate measure. For contracts deep out of the money, the relevant risk premium is the tail component – the compensation for bearing variance in the wings of the distribution – which the literature places at multiples of ATM VRP – conservatively 1.5 to 2.5× at 10-delta (Bollerslev et al., 2011). Several prediction market contracts in equity indices and gold reference deep OTM barriers (e.g., "S&P 500 below 3,300" when the index trades at 5,700). The ATM VRP applied to these events understates the structural wedge.

This understatement affects interpretation more than threshold verdicts. The OTM equity events that lose in §V lose on PM liquidity: thin markets with \$1–2.3M in volume. With properly measured tail VRP, these events would appear wedge-sufficient but liquidity-constrained – the structural wedge exceeds PM friction, but PM friction is too high to demonstrate it because the PM market is too thin. The distinction matters for the Seesaw mechanism: the binding constraint for OTM equity events is depth, not the size of the wedge. Their gradient positions in Table IV.22 should be read as lower bounds on the true structural wedge.

Equity Indices

Thirteen event-contracts cover S&P 500 and Nasdaq 100 year-end outcomes on Kalshi, measured against ES and NQ futures options. Contracts are structured as barrier outcomes: "S&P 500 above 5,700 at year-end," "Nasdaq below 13,250 at year-end," and similar.

Index	**Events**	**ATM IV (%)**	**RV (%)**	**VRP (%)**
S&P 500 (ES)	8	14.30	12.24	2.06
Nasdaq 100 (NQ)	5	18.77	17.03	1.74

VRP of 1.74% to 2.06% is consistent with the variance risk premium literature and confirms that the pipeline produces textbook results for well-studied asset classes. All events within each index share the same VRP because they reference the same year-end expiry and option chain. The thirteen event-contracts represent two independent VRP observations and thirteen PM liquidity observations. The structural wedge is held constant while PM friction varies across strikes, isolating the liquidity channel.

Gold

Three event-contracts cover gold price targets on Polymarket, measured against GC futures options.

Event	**IV (%)**	**RV (%)**	**VRP (%)**	**PM Volume**
Gold $\geq$\$3,200 EOY 2024	20.8	13.6	7.22	\$3.2M
Gold $\geq$\$3,200 EOY 2025	32.5	18.2	14.33	\$1.4M
Gold hits \$5,500 by Jan 2026	64.0	44.4	19.57	\$1.0M

Gold VRP ranges from 7.22% to 19.57%, placing it in the high-VRP regime alongside BTC. The Gold $\geq$\$3,200 EOY 2024 event is the cleanest measurement: gold traded near \$2,600 at observation, making the target moderately out of the money but not extreme. It is also the only gold event where PM liquidity is sufficient for a meaningful Seesaw comparison (\$3.2M in volume). The remaining events combine high VRP with thin PM liquidity. The binding constraint is liquidity, not the wedge.

The Gold \$5,500 target illustrates the moneyness limitation. With spot at approximately \$2,600, this is a deep out-of-the-money barrier. ATM VRP of 19.57% already suggests a wide structural wedge. The wing VRP at the relevant moneyness – roughly analogous to a 10-delta call – would be substantially higher, conservatively 1.5 to 2.5 times the ATM figure (Bollerslev et al., 2011). The gradient position of gold's upper events is compressed by ATM measurement; the true structural wedge is wider than reported.

Silver

Three event-contracts cover silver price targets on Polymarket, measured against SI futures options.

Event	IV (%)	RV (%)	VRP (%)	PM Volume
Silver ≥$110 Jan 2026	42.9	53.1	−10.19	$2.8M
Silver ≥$120 Jan 2026	42.9	53.1	−10.19	$2.9M
Silver ≥$150 Jan 2026	42.9	53.1	−10.19	$2.6M

Silver is the structural control. VRP is −10.19% over our measurement window: realized volatility substantially exceeded implied, meaning options were cheap rather than expensive. Vol-sellers lost money. All three events share the same VRP because they reference the same option chain – this is one independent VRP observation with three PM liquidity observations.

The framework correctly identifies this regime. With negative VRP, the structural wedge does not exist. There is no embedded cost in derivatives for prediction markets to undercut. Even with infinite PM liquidity, prediction markets cannot offer a cost advantage when derivatives do not overcharge. §V reports the threshold classification as PM-loses for all three events; the loss is structural rather than frictional.

A caveat on the sign: the −10.19% is measured over a single window. Silver can exhibit positive VRP in other periods – commodity vol surfaces are regime-dependent and the structural sign of silver VRP requires longer observation to establish with confidence. What the framework demonstrates here is that *given* negative VRP in a measurement window, PM cannot win. The finding is conditional, not a permanent classification of silver. If the framework only identified categories where PM wins, selection bias would be a concern. Silver demonstrates that it discriminates: it identifies categories where PM cannot win because the economic precondition for displacement is absent.

IV.6. Proxy Categories

Five categories require one-hop proxy measurement (Tier 3). The native derivative instrument either does not exist in observable form or lacks the liquidity to support VRP extraction. Instead, VRP is measured from the nearest liquid derivative that absorbs the event. This introduces basis risk: the proxy reacts to the catalyst but also reacts to dynamics unrelated to it.

The proxy contamination is substantial and one-directional. The measured VRP overstates event-specific VRP because the proxy derivative embeds volatility from non-event sources. PM-loses verdicts are conservative: if the inflated proxy VRP is insufficient for PM to win, the true event-specific VRP is even more insufficient. But gradient positions for Tier 3 categories are inflated relative to what event-specific VRP would show. We report these categories for gradient population and completeness, not as equivalent evidence to Tier 1 and Tier 2 categories. The reader should evaluate the gradient primarily against the five high-reliability categories.

A structural observation applies to all five Tier 3 categories: PM liquidity at current volumes is insufficient for a meaningful Seesaw comparison regardless of where the true VRP sits. ECB prediction markets have \$0.8M to \$1.1M in volume. CPI contracts outside the headline release have under \$400K. Whether ECB VRP is the measured 2.01% or a more plausible rate-specific figure of 0.3–0.6% does not change the threshold verdict. The proxy noise affects gradient positioning, not threshold classification.

Central Banks: ECB and BOJ

ECB rate decisions on Polymarket are measured against 6E (EUR/USD futures) options. BOJ rate decisions on Polymarket are measured against 6J (JPY/USD futures) options. The proxy structure is identical: the central bank decision moves the exchange rate, and FX options absorb that move. But the exchange rate also absorbs dollar dynamics, risk appetite, trade flows, carry positioning, and intervention expectations. The rate decision is one driver among many.

The contamination is severe enough to warrant explicit quantification of the mismatch. An ECB cut of 25bp might move EUR/USD 30–50 pips on a typical announcement day. The same currency pair routinely moves 100 pips on an unrelated geopolitical headline. FX VRP around an ECB meeting reflects the full distribution of potential EUR/USD moves – of which the ECB decision is perhaps 25–40% of the variance. If native Euribor options were observable through our data infrastructure, the rate-decision-specific VRP would likely sit at 0.3–0.6% for ECB – essentially FOMC territory. Similarly, the BOJ January 2026 reading of 5.03% is dominated by JPY/USD volatility around the Bank of Japan's rate normalization path; the BOJ binary decision itself may contribute 1–2% of the measured figure.

Event	Proxy	Proxy IV (%)	Proxy RV (%)	VRP (%)	PM Volume
ECB Sep 2025	6E	8.40	6.39	2.01	\$1.0M
ECB Oct 2025	6E	7.19	4.45	2.74	\$1.1M
ECB Dec 2025	6E	5.03	3.47	1.56	\$0.8M
BOJ Dec 2025 (3 contracts)	6J	8.95	6.72	2.23	\$1.6–1.9M
BOJ Jan 2026 (2 contracts)	6J	12.26	7.23	5.03	\$1.5–1.6M

These categories appear in the gradient table with explicit notation that VRP reflects full proxy-instrument volatility. The VRP column for ECB and BOJ should not be read as a claim about the variance risk premium of rate decisions. It is a claim about the FX volatility that happened to coincide with a meeting date.

Macro Releases: CPI and GDP

CPI releases on Kalshi are measured against ZN (10-year Treasury note futures) options. The proxy is stronger than FX-to-central-bank: CPI is a primary driver of Treasury volatility, and a hot print moves the long end directly. The transmission channel is tighter.

Event	**ZN IV (%)**	**ZN RV (%)**	**VRP (%)**	**PM Volume**
CPI YoY Apr 2025	6.11	5.47	0.63	$2.3M
CPI YoY May 2025	7.52	6.24	1.28	$0.4M
CPI YoY $\geq$3% Jun 2025	5.06	4.01	1.05	$0.2M
CPI Jul 2025	6.95	4.82	2.13	$64K

CPI VRP of 0.63% to 2.13% places it in the low-to-moderate range, close to FOMC. This is consistent with the thesis: CPI releases are partially anticipated by consensus estimates, and Treasury options are liquid enough that vol-sellers compete aggressively. PM volume is thin across all CPI events; the April 2025 event has the most liquidity at $2.3M, still below the $3M institutional reference size.

GDP releases on Kalshi are similarly measured against ZN options but with a weaker transmission channel – GDP's impact on Treasuries is slower and more diffuse than CPI's – and negligible PM volumes ($145K to $360K). VRP is variable: -2.48% for Q1 2025 (where the print surprised and realized vol exceeded implied), +2.15% for Q3 2025, +3.78% for Q2 2025. At these PM volumes, execution cost estimation is unreliable and the Seesaw comparison is not meaningful. GDP is included for gradient completeness; the threshold verdicts (all PM-loses on high friction) reflect negligible PM depth, not GDP's position on the VRP gradient.

Retail Gas

Gas price contracts on Kalshi are measured against RB (RBOB gasoline futures) options.

Event	**RB IV (%)**	**RB RV (%)**	**VRP (%)**	**PM Volume**
Gas Apr 2025	50.0	36.9	13.1	$1.3M
Gas May 2025	34.2	24.5	9.7	$2.0M
Gas Jun 2025	44.9	41.2	3.7	$2.2M
Gas Jul 2025	53.6	22.4	31.2	$1.9M

Retail gas VRP ranges from 3.7% to 31.2%, the widest spread of any category. **All four events carry quality warnings.** The pipeline flagged stale or one-sided quotes near the event date on every retail gas observation, indicating that RB option liquidity was thin at the relevant tenors. The July 2025 reading of 31.2% – RBOB implied vol of 53.6% against realized vol of 22.4% – is almost certainly inflated by stale option quotes rather than genuine risk premium of that magnitude. A stale vol surface is not a priced vol surface; it is a number that has not been updated to reflect current market conditions. The gradient table reports retail gas VRP with an explicit quality caveat. Point estimates should not be treated as reliable measures of gasoline variance risk premium.

IV.7. Elections Data

Contested elections are the predicted extension from §II.5: a theoretically high-VRP category – approximately 3 to 5 percent – where prediction markets have achieved exceptional depth. Seventeen event-periods across twelve countries test the prediction using country equity ETF options as the derivative proxy.

The Proxy Chain

Election outcome → country equity market reaction → country equity ETF → ETF options chain → ATM IV → VRP.

This is heavier contamination than any Tier 3 category in the main sample. Country equity ETFs absorb macro conditions, trade policy, global risk appetite, sector rotation, and currency dynamics alongside the election. The election is one driver among dozens.

The pipeline quantifies the contamination through an election event premium:

$$EP_e = IV_e^+ - IV_e^- \tag{12}$$

where IV_e^+ is implied volatility on the nearest expiry after the election and IV_e^- is implied volatility on the nearest expiry before it. The ratio EP_e/IV_e^{pre} measures the election's share of total pre-election IV.

Event	**Country**	**ETF Proxy**	**ep_ratio**	**Election Share of IV**
IRL PM 2024	Ireland	EWG*	0.345	34.5%
NOR Parliament 2025	Norway	EWU*	0.338	33.8%
PRT PM 2025	Portugal	EWG*	0.259	25.9%
DEU Chancellor 2025	Germany	EWG	0.239	23.9%
ARG Deputies 2025	Argentina	ARGT	0.183	18.3%
PRT President 2026	Portugal	EWG*	0.168	16.8%
DEU Parliament 2025	Germany	EWG	0.132	13.2%
IRL President 2025	Ireland	EWG*	0.116	11.6%

Eight of seventeen event-periods pass the event isolation threshold ($EP_e/IV_e^{pre} \geq 0.10$). The threshold is conservative: no event-period produces an ep_ratio between 0.08 and 0.116, and the qualitative findings (twelve PM-wins, five losses) are invariant to perturbation of the cutoff within this gap. Even on these headline-clean events, the election accounts for 12–35% of total IV. Nine event-periods fail: IV moved opposite to the election premium prediction on most, meaning the proxy is not transmitting election-specific risk. Failed events are reported in the full table but excluded from the headline SVEP distribution.

**Proxy-substituted.* EIRL, PGAL, and ENOR are unavailable in OPRA; EWG and EWU serve as substitutions, adding a second layer of proxy noise.

Route B: Standardized Comparison

Raw VRP from country equity ETFs is not comparable across countries or to other categories in the sample. Poland's EPOL has baseline IV of 102.5%; South Korea's EWY has baseline IV of 35.7%. Route B normalizes both sides of the Seesaw by each event-period's own baseline implied volatility:

$$SVEP_e = \frac{VRP_e}{IV_e^{base}} \tag{13}$$

where IV_e^{base} is the median ATM IV over $T-60$ to $T-30$. PM execution cost is scaled by the same denominator. The Seesaw verdict: PM wins when $SVEP_e > M_{PM,e}^{scaled}(\$3M)$. The methodology is documented in the Elections Methodology Annex.

Event Sample

Event	**Country**	**ETF**	**PM Volume**	**Exec Cost \$3M (%)**
IRL President 2025	Ireland	EWG*	\$115.1M	0.04
KOR President 2025	S. Korea	EWY	\$29.0M	0.24
CAN PM 2025	Canada	EWC	\$23.2M	0.49
PRT President 2026	Portugal	EWG*	\$18.9M	0.17
CHL President 2025	Chile	ECH	\$13.3M	0.23
ARG Deputies 2025	Argentina	ARGT	\$9.9M	0.33
DEU Chancellor 2025	Germany	EWG	\$9.4M	0.34
POL President 2025	Poland	EPOL	\$7.9M	0.88
DEU Parliament 2025	Germany	EWG	\$7.2M	0.45
CAN GOV 2025	Canada	EWC	\$3.0M	3.33
IRL PM 2024	Ireland	EWG*	\$2.6M	11.53
NOR Parliament 2025	Norway	EWU*	\$2.0M	15.07
PRT PM 2025	Portugal	EWG*	\$2.0M	14.88
AUS Parliament 2025	Australia	EWA	\$1.7M	17.57
TWN President 2024	Taiwan	EWT	\$1.5M	20.66
MEX President 2024	Mexico	EWW	\$0.7M	40.28
GBR PM 2024	United Kingdom	EWU	\$0.3M	50.29

PM volume ranges from \$295K (GBR, July 2024) to \$115M (Ireland President, October 2025). The three earliest events (TWN, MEX, GBR – all 2024) have under \$1.5M. The seven 2025 events above \$7M each represent substantial institutional-scale markets. Countries contributing two event-periods (Canada, Germany, Ireland, Portugal) have materially distinct catalysts or resolution dates constituting separate political processes.

VRP Results

Twelve PM-wins. Five PM-loses.

The raw VRP column reports full equity ETF VRP, not election-specific variance risk premium. Poland's 72.58% is the VRP of EPOL coinciding with the presidential election, against a baseline IV of

Event	Country	VRP (%)	IV_{base} (%)	SVEP	PM \$3M	Verdict
POL President 2025	Poland	72.58	102.52	0.708	0.009	PM wins
NOR Parliament 2025	Norway	22.75	29.84	0.762	0.505	PM wins
ARG Deputies 2025	Argentina	23.99	38.51	0.623	0.009	PM wins
CAN PM 2025	Canada	29.57	51.00	0.580	0.010	PM wins
PRT PM 2025	Portugal	17.33	32.35	0.536	0.460	PM wins
CHL President 2025	Chile	20.52	42.80	0.480	0.006	PM wins
AUS Parliament 2025	Australia	23.27	49.09	0.474	0.358	PM wins
GBR PM 2024	UK	11.36	24.40	0.466	2.061	PM loses
DEU Parliament 2025	Germany	14.67	31.90	0.460	0.014	PM wins
MEX President 2024	Mexico	7.22	17.01	0.425	2.368	PM loses
CAN GOV 2025	Canada	12.18	30.12	0.405	0.111	PM wins
TWN President 2024	Taiwan	13.03	22.30	0.584	0.927	PM loses
IRL President 2025	Ireland	10.02	28.19	0.355	0.001	PM wins
DEU Chancellor 2025	Germany	10.46	36.80	0.284	0.009	PM wins
KOR President 2025	S. Korea	7.66	35.75	0.214	0.007	PM wins
IRL PM 2024	Ireland	2.07	26.68	0.078	0.432	PM loses
PRT President 2026	Portugal	−8.87	25.91	−0.342	0.007	PM loses

102.5%. SVEP contextualizes: the election VRP equals 71% of EPOL's own background volatility. The raw figure is not comparable to BTC's 4.83% or FOMC's 0.52% without standardization.

SVEP across the twelve PM-wins ranges from 0.214 (Korea) to 0.762 (Norway), with a cross-event median of 0.480. The five losses: one structural (PRT President 2026, SVEP −0.342, same pattern as silver), four liquidity-constrained (GBR, MEX, TWN, IRL PM 2024 – all under \$2.6M in PM volume with positive SVEP). §V integrates these into the binding constraint taxonomy.

Gradient Positioning

Elections enter Table IV.22 with a methodological caveat. SVEP is the primary metric; raw VRP is reported for completeness. The gradient position is consistent with the §II.5 prediction: above FOMC and the moderate-VRP categories, likely between BTC and equity indices. Exact placement is uncertain due to proxy contamination – the same contamination that constitutes the Observability Wedge at its most extreme for this category. No election volatility surface exists. No dealer constructed one. The measurement difficulty is the structural prediction of the Cosmos/Taxis ontology: thin infrastructure produces opacity as well as high structural cost.

IV.8. The VRP Gradient

The eleven categories span the VRP spectrum from −10% (silver) to +31% (retail gas, single event with quality caveat), with the central mass between 0.5% (FOMC) and 8.7% (BTC annual). Elections are positioned by SVEP rather than raw VRP due to cross-country proxy heterogeneity (§IV.7).

†FX proxy; measured VRP reflects full exchange-rate volatility. Event-specific rate-decision VRP likely 50–70% lower; ECB plausibly 0.3–0.6%, BOJ plausibly 0.5–1.5%. Gradient position is an upper bound. ‡Quality warnings on all events; July 2025 estimate (31.2%) likely dominated by stale option quotes. Point estimates unreliable. ††Elections VRP reported as SVEP (dimensionless ratio) rather than

Category	N	VRP Obs.	Mean (%)	Median (%)	Range (%)	Tier	Wedge (%)	Reliability
Silver	3	1	−10.19	−10.19	–	1	−10.12	High (control)
FOMC	12	6	0.52	0.50	0.42–0.70	2	0.49–0.77	**High**
CPI	4	4	1.27	1.17	0.63–2.13	3	0.70–2.20	Low
GDP	3	3	1.15	2.15	−2.48–3.78	3	−2.41–3.85	Low
Eq. Indices	13	2	1.90	1.90	1.74–2.06	1	1.81–2.13	**Moderate**
ECB	3	3	2.10†	2.01†	1.56–2.74†	3	1.63–2.81†	Low
BOJ	5	2	3.35†	2.23†	2.23–5.03†	3	2.30–5.10†	Low
BTC	20	6	4.83	4.10	−4.61–15.65	1	−4.54–15.72	**High**
Gold	3	3	13.71	14.33	7.22–19.57	1	7.29–19.64	**Moderate**
Retail Gas	4	4	14.43‡	11.40‡	3.71–31.16‡	3	3.78–31.23‡	Low
Elections	17	17	††	††	††	3+	††	Low-Moderate

raw percentage due to cross-country proxy heterogeneity. Median SVEP of 0.480 across PM-wins is consistent with predicted 3–5% gradient position. Raw VRP ranges from −8.9% to 72.6% but reflects full equity ETF dynamics, not election-specific risk. See §IV.7. The gradient is ordered by median VRP, which is more robust than mean to single outliers (BTC's −4.61% November 2024 horizon, GDP's −2.48% Q1 2025, retail gas's 31.2% July 2025). The ordering is identical under either central tendency measure for all categories except GDP, whose mean (1.15%) and median (2.15%) diverge due to the single negative observation.

The "Ind. VRP" column reports independent VRP observations – the actual degrees of freedom for the structural side of the comparison. The "Reliability" column reflects joint measurement quality from Table IV.2. **The gradient hypothesis should be evaluated primarily against the High and Moderate reliability categories**: Silver, FOMC, Equity Indices, BTC, and Gold – five categories with eighteen independent VRP observations and fifty-one event-contracts. Elections provide a gradient-consistent result at Low-Moderate reliability, with eight headline-clean event-periods out of seventeen tested. The Low reliability categories are consistent with the gradient but should not be weighted as equivalent evidence.

Three observations about the gradient, before threshold analysis.

First, the gradient has structure. Categories do not cluster at random across the VRP spectrum. Rate decisions (FOMC, and plausibly ECB and BOJ once proxy contamination is removed) occupy the low end, consistent with central bank communication reducing genuine uncertainty. Equity indices produce textbook VRP in the 2% range. Crypto and commodities occupy the high end, consistent with persistent volatility and elevated tail risk. The gradient reflects underlying economics, not measurement noise.

Second, VRP varies substantially within categories. BTC ranges from −4.61% to +15.65%. GDP ranges from −2.48% to +3.78%. Retail gas ranges from 3.71% to 31.16% (though the upper end is unreliable). The segmentation hypothesis operates at the category level – high-VRP categories cross first – but §V examines whether event-level VRP predicts event-level threshold status, a finer-grained test than the category-level ordering.

Third, the gradient of VRP corresponds to a gradient of measurement difficulty. Categories with

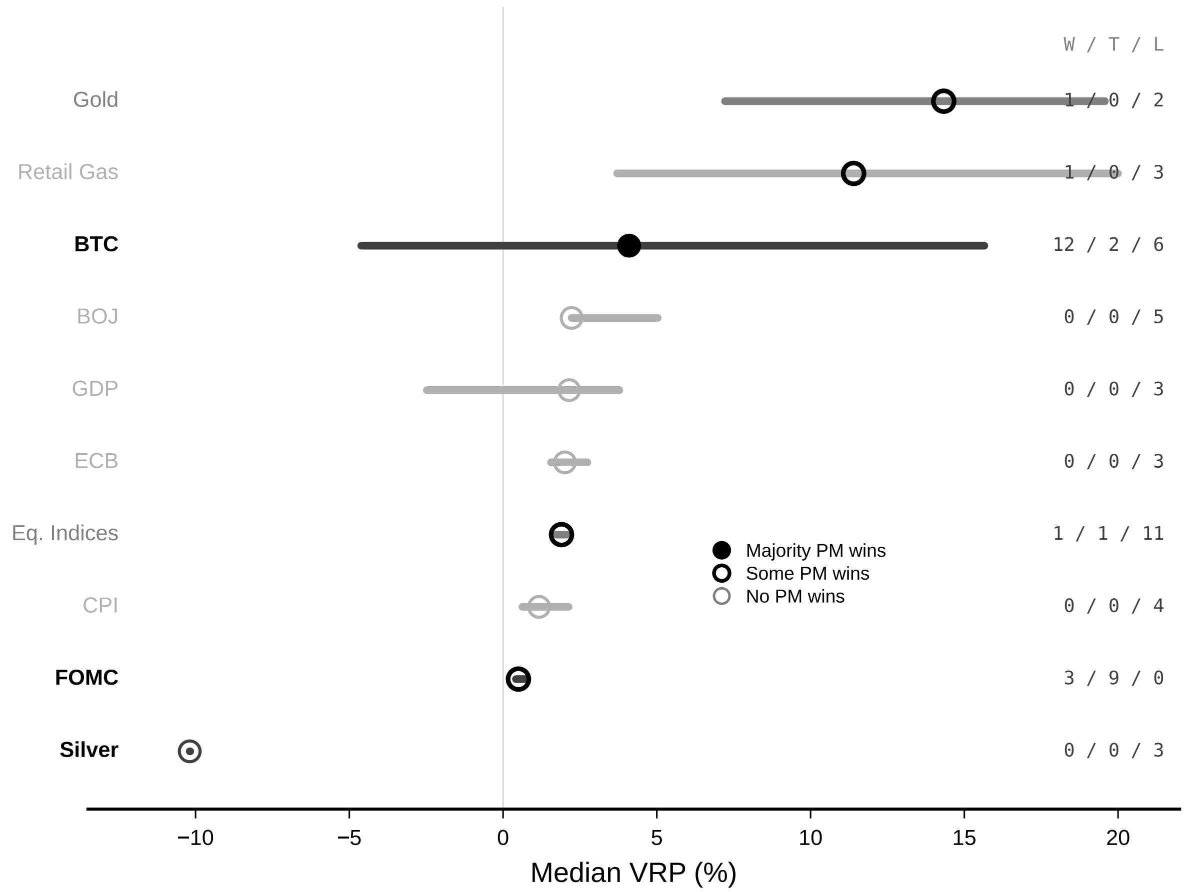

Figure 3. VRP gradient across categories, ordered by median VRP. Range bars show minimum to maximum observed VRP within each category. Threshold verdicts (W/T/L) annotated on the right. High-reliability categories (bold labels) exhibit clean gradient ordering: PM wins cluster at high VRP, losses cluster at low VRP. Low-reliability categories (light labels) are gradient-consistent but carry wider measurement uncertainty.

the lowest VRP (FOMC, CPI) are embedded in the deepest and most transparent derivative infrastructure, where VRP is compressed by competitive vol-selling. Categories with the highest VRP (gold, retail gas, elections) sit in thinner derivative infrastructure where fewer participants compete to sell vol. The same conditions that produce high VRP also produce noisier VRP measurement: thin option chains, stale quotes, wider bid-ask spreads in the derivative market itself. Elections sit at the extreme: no election volatility surface exists, and the proxy chain requires country equity ETFs that absorb 65–90% non-election dynamics.

This correspondence is not coincidental. It is a prediction of the Observability Wedge. Constructed order is most visible where it is most actively maintained – which is where vol-selling is most competitive, which is where VRP is lowest. Constructed order becomes less visible as maintenance withdraws – which is where VRP is highest, which is where measurement requires proxies and carries wider uncertainty bands. The difficulty of measuring VRP in the upper portion of the gradient is evidence for the framework, not a limitation of it. If high-VRP categories were as transparently priced as FOMC, the Observability Wedge would not exist.

Threshold Preview

The data assembled in this section feed the threshold analysis in §V. As a preview of the gradient's predictive power, Table IV.23 summarizes the pipeline's threshold classification at the $3M institutional reference size.

Category	Wins	Thresh.	Loses (Marg.)	Loses (High)	Total	Binding Constraint
FOMC	3	9	0	0	12	Narrow wedge, high volume
BTC	12	2	0	6	20	PM liquidity (thin contracts)
Equity Indices	1	1	1	10	13	PM liquidity (OTM contracts)
Gold	1	0	0	2	3	PM liquidity
Silver	0	0	0	3	3	Structural (negative VRP)
CPI	0	0	0	4	4	PM liquidity
ECB	0	0	0	3	3	PM liquidity
BOJ	0	0	0	5	5	PM liquidity
GDP	0	0	0	3	3	PM liquidity
Retail Gas	1	0	0	3	4	PM liquidity
Elections	12	0	0	5	17	PM liquidity + structural (1)
Total	**30**	**12**	**1**	**44**	**87**	

The "Binding Constraint" column classifies *why* each category arrives at its threshold status – a distinction §V develops in full. Three types emerge. *Structural losses*: Silver and PRT President 2026. The wedge does not exist; PM cannot win regardless of liquidity. *Narrow-wedge competition*: FOMC. VRP is low (0.42–0.70%) but PM volume is enormous ($96M–$365M per contract), compressing execution costs to the point where even a narrow wedge is capturable on the highest-volume contracts. *Liquidity-constrained losses*: The remaining categories where PM loses. VRP may be sufficient – BTC thin contracts, equity OTM contracts, all proxy categories, 2024-vintage election markets – but PM volume is too low to compress execution costs below the structural wedge. These are not verdicts about the wedge; they are verdicts about PM maturity.

The threshold status is also size-dependent. At the $1M reference size, thirty-six events cross the PM-wins threshold (versus thirty at $3M): six additional events – primarily BTC thin contracts and select commodity and election events – become cost-competitive at smaller position sizes. This confirms the size-dependence predicted by §III's Q^* derivation and is analyzed in §V.

The gradient's prediction is partially visible even in this preview: high-VRP categories (BTC, gold, elections) produce PM wins; low-VRP categories (FOMC) produce narrow competition; and negative-VRP categories (silver) produce structural losses. §V develops the full analysis, including within-category gradient tests, the volume multiplier required for currently-losing events to cross threshold, and the twenty-seven events where PM spreads alone exceed the structural wedge – rendering the Seesaw structurally inoperable regardless of volume growth.

The data are assembled. §V tests whether the gradient predicts the sequence of displacement.

V. Findings

Prediction market execution costs fell across every category with sufficient volume during the sample period. The compression was dramatic in the two deepest categories – 78% for FOMC, 57% for BTC – and visible in thinner categories where even modest volume growth reduced spreads and impact. Across all eleven categories, contracts above approximately $5M in PM volume show execution costs well below the structural wedge. Whether this trajectory tips the Seesaw depends on the wedge it must overcome. The VRP gradient determines the sequence.

V.1. The VRP Gradient

The segmentation hypothesis predicts that high-VRP categories cross the displacement threshold before low-VRP categories. Larger wedges absorb more execution cost before the Seesaw tips against PM. The gradient should predict threshold status: PM-wins at the top of the VRP spectrum, PM-losses at the bottom, visible transition in between (Figure 3).

Eleven categories spanning five asset classes test this prediction against approximately fifty-one independent VRP observations. Not all categories carry equal evidentiary weight. Five – Silver, FOMC, Equity Indices, BTC, and Gold – permit high-reliability comparison: their derivative instruments price event risk directly or through validated decomposition (Tiers 1 and 2), and their prediction markets have sufficient depth for meaningful execution cost estimation. These five provide eighteen independent VRP observations across fifty-one event-contracts. They are the primary evidence.

Table V.1: Gradient Evaluation (High-Reliability Categories)

Category	Tier	Ind. VRP Obs.	Median VRP (%)	VRP Range (%)	PM-Wins	Threshold	PM-Loses
Silver	1	1	-10.19	–	0	0	3
FOMC	2	6	0.50	0.42–0.70	3	9	0
Eq. Indices	1	2	1.90	1.74–2.06	1	1	11
BTC	1	6	4.10	-4.61–15.65	12	2	6
Gold	1	3	14.33	7.22–19.57	1	0	2

The ordering holds without exception. Negative VRP produces structural losses. Low VRP produces narrow-wedge competition where PM wins only on the highest-volume contracts. Moderate VRP produces liquidity-dependent outcomes. High VRP produces PM-wins wherever depth is sufficient. The gradient predicts not just which categories PM can contest but the *character* of competition at each level.

Silver anchors the low end. VRP is -10.19%: realized volatility exceeded implied, meaning options were cheap. The structural wedge does not exist. PM cannot undercut a cost that is not there. All three events lose structurally. If the framework only identified winners, selection bias would be a concern. Silver shows it discriminates.

FOMC occupies the narrow-wedge regime. Mean VRP of 0.52

BTC occupies the high-VRP regime. Median VRP of 4.10

Gold confirms the pattern at higher VRP. Three events, one with sufficient PM liquidity ($3.2M), which crosses. Two with thin PM volume, which do not. The binding constraint is depth, not the wedge.

Equity indices sit in the intermediate range. VRP of 1.74–2.06

Five additional categories populate the gradient's intermediate and upper ranges at lower reliability. CPI (median VRP 1.17

Elections extend the gradient into political risk, the eleventh category and fifth asset class. The framework predicted elections would cluster between BTC and the moderate-VRP categories at approximately 3–5

The gradient reflects underlying economics, not measurement noise. Rate decisions occupy the low end – central bank communication reduces genuine uncertainty, deep derivative infrastructure compresses VRP through competitive vol-selling. Equity indices produce textbook VRP in the 2% range. Crypto, commodities, and elections occupy the high end – persistent volatility, elevated tail risk, thinner infrastructure where fewer participants compete to sell vol.

The same conditions that produce high VRP also produce noisier VRP measurement. Low-VRP categories are embedded in the deepest, most transparent derivative markets; high-VRP categories sit in thinner infrastructure with wider spreads and staler quotes. This correspondence is a prediction, not a coincidence. §V.6 develops it as the Observability Wedge.

The gradient predicts the sequence. But "different markets have different costs" is a trivially true statement, and VRP correlating with threshold status could reflect an omitted variable rather than the Seesaw mechanism. The gradient is necessary but not sufficient. The mechanism must be isolated. Two categories provide a test.

V.2. Mechanism Isolation

Cross-category comparison confounds the structural wedge with everything else that differs between markets: participant composition, regulatory environment, platform maturity, information structure. A skeptic can dismiss the gradient as "different markets have different costs" without engaging the Seesaw mechanism.

Two categories provide a cleaner test. Within FOMC and within BTC, multiple contracts share the same VRP while PM liquidity varies. The structural arm is held constant. Only the friction arm moves. If the Seesaw operates, contracts with sufficient PM depth should cross while those with insufficient depth should not – even though both face the same structural wedge.

BTC: January 2026. Five contracts at the same horizon share an identical structural wedge of 4.08% (Figure 4). The $150K touch contract attracted $32.8M in PM volume; execution cost at $3M is 0.13

Table V.2: BTC January 2026 – Mechanism Isolation

Strike	PM Volume	Exec Cost $3M (%)	Structural Wedge (%)	Gap (%)	Verdict
$150K	$32.8M	0.13	4.08	+3.95	PM wins
$100K	$13.3M	0.93	4.08	+3.15	PM wins
$105K	$7.1M	0.86	4.08	+3.22	PM wins
$110K	$4.9M	0.85	4.08	+3.23	PM wins
$125K	$2.2M	13.83	4.08	−9.75	PM loses

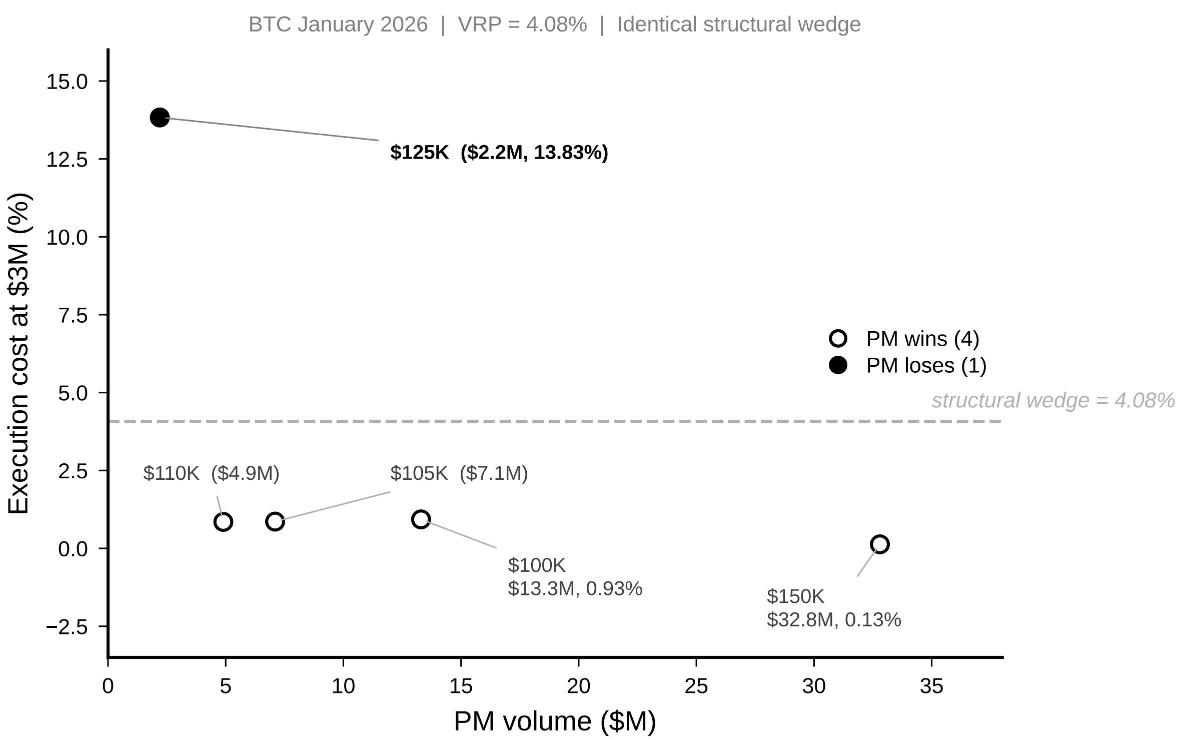

Figure 4. BTC January 2026 mechanism isolation: five contracts sharing an identical 4.08% structural wedge (dashed line) with PM volume ranging from $2.2M to $32.8M. Execution cost at the $3M reference size varies from 0.13% to 13.83%. Four contracts with sufficient PM depth cross; the single thin contract loses. The structural arm is held constant; only the friction arm moves.

Four contracts with PM volume above $4.9M cross with decisive gaps. The single contract below $3M loses by nearly ten percentage points. Same wedge. Same VRP. The only variable that moved is PM liquidity, and it moved the verdict.

November 2025 replicates the experiment at a slightly higher wedge (4.29

This is not “different markets have different costs.” Same market, same cost structure, same moment. What varies is whether enough participants showed up to compress friction below the wedge. The Seesaw operates at contract level: structural arm fixed, friction arm variable, verdict determined by their relative magnitude.

FOMC: January 2026. Four contracts at the same meeting share an identical VRP_{adj} of 0.44

The January 2026 FOMC experiment operates in a different regime from BTC. The structural wedge is narrow – 0.51

The contrast with earlier FOMC meetings sharpens the finding. The January 2025 meeting generated $96M in PM volume on a single contract, with execution costs of approximately 0.08

The joint finding. BTC isolates the mechanism in the high-VRP regime: large wedge, wide variation in PM depth, clear separation between contracts that cross and contracts that don’t. FOMC isolates it in the low-VRP regime: thin wedge, enormous PM volume, confirmation that even narrow wedges are capturable at sufficient depth. Together, they demonstrate that the Seesaw operates at both ends of the gradient, that VRP determines the height of the bar, and that PM liquidity determines whether any given contract clears it.

The mechanism is isolated. §V.3 applies it across the full sample.

V.3. Threshold Results

The Seesaw condition – structural wedge exceeds PM execution cost – is applied to all eighty-seven event-contracts at the $3M institutional reference size (Figure 5). The mechanism has been isolated; the gradient has been established. What follows are the verdicts.

Table V.3: Threshold Status by Category (at $3M)

Category	**PM Wins**	**Threshold**	**PM Loses (Marginal)**	**PM Loses (High Friction)**	**Total**
BTC	12	2	0	6	20
FOMC	3	9	0	0	12
Silver	0	0	0	3	3
Equity Indices	1	1	1	10	13
Gold	1	0	0	2	3
CPI	0	0	0	4	4
ECB	0	0	0	3	3
BOJ	0	0	0	5	5
GDP	0	0	0	3	3
Retail Gas	1	0	0	3	4
Elections	12	0	0	5	17
Total	**30**	**12**	**1**	**44**	**87**

Classification criteria: "PM wins" when structural wedge exceeds execution cost by more than 0.5 percentage points. "Threshold" when the gap is within ±0.5 percentage points. "PM loses (marginal)" when execution cost exceeds structural wedge by 0.5 to 2.0 percentage points. "PM loses (high friction)" when execution cost exceeds structural wedge by more than 2.0 percentage points. The ±0.5pp boundary reflects measurement uncertainty in both the VRP estimate and the execution cost estimate; verdicts within this range should be held loosely. The qualitative findings are boundary-invariant: BTC wins at gaps of +3 to +15 percentage points survive any reasonable perturbation of the classification threshold, FOMC threshold events remain at threshold under tighter or wider bands, and the constraint taxonomy is unchanged.

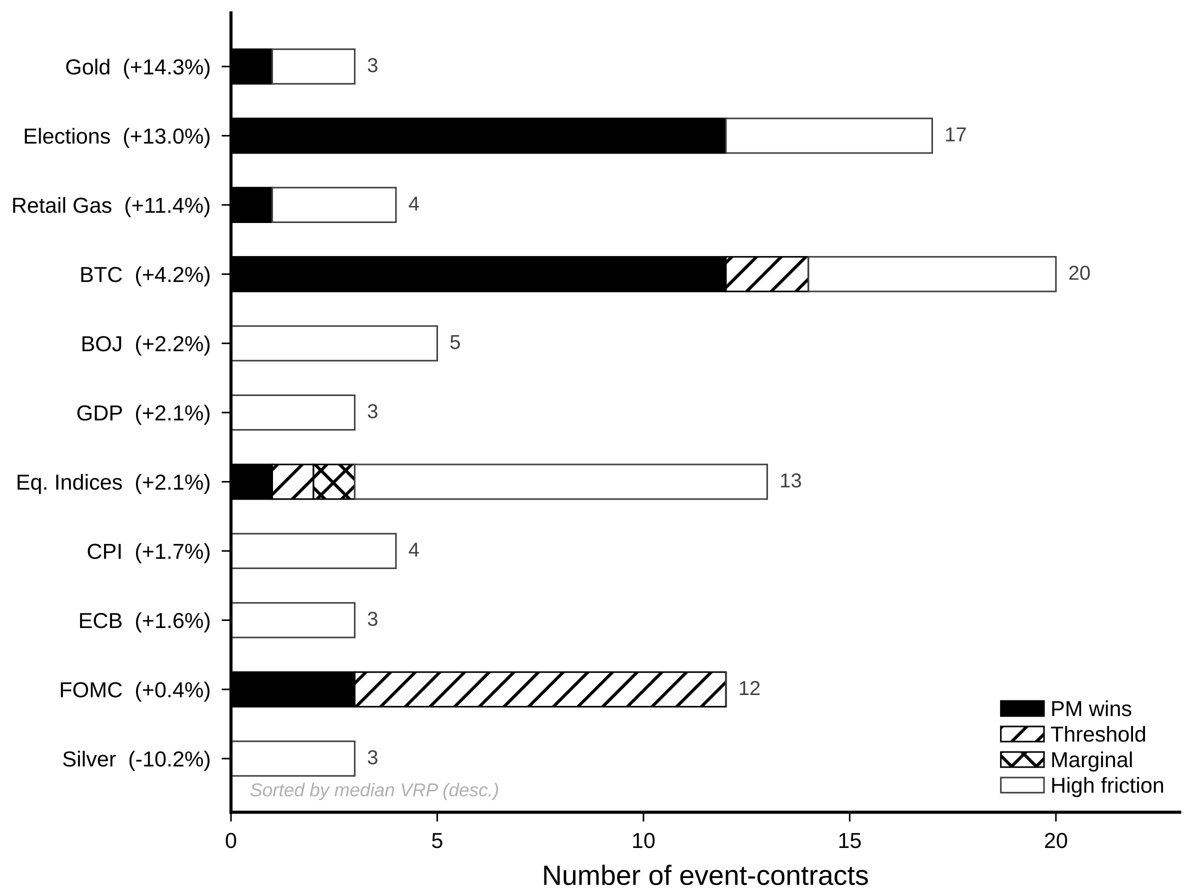

Figure 5. Threshold verdicts by category at the $3M reference size, sorted by median VRP (descending). Stacked bars show PM wins (solid), threshold (hatched), marginal losses (cross-hatched), and high-friction losses (white). Total event counts annotated on the right. The transition from all-win (BTC) to all-loss (Silver) follows the VRP gradient.

Thirty PM-wins at $3M. The verdicts cluster where the gradient predicts: BTC (12), elections (12), and FOMC (3) account for twenty-seven of thirty. The remaining three – one each from equity indices, gold, and retail gas – are high-VRP events where even moderate PM depth sufficed.

The verdicts are size-dependent. At the $1M reference size, thirty-six events cross – six more than at

$3M. Impact costs scale with position size; the structural wedge does not. The six additional events that cross at smaller scale are primarily BTC thin contracts and select commodity and election events where the wedge comfortably exceeds spread cost but impact at $3M pushes total execution above the threshold. This confirms the Q* size-dependence derived in §III: the position size at which the Seesaw tips is a function of the structural wedge and the depth of the PM market. For some events, the Seesaw has tipped at institutional-but-not-large-institutional scale. Full contract-level detail for the size-dependence analysis is reported in Annex Tables A.8 and A.9.

Twenty-seven of eighty-seven events face a more fundamental constraint. At these events, the PM bid-ask spread alone – before any depth or impact consideration – exceeds the entire structural wedge. The Seesaw is structurally inoperable regardless of volume growth. Volume compresses impact but does not eliminate the spread. Market-maker competition compresses spreads, but this is a slower dynamic than depth improvement, and these twenty-seven events sit far enough from threshold that both mechanisms would need to operate substantially before crossing becomes possible. §V.4 develops this distinction.

We present category-level results for the five high-reliability categories, beginning with the strongest evidence.

V.3.1. BTC

Twelve of twenty BTC contracts cross at $3M. Two sit at threshold. Six lose.

The twelve wins are concentrated in contracts with PM volume above $5M – the depth threshold identified in §V.2. The gaps are decisive: +3.0 to +15.5 percentage points in high-VRP horizons. Even in the lowest positive-VRP horizon (January 2026, VRP 4.00

The two threshold contracts are intermediate cases: PM volume between $3M and $5M, execution costs compressing toward but not clearly below the wedge. Measurement uncertainty prevents a clean verdict.

The six losses partition into five thin contracts with PM volume under $3M and one structural loss (November 2024, VRP -4.61

The two BTC marginal periods deserve attention. November 2024 produced the single negative-VRP horizon (-4.61

V.3.2. FOMC

Three of twelve FOMC contracts cross. Nine sit at threshold. None lose outright.

This is the finding that has changed most from V1, which reported zero of six PM-wins. The shift reflects two developments: the expansion from six single-contract meetings to twelve contracts across six meetings (including multi-contract events), and the extraordinary growth in FOMC PM volume. The January 2026 meeting generated over $1 billion in PM flow across four contracts – on an offshore platform, with no institutional custody infrastructure. The plumbing has not arrived. The demand has.

The three wins are high-volume contracts at recent meetings. The nine threshold contracts are not failures – they are narrow misses in a category where the structural wedge averages 0.59

The volume trajectory within FOMC is itself informative (Figure 6). January 2025 generated $96M on a single contract. September 2025 generated $204M across two. January 2026 generated $1,007M across four. This is not steady-state liquidity. It is a market discovering that the instrument is useful. Whether this volume growth is sustainable depends on the structural economics of PM liquidity provision – addressed in §V.5.

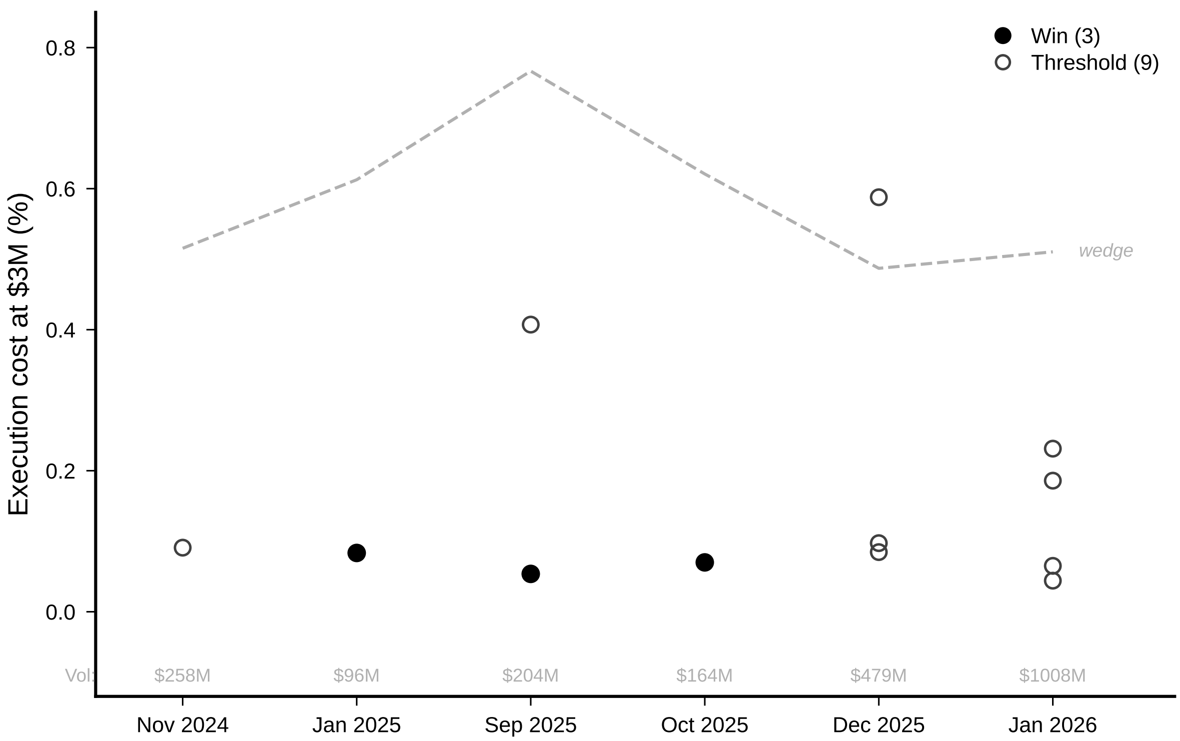

Figure 6. FOMC contract execution cost versus structural wedge over time (November 2024 through January 2026). Volume annotations show cumulative PM flow per meeting ($258M to $1,008M). Filled circles indicate PM wins; open circles indicate threshold events. The dashed line marks the structural wedge. The trajectory shows execution costs falling below the wedge on the most liquid contracts as PM volume grew fourfold over 14 months.

A note on what FOMC demonstrates. The segmentation hypothesis predicted that low-VRP categories would be the last to cross. FOMC is crossing – on its most liquid contracts, at the narrowest margins. This is not a revision of the prediction; it is a refinement. The prediction was that high-VRP crosses *first*, not that low-VRP *never* crosses. FOMC's narrow-wedge competition shows the Seesaw tipping at the margin in the category where tipping is hardest. If PM volume continues to grow, the nine threshold events become wins. The binding constraint is not the wedge – the wedge exists, however thin. The binding constraint is execution cost, and execution cost is falling.

V.3.3. Silver

Zero of three events cross. The framework correctly identifies the category where PM cannot win.

Silver VRP is -10.19

All three silver events share the same VRP – one independent observation – measured against SI futures options (Tier 1, direct). PM volumes are moderate ($2.6–2.9M), sufficient for execution cost estimation but irrelevant to the verdict. The loss is structural, not frictional.

A caveat on permanence: the -10.19

V.3.4. Equity Indices and Gold

Equity indices: one PM-win, one threshold, one marginal loss, ten high-friction losses across thirteen contracts. VRP of 1.74–2.06

Gold: one PM-win, two losses across three events. The win is the Gold ≥$3,200 EOY 2024 event – the cleanest gold measurement ($3.2M PM volume, 7.22

V.3.5. Elections

Twelve PM-wins, five losses across seventeen event-periods in twelve countries. The wins are concentrated in 2025 events where PM depth had matured: Ireland President ($115M volume, 0.04

V.3.6. Tier 3 Categories

CPI, ECB, BOJ, GDP, and Retail Gas collectively produce one PM-win and eighteen losses across nineteen events. The single win is a retail gas event where VRP (even with quality caveats) is high enough to absorb thin-market execution costs. The eighteen losses are all on PM liquidity: ECB contracts at $0.8–1.1M, GDP contracts under $400K, CPI outside the headline release under $400K. Proxy contamination overstates their VRP, making these PM-loses verdicts conservative – if PM cannot win against an inflated wedge, it would lose against the true event-specific wedge even more decisively. Their gradient positions are reported in Table IV.22 for completeness but should not be weighted as equivalent to the direct-measurement categories.

The aggregate picture: thirty PM-wins concentrated in high-VRP categories with deep PM liquidity. Twelve threshold events concentrated in FOMC, where the wedge is narrow and PM volume is growing. Forty-five losses driven overwhelmingly by insufficient PM depth, with silver and one election event as the structural exceptions. The data support cost-competitiveness on specific contracts. They do not yet support observed displacement.

A note on language, carried forward from the V1 analysis: "PM wins on cost" means prediction markets offer lower total cost for accessing the same risk exposure at the reference position size. It does not mean institutional flow has migrated, dealers have withdrawn, or displacement has occurred.

Those are predictions that follow from the cost finding. The cost comparison is what we measure. Displacement is what we predict.

The forty-five losses are not uniform. Three distinct constraint types explain why the cost comparison does not favor PM at these events, and predict different futures for each.

V.4. Binding Constraint Taxonomy

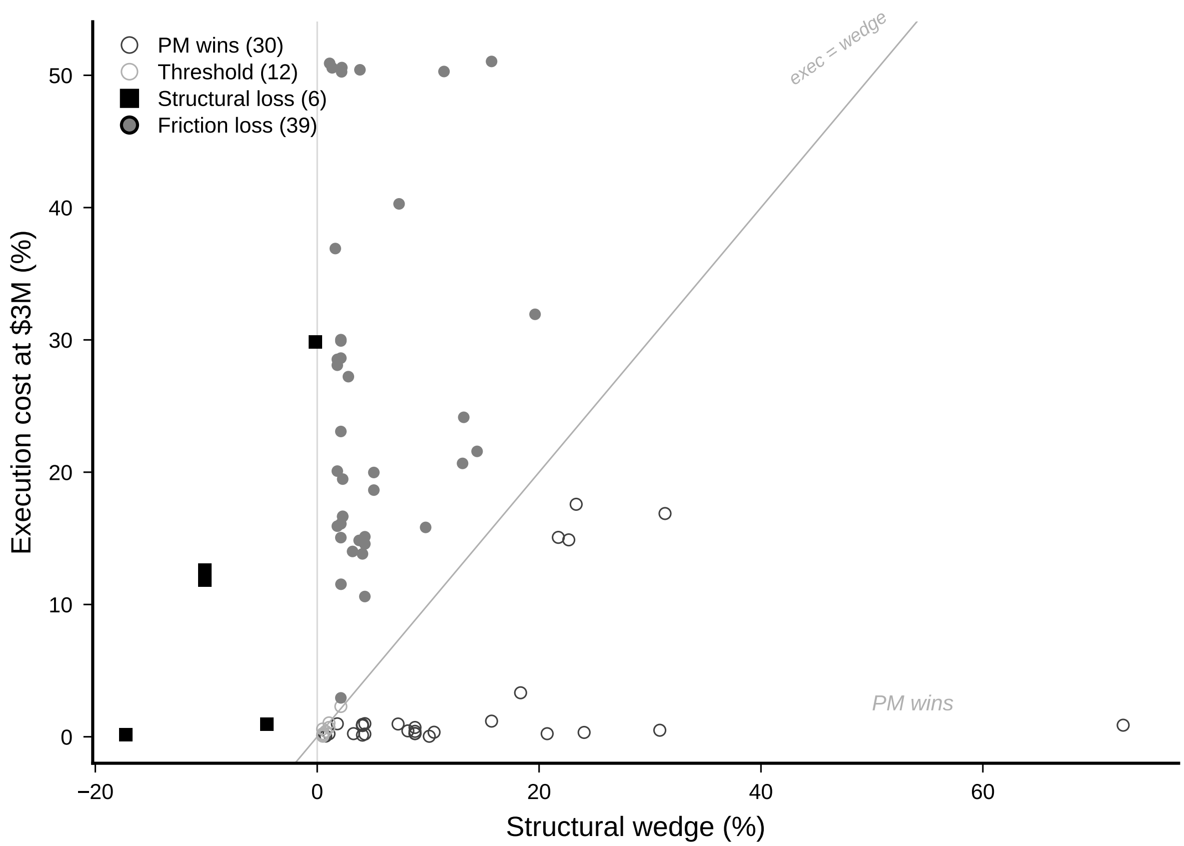

Figure 7. Structural wedge versus PM execution cost for all 87 event-contracts at the $3M reference size. The 45-degree line marks where execution cost equals the structural wedge. Events below the line favor prediction markets. Four marker types distinguish outcomes: PM wins (filled dark circles, 30 events), threshold (open dark circles, 12), structural losses (filled squares, 6), and liquidity-constrained losses (open light circles, 39). The FOMC cluster near the origin is expanded in Figure 8.

The forty-five events that do not cross at $3M are not uniform failures (Figure 7). Three distinct constraint types explain why the cost comparison does not favor PM, and each predicts a different future.

Structural losses. The structural wedge is zero or negative. VRP is absent: realized volatility meets or exceeds implied, meaning options are cheap. PM cannot undercut a cost that does not exist. No amount of PM volume growth, spread compression, or market-maker competition changes the verdict. The constraint is the underlying economics of the asset class, not market development.

Silver is the clear case: VRP of -10.19

Spread-constrained losses. The structural wedge is positive but the PM bid-ask spread alone meets or exceeds the wedge minus the 0.5 percentage point buffer. Volume growth compresses impact but does not eliminate the spread floor. These events cannot cross through organic flow alone – they require spread compression, which is driven by market-maker competition rather than trade volume per se.

Twenty-two events fall here. The composition is revealing:

Nine are FOMC Threshold events. FOMC has enormous PM volume – \$72M to \$365M per contract – but spreads of 0.04

Seven are equity index OTM contracts. PM volumes of \$1–3.6M with spreads of 0.07

Four are CPI and GDP events. PM volumes under \$400K on three of four, with spreads that dwarf the structural wedge. These are nascent markets where the market-making business model has not yet arrived.

Two are remaining cases across categories.

The FOMC spread-constrained events deserve particular attention. Spreads compressed substantially over the sample period: early FOMC contracts (Nov 2024) show effective spreads of 0.08

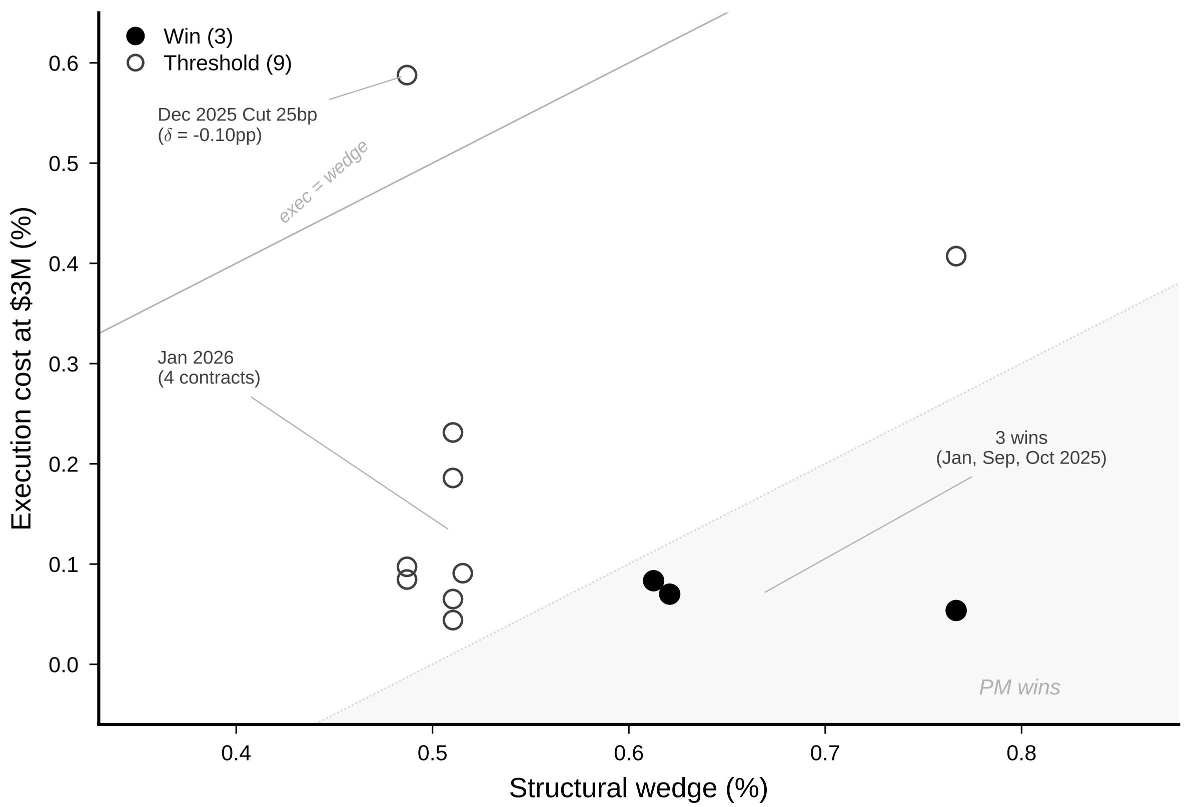

Figure 8. Zoomed view of the 12 FOMC contracts from Figure 7, showing the 0.35%–0.85% wedge range where narrow-wedge competition occurs. Filled circles indicate PM wins (3); open circles indicate threshold events (9). The PM-win boundary (±0.5pp) and the 45-degree exec = wedge line are shown. Several threshold events fall within basis points of the PM-win classification boundary.

Liquidity-constrained losses. The structural wedge exceeds the spread floor – the market-making business model works – but PM depth is insufficient to absorb a $3M position without excessive impact. Volume growth alone tips the verdict. The required volume multiplier is calculable from the depth curve under the assumption that depth scales linearly with market volume.

Seventeen events fall here. Four are 2024-vintage election markets – GBR ($295K), MEX ($748K), TWN ($1.5M), IRL PM 2024 ($2.6M) – where SVEP is positive and substantial on three of four but PM depth had not yet developed. The remaining thirteen span the original ten categories. The multiplier ranges from 1.02× (nearly sufficient) to 8.89× (far from sufficient):

Table V.4: Volume-Constrained Events – Required Growth

Event	Category	Current Vol	Required Vol	Multiplier
BTC_120K_NOV_2025	BTC	$2.94M	$3.00M	1.02×
BTC_150K_OCT_2025	BTC	$4.99M	$6.45M	1.29×
BTC_125K_JAN_2026	BTC	$2.20M	$3.00M	1.36×
BTC_150K_NOV_2025	BTC	$2.07M	$3.00M	1.45×
BTC_130K_NOV_2025	BTC	$2.02M	$3.00M	1.49×
GAS_JUN_2025	RETAIL_GAS	$2.17M	$3.29M	1.52×
GAS_MAY_2025	RETAIL_GAS	$2.01M	$3.16M	1.57×
BOJ_DEC_INC_25BPS	BOJ	$1.86M	$3.13M	1.69×
GAS_APR_2025	RETAIL_GAS	$1.34M	$2.69M	2.01×
ECB_OCT_2025	ECB	$1.10M	$3.00M	2.72×
SPX_GTE_5700_DEC	EQ. INDICES	$3.40M	$13.43M	3.95×
BTC_130K_OCT_2025	BTC	$5.69M	$34.45M	6.06×
BTC_100K_DEC_2024	BTC	$0.24M	$2.13M	8.89×

Seven of the thirteen are BTC contracts with multipliers under 2×. These are the nearest edge of the displacement frontier – one doubling of thin-contract volume, in a category where liquid contracts already reach $33M, would tip most of them. BTC_120K_NOV_2025 at 1.02× is effectively already at threshold; the linear depth assumption at this proximity is well within its defensible range.

The further cases (SPX_GTE_5700 at 3.95×, BTC_130K_OCT at 6.06×) require extrapolation beyond observed depth, where the linearity assumption is less reliable. Depth may scale better than linearly at low volumes (market-maker entry) or worse than linearly at high volumes (capital constraints). The multipliers should be read as indicative of distance from threshold, not as precise engineering targets.

The displacement frontier. The three constraint types partition the loss space into regions with different dynamics:

Structural losses sit outside the frontier entirely. They cannot cross without a regime shift in the underlying asset's volatility structure – a macro question, not a market development question.

Spread-constrained losses sit at the frontier but require a specific development: market-maker entry and competitive spread compression. This is the Cantillon-Yin dynamic – venue competition attracting professional liquidity providers whose presence compresses costs, which attracts more flow. The FOMC trajectory is suggestive: $96M in PM volume at a single January 2025 meeting, over $1 billion across four contracts at January 2026. Volume grew tenfold. Spreads compressed. Three contracts tipped from threshold to PM-wins. Nine remain at threshold, most within fractions of a percentage point of crossing. The claim is not that FOMC will cross entirely – it is that the trajectory over twelve months moved the category from zero PM-wins to three, with nine more at the boundary.

Liquidity-constrained losses sit inside the frontier. The spread is already below the wedge; depth is the only remaining barrier. These tip with organic volume growth – and BTC liquid contracts demonstrate that the category can sustain \$5–33M in per-contract volume. The thin BTC contracts that lose are not in a different market from the liquid ones that win. They are the same market at lower salience.

The taxonomy predicts that displacement, if it proceeds, proceeds unevenly: first where the wedge is wide and PM already deep (BTC liquid contracts – already crossed), then where the wedge is wide but PM thin (BTC thin contracts – nearest to crossing), then where the wedge is narrow and PM volume is growing (FOMC – actively tipping), and last where the wedge is narrow and PM nascent (ECB, BOJ, CPI, GDP – far from crossing). This sequence is not projected from growth rates. It is read directly from the structural position of each constraint type relative to the frontier.

V.5. Why the Mechanism Sustains

The cost comparison raises a question: who takes the other side?

If prediction markets are cheaper for hedgers, someone provides liquidity. Fragile liquidity (retail enthusiasm, election-year speculation) would make the threshold crossing temporary. Structural liquidity (professional participants earning sustainable returns) makes it persist.

Two independent empirical studies address this. Bürgi et al. (2025) analyze 313,972 Kalshi contracts from November 2021 through April 2025. Reichenbach and Walther (2025) analyze 124 million Polymarket trades from nearly one million traders representing \$48 billion in volume. The platforms, samples, and methodologies differ. The findings converge.

Maker/Taker Economics

Bürgi et al. document that Makers (limit order posters) earn average returns of -12

Reichenbach and Walther find that only 30

The 70

Cosmos, Not Taxis

Professional participation does not convert Cosmos to Taxis. The distinction concerns how price is produced.

A market maker posting limit orders to Polymarket remains a participant in emergent order. Their quotes enter the public book alongside retail orders. Price emerges as the exhaust of the matching process. They compete for flow on a transparent order book. The distinction rests on the mechanism of price production, not the sophistication of participants. A market maker running quantitative models posts to the same public order book as a retail participant. The price emerges from order flow interaction, not from a maintained model that the market maker controls access to. Sophistication of participants is compatible with emergence of the price.

The execution cost compression documented across the sample period (78

Adverse Selection Reframed

Standard adverse selection concern: informed flow arrives, uninformed liquidity providers get picked off, LPs withdraw, spreads widen, market fails.

Prediction market adverse selection differs. In OTC derivatives, informational asymmetry derives from differential access to order flow. Dealers see what clients cannot. In prediction markets, informational asymmetry derives from domain knowledge. A participant who spent fifteen years in Iowa politics may hold genuine edge in Iowa election markets. A participant with no cryptocurrency expertise will lose in BTC threshold markets.

Edge comes from knowledge. Participants self-select into markets where they have information advantage. Surplus flows to people who know things.

Transition Risk

Institutional adoption may temporarily widen spreads. New flow with correlated information (hedgers arriving simultaneously to cover the same exposure) creates adverse selection pressure on existing liquidity providers. This is the J-curve: spreads widen during transition before new liquidity arrives to service the flow.

The J-curve is transition risk. If the structural cost advantage persists, capital follows. Cantillon and Yin (2007) documented this pattern in the Bund futures migration: DTB attracted newcomers, the market expanded, liquidity concentrated, and the incumbent collapsed. The coordination problem solved itself through market expansion.

Synthetic Exposure

BTC threshold markets form a discrete approximation of a volatility surface. Contracts at $75K, $80K, $90K, $100K, $110K, $125K each pay $1 if touched by month-end. Taken together, they describe the market's expected distribution across price levels.

An institution seeking exposure to "BTC between $100K and $110K by March" can construct the payoff from two contracts:

- Long $100K touch at $0.65
- Short $110K touch at $0.40

Net cost: $0.25. Payoff: $1 if BTC reaches $100K but not $110K, $0 otherwise. This is a vertical spread, the same structure that options desks build from calls or puts.

Other structures follow the same logic:

The discrete strikes create approximation error. A trader wanting exposure at $107K cannot get it directly. But for institutional hedging, where precision matters less than cost and execution, the discrete grid is sufficient. Corporate treasurers hedge FX in round-number notionals. The same logic applies.

Structure	Construction
Bull spread	Long lower strike, short higher strike
Range binary	Long touch at floor, short touch at ceiling
Straddle equivalent	Long upside touch + long downside touch at equidistant strikes

V.6. The Observability Wedge

The previous sections measured what could be measured. This section concerns what could not.

V.6.1. The Absent Market

The native instrument for FOMC binary hedging is Fed Funds options (OZQ). We did not use it. We could not use it.

Table V.5: Evidence of Fed Funds Options Market Absence

Source	Finding
Databento (CME GLBX MDP3)	Zero rows returned for OZQ order book data
CME Interest Rate Liquidity Review 2024	Fed Funds options not mentioned in volume analysis
CFTC Bank Participation Reports	Excluded (<5 banks hold reportable positions)
Clarus Analytics	Volume described as "tiny"; excluded from analysis
CME trading floor	Fed Funds options pit closed May 4, 2021

The contract exists. CME lists it. Specifications are published. Clearing is available. But there is no continuous order book, no publicly quoted spread, no observable depth, no resting bids and offers.

What liquidity exists occurs off-book: block trades, exchange for related positions (EFRPs), bilaterally negotiated and subsequently reported through CME for clearing. The price is produced inside dealer relationships, not aggregated from public order flow. A client requesting a Fed Funds options quote receives a price derived from the dealer's model, adjusted for inventory position and balance sheet cost, with no public reference price against which to benchmark.

The instrument exists. The market does not.

This is a precise claim. A market, in the microstructure sense, instantiates specific objects: quotes, depth, continuous price formation, queue dynamics, cancellation flow. These objects permit external observation and analysis. Fed Funds options do not instantiate them. The contract is a market in the legal sense (listed, regulated, clearable) but not in the operational sense. It is a contract that is not a market.

V.6.2. The Measurement Asymmetry

To measure options-side costs for FOMC binary risk, we built a seven-step inferential apparatus:

1. Accept that the native instrument has no observable market

2. Select a proxy (SR3) with different tenor and structure
3. Address multi-meeting contamination through event-window isolation
4. Apply Black-76 to extract implied volatility from settlement prices
5. Convert from normal (basis points) to percentage terms
6. Calculate realized volatility from Fed Funds futures in matched windows
7. Compute VRP as the difference, acknowledging decomposition uncertainty

To measure prediction market costs:

1. Look at the order book

The contract "Fed holds at January FOMC" displays a price. 74¢ means 74

The methodological burden to measure the cost of constructed order mirrors the transactional burden to access constructed order. Both require apparatus. Both require insider knowledge or painstaking reconstruction. Both reflect costs that spontaneous order, by construction, avoids.

V.6.3. The Inference

One might object that we are rationalizing a data limitation as a finding. The objection would hold if Fed Funds options were simply illiquid, if low trading interest explained the absence of observables. But the evidence suggests something different: pricing does occur. Dealers quote to clients. Trades clear through CME. Open interest exists, however small. The apparatus is maintained; it simply does not emit public traces.

This is not contingent illiquidity but architectural opacity.

The distinction matters. Contingent illiquidity implies the observables could return if interest picked up. Architectural opacity implies the structure itself prevents public instantiation. Fed Funds options exhibit both architectural and economic opacity: bilateral market structure routes flow away from public venues, and limited speculative appeal reduces the incentive to maintain one. Whether the apparatus cannot or does not emit public prices, the result for the external observer is identical. The cost comparison cannot be computed directly. The market exists for relationship participants and does not exist for external observers: not because no one is trading, but because the apparatus is designed to function inside relationships rather than emit public prices. The apparatus reveals its maintenance costs to those inside (clients who receive quotes, dealers who calibrate surfaces) while emitting nothing to those outside.

The distinction matters because the framework predicts exactly this pattern. Constructed order, when maintenance withdraws from public markets, continues to function inside relationships while ceasing to instantiate the objects that make something observable as a market. This is not a novel claim. Prior work on opacity in wholesale finance showed that bilateral markets dependent on dealer apparatus can maintain pricing inside relationships while ceasing to exist as observable market (Duffie, 2012). The Global Financial Crisis demonstrated the systemic consequences: risks accumulated in structures that didn't emit public prices until catastrophic failure made them visible.

Fed Funds options are a small instantiation of this pattern. The absence of observables is not a data limitation. It is the predicted failure mode of Taxis at its boundary.

V.6.4. State Observability

The asymmetry is not about information transparency. Neither Cosmos nor Taxis makes the information behind prices accessible: the beliefs, the private knowledge, the reasoning that generates order flow. What differs is state observability.

Cosmos makes the current state of the market visible: the price, the depth, the spread, the queue. You can see where the market is, even if you cannot see why it is there. The order book is a public object, updated continuously, observable by anyone. This is not a design choice that could easily have been made differently. Public aggregation of dispersed orders is constitutive of what makes an order book an order book. The observability is the mechanism.

Taxis, when maintenance withdraws from public instantiation, provides neither. The price exists, but only inside the apparatus. From outside, you do not know where the market is. You do not know if a market exists in the operational sense. You can infer, approximate, construct proxies. But the state is not observable. It must be reconstructed.

This is the Observability Wedge. Not a second cost to be added to the Vega Wedge, but a precondition, a pre-wedge that exists even when the Vega Wedge cannot be computed. Where observation requires insider access, the Vega Wedge is not large; it is undefined. You cannot compute a cost differential for an instrument that does not emit public prices.

The Vega Wedge for Fed Funds options is not literally infinite; we approximated it via proxy instruments at 0.69

V.6.5. Failure Modes

Spontaneous and constructed order fail differently.

Cosmos can be thin, noisy, even manipulable, but it remains observable. You can see it being wrong. You can see spreads widen. You can see depth vanish. You can watch prices move on small trades and know the market is unreliable. The failure is visible, legible, available for analysis and response.

Taxis can function (pricing occurs, trades clear, the apparatus is maintained) and yet be externally uninspectable because the functioning happens inside relationships. The failure mode is not degradation but disappearance. When the apparatus withdraws from public markets, the market does not become thin. It becomes private. The price continues to exist for those inside. It ceases to exist for those outside.

One fails loudly, on-screen. The other fails quietly, by not printing anything you can analyze. The

Demiurge of the volatility surface continues to speak, but only to those inside the apparatus. From outside, silence.

This is not a claim about which failure mode is worse. It is a claim about what kind of public object a price is. Cosmos produces prices as a side effect of aggregation; the order book is the exhaust of the matching process, publicly visible by construction. Taxis produces prices inside an apparatus; publicity is optional, contingent on the maintainer's choice to emit. The mode of existence differs, not merely the mechanism of production.

Zhu (2014) examined the relationship between opacity and price formation, finding that dark venues can harm price discovery when they capture uninformed order flow. The dynamic is relevant here: Taxis at its boundary captures informed flow (relationship clients) while emitting nothing to inform the broader market. The price discovery function is privatized.

The GFC was, among other things, a demonstration of what happens when Taxis fails quietly at scale. The risks accumulated in bilateral relationships (credit derivatives, repo chains, securitization structures) that did not emit public prices. The apparatus functioned for those inside. It was invisible to those outside. The failure was invisible until it was catastrophic.

V.6.6. The Generalization

Fed Funds options are a single case. We have documented that this particular instrument, at this particular time, exhibits the predicted failure mode of constructed order at its boundary. We do not claim this case proves a universal pattern.

The claim is narrower: for instruments dependent on dealer apparatus, architectural opacity is a predictable failure mode when maintenance thins. Fed Funds options demonstrate the pattern with unusual clarity because they occupy a boundary: listed, contractually valid, functionally non-existent as observable market. The apparatus is maintained (dealers quote to clients, trades clear) but the public market has vanished.

This boundary case is diagnostic. It shows what Taxis looks like when the maintainer is present but the public instantiation is absent. It reveals the distinction between legal existence and market existence, between the contract and the market, between the price that exists inside relationships and the price that exists as public object.

Future work should examine whether the pattern generalizes to other thinly-maintained Taxis instruments: corporate bonds at the illiquid end, exotic structures that never traded publicly, bespoke OTC positions that exist only in bilateral relationships. The topology of observability (what can be seen from outside, what requires insider access) is prior to the cost calculations the Vega Wedge performs. We have mapped one point on this topology. The full terrain remains to be surveyed.

The seven steps required to approximate what cannot be directly observed are not methodology. They are the phenomenon.

The provenance hierarchy developed in §IV generalizes the pattern beyond Fed Funds options. Tier 1 categories (BTC, equity indices, gold, silver) emit public observables from their derivative markets:

continuous order books, quoted spreads, transparent settlement. VRP measurement requires one step – observe the market. Tier 2 (FOMC) requires the seven-step apparatus documented above. Tier 3 categories (ECB, BOJ, CPI, GDP, retail gas) require proxy measurement through instruments one hop removed from the event itself – EUR/USD options for ECB decisions, Treasury options for CPI releases – introducing contamination from non-event dynamics. The progression from Tier 1 through Tier 3 is a gradient of observability: each tier represents an additional layer of apparatus required to reconstruct what the derivative market does not directly emit. The provenance hierarchy is a legibility hierarchy in the sense Scott (1998) intended. Tier 1 categories make themselves legible: the order book presents its prices to anyone who looks. Tier 3 categories require the observer to reconstruct legibility through proxy, introducing the same distortions Scott documented when central authorities imposed simplification schemes on complex local realities. The Observability Wedge is the point where legibility fails entirely. A simpler reading holds that the observability gradient reflects derivative market depth rather than anything ontological. Deep markets produce transparent data; thin markets produce noisy data. The correlation is real. But depth alone does not explain Fed Funds options, where institutional demand for FOMC hedging is substantial yet the market is architecturally opaque. The apparatus is maintained. It does not emit. Depth explains gradations of measurement noise across the spectrum. It does not explain the categorical absence of public price formation in a product with active institutional use. The Observability Wedge is not confined to a single absent market. It is the structural feature of how Taxis relates to its own prices across the full spectrum of institutional entrenchment.

V.7. The Smile Analysis

The volatility smile – implied volatility varying across strike prices – is a signature feature of derivatives markets, reflecting the market's assessment of tail risk. It is typically produced by dealer apparatus: models interpolating across strikes, inventory shaping quotes, dynamic hedging transmitting information.

BTC touch markets test whether the smile requires apparatus. It does not (Figure 9).

V.7.1. Methodology

Polymarket's BTC touch markets offer contracts at discrete thresholds: \$75K, \$80K, \$90K, \$100K, \$110K, \$125K, \$150K, etc. Each contract pays \$1 if BTC reaches the threshold by month-end, \$0 otherwise. The price is the market's implied probability of reaching that level.

We extract implied volatility from these touch probabilities using a GBM-based inversion. Under geometric Brownian motion with constant volatility σ, the probability of an up-touch (spot S reaching barrier B > S before expiry T) is given by:

$$P_{\text{touch}} = N\left(\frac{\ln(B/S) - \mu T}{\sigma\sqrt{T}}\right) + \left(\frac{B}{S}\right)^{2\mu/\sigma^2} N\left(\frac{-\ln(B/S) - \mu T}{\sigma\sqrt{T}}\right) \tag{14}$$

Where N(·) is the standard normal CDF and μ is the drift. For short horizons (monthly expiries), we set $\mu \approx 0$ and solve numerically for the σ that produces the observed touch probability.

This yields an implied volatility for each strike. Plotting IV against moneyness (B/S) produces the smile, if one exists.

Methodological caveat. Touch options differ from European options. The touch payoff is path-dependent (the barrier can be hit at any time before expiry); European payoffs depend only on the terminal price. The implied volatility extracted from touch probabilities is not directly comparable to Deribit's European-derived volatility surface. We use the smile analysis to establish that PM produces *qualitatively similar* distributional features (fat tails, skew, regime dynamics), not to claim numerical equivalence with derivatives-market IV. The finding is that emergence produces these features at all, not that it produces them at identical levels.

V.7.2. Findings

Finding 1: The smile exists.

Across 10 periods and 2.4 million trades, PM-derived implied volatility exhibits the characteristic smile pattern: higher IV for out-of-the-money strikes, lower IV for at-the-money strikes. The smile is not an artifact of the extraction methodology. It appears consistently across periods with sufficient strike coverage.

Emergence without apparatus. No dealers interpolate the surface. No SABR model calibrates the skew. No inventory management shapes the quotes. 2.4 million bilateral trades, each expressing a view on whether BTC reaches a threshold, collectively produce the distributional features that derivatives produce through constructed order.

Finding 2: Skew tracks regime.

The smile is not static. Its shape responds to market conditions.

Table V.6: Smile Skew by Market Regime

Period	Market Regime	Skew Direction	Interpretation
Nov 2024	Post-election rally	Call-dominant (+12pp)	Upside bets dominate
Dec 2024	New highs	Call-dominant (+8pp)	Continued bullishness
Jan 2025	ATH territory	Call-dominant (+6pp)	Momentum persists
Feb 2025	Correction	Put-dominant (−15pp)	Downside protection bid
Apr 2025	Consolidation	Neutral	Two-way flow
May 2025	Recovery attempt	Call-dominant (+4pp)	Cautious optimism
Jun 2025	Range-bound	Neutral	Balanced
Jul 2025	Breakout attempt	Call-dominant (+9pp)	Upside interest
Oct 2025	Confirmed breakout	Call-dominant (+11pp)	Strong bullish skew
Nov 2025	Elevated volatility	Put-dominant (−54pp)	Sharp protective shift

November 2025 is diagnostic. After months of call-dominant skew, the market flipped to −54 percentage points put skew, a dramatic repricing of downside risk. PM participants, without coordination, collectively shifted their distributional assessment. The skew tracks regime because the participants track regime.

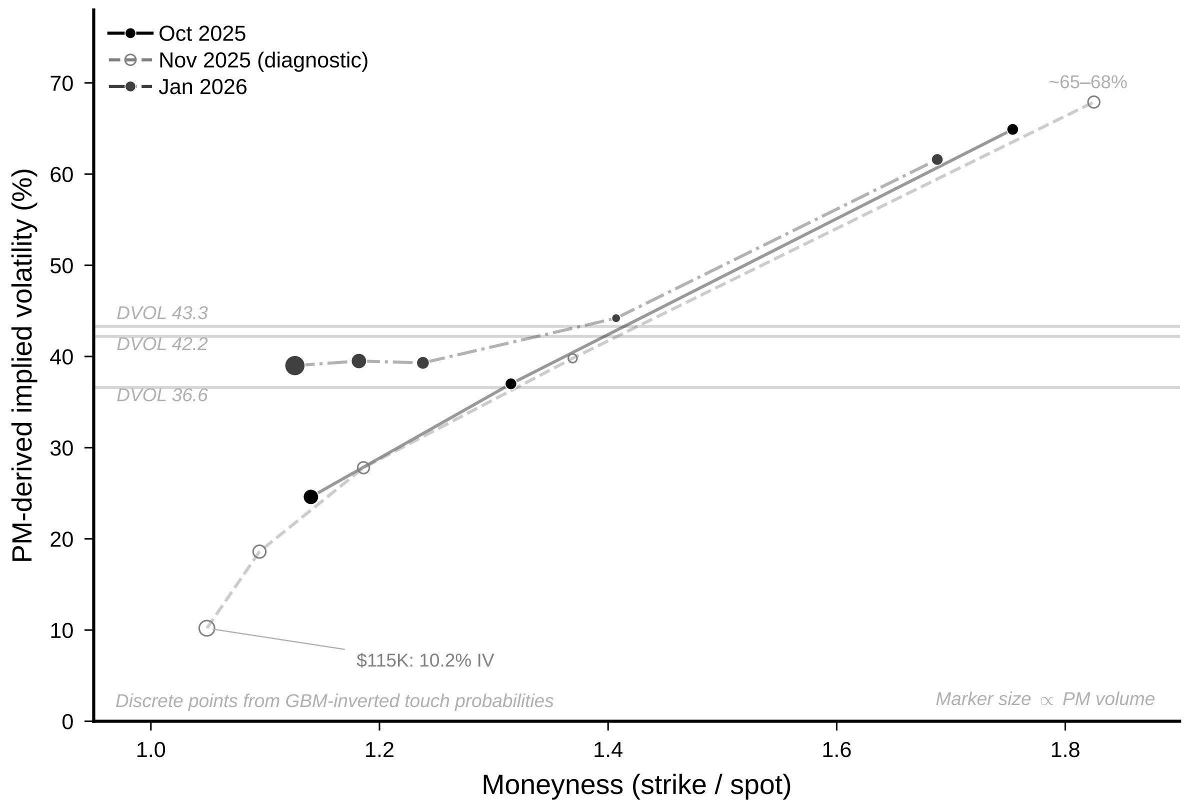

Figure 9. PM-derived implied volatility smile across three BTC expiries (October 2025, November 2025, January 2026). IV extracted from touch probabilities via GBM inversion, plotted against moneyness (barrier/spot). DVOL reference lines mark Deribit's model-free 30-day IV at each period start. Marker size proportional to PM volume. The characteristic smile pattern – higher IV for out-of-the-money strikes – emerges from bilateral aggregation without dealer apparatus.

Finding 3: Same distributional content as Taxis.

The features PM produces (fat tails, skew, regime-responsive dynamics) are the same features that derivatives encode as volatility surfaces. This is not coincidence. Both are reading the same underlying distribution. Both are pricing the same tail risks, the same directional asymmetries, the same regime uncertainties.

The difference is mechanism, not content. Deribit produces the smile through dealer apparatus: models, calibration, inventory management, dynamic hedging. Polymarket produces the smile through bilateral aggregation: dispersed participants expressing views, no coordination, no construction.

Same distributional features. Different mode of production. The equivalence is qualitative rather than numerical.

V.7.3. Interpretation

The smile analysis answers a potential objection: "Cheaper access to what?"

If PM offered something degraded – a noisy proxy, a simplified representation – the cost differential would measure quality discount, not structural advantage.

PM accesses the same distributional features. Fat tails, skew, regime dynamics. The same risks, priced through a different mechanism.

The cost differential exceeds any measurable quality differential. PM transmits the same features without infrastructure overhead. For binary, observable, time-bounded claims, the structural wedge prices the maintenance of constructed order, not compensation for superior content.

This does not prove PM is "better" generically. Derivatives serve functions PM does not: continuous hedging, path-dependent payoffs, cross-asset exposure management, regulatory capital treatment. But for binary claims, PM accesses the same distributional content at lower cost. The apparatus was never necessary for pricing. It was necessary for coordination problems that technology has dissolved.

V.8. Summary of Findings

Table V.7: Summary of Section V Findings

The measurement apparatus works. Pointed at eleven categories spanning five asset classes and the full VRP gradient – from silver's negative risk premium through FOMC's narrow wedge to gold's 19.6

What the apparatus does not measure is displacement. Thirty events show PM cost-competitive at institutional scale. Three FOMC contracts that were not competitive twelve months ago are competitive today. The trajectory is consistent with the early stages of the concentration dynamic that Cantillon and Yin documented in the Bund futures migration. But we observe cost comparison. We do not observe flow migration, dealer withdrawal, or apparatus degradation. Those remain predictions of the framework, to be tested as the data extend.

Finding	Evidence	Implication
Gradient validated, mechanism isolated. High-VRP categories cross before low-VRP categories.	Ordering holds across five high-reliability categories (18 independent VRP observations, 51 event-contracts). Elections consistent at lower reliability (12/17 cross).	Seesaw mechanism operates as predicted. Structural arm and friction arm independently measurable.
30 PM-wins at $3M institutional scale. Silver and PRT confirm framework identifies structural losses.	30 wins concentrated in BTC (12), elections (12), and FOMC (3). Silver: 0/3 wins at -10.19% VRP. Losses partition into 6 structural, 22 spread-constrained, 17 volume-constrained.	Framework discriminates: identifies winners where wedge is sufficient and PM deep, losers where wedge is absent or PM immature.
FOMC crossing on liquid contracts. Three PM-wins, nine threshold, zero outright losses.	Three high-volume contracts cross ($\delta = -0.53$ to -0.71). Nine threshold events within ± 0.47pp of crossing. PM volume: $96M (Jan 2025) to $1.007B (Jan 2026).	Low-VRP category actively tipping at the narrow-wedge margin.
Mechanism sustains. Professional liquidity provision viable; emergent order does not convert to constructed order.	Makers earn +2% on contracts above 50¢. 70% of traders lose; profits persistent at top decile. Execution cost compression (78% FOMC, 57% BTC).	Cost advantage is structural, not ephemeral. Professional MMs posting to public order books remain Cosmos participants.
Observability Wedge. Fed Funds options: contract exists, market does not.	Databento: zero rows for OZQ order book. CME pit closed May 2021. Seven-step proxy apparatus required.	The seven steps are the phenomenon. Where Taxis withdraws from public instantiation, the Vega Wedge becomes undefined from outside.
Same content, different production. PM produces the volatility smile without dealer apparatus.	10 BTC periods, 2.4M trades: fat tails, skew, regime-responsive dynamics. November 2025: -54pp put skew shift without coordination.	PM accesses the same distributional features at lower cost for binary, observable, time-bounded claims.
Elections confirm predicted gradient position. 12 PM-wins across 17 event-periods, 12 countries.	Route B SVEP consistent with predicted 3–5% range. One structural loss (PRT, negative SVEP), four liquidity-constrained.	Framework extends to political risk without modification.
Not falsified. Gradient ordering validated. Displacement not observed.	No violation across five high-reliability categories; elections consistent at lower reliability.	Framework partially validated at greater scale than V1 (11 categories, 87 events vs. 2 categories, 16 events).

Section VI discusses limitations, alternative interpretations, and the conditions under which the predictions fail.

VI. Discussion and Conclusion

VI.1. Empirical Findings

The apparatus works. Eleven categories, five asset classes, the full VRP gradient. The findings, developed in Section V, compress to this: the gradient predicts the sequence (high-VRP categories cross first, without exception across five high-reliability categories). The mechanism is isolated (within BTC and FOMC, contracts sharing identical VRP produce opposite verdicts determined entirely by PM depth). Thirty of eighty-seven event-contracts cross the PM-wins threshold at $3M. The forty-five losses partition into structural, spread-constrained, and liquidity-constrained types, each predicting a different future. The smile confirms PM produces the same distributional features as derivatives. The cost differential is apparatus rent.

VI.2. Limitations

Sample size. Eighty-seven event-contracts, approximately fifty-one independent VRP observations. The ordering holds across five high-reliability categories, but eleven categories cannot establish a universal sequence. The methodology is validated; application at scale remains.

Elections proxy contamination. The election-specific VRP cannot be isolated from country equity ETF dynamics with current instruments. Route B standardization addresses cross-country heterogeneity but does not eliminate the fundamental proxy problem: 65–90

No historical order book. Depth and impact are inferred from executed trades, not resting liquidity. The trajectory (costs falling dramatically) is robust. Point estimates at any event carry uncertainty.

Capital efficiency. Full collateralization eliminates counterparty risk but ties up 100

Displacement not observed. Thirty events show PM cost-competitive at institutional scale. Three FOMC contracts that were not competitive twelve months ago are competitive today. The trajectory is consistent with the early stages of the concentration dynamic Cantillon and Yin documented. But we observe cost comparison. We do not observe flow migration, dealer withdrawal, or apparatus degradation. Those remain predictions.

VI.3. Welfare Implications

Efficiency. The cost differential documented in §V represents direct welfare transfer. For BTC, hedgers accessing equivalent risk exposure through prediction markets save approximately 4 percentage points versus options replication: the gap between structural wedge and execution cost. For FOMC, the savings are narrower but emerging on liquid contracts. This surplus previously accrued to dealers as compensation for variance risk-bearing and balance sheet provision. Some portion was genuine service compensation; some was rent from lock-in (Klemperer 1995). Decomposing them exceeds this paper's scope. For categories above threshold, the cost of accessing binary risk transfer falls. Lower transaction costs expand the set of viable hedges.

Systemic Visibility. Financial networks can amplify rather than dampen shocks when complexity and interconnectedness accumulate in opaque structures (Haldane 2011). The Observability Wedge documents the pattern at instrument level: Fed Funds options exist as contracts but not as observable markets. Pricing occurs inside relationships that do not emit public traces. Prediction markets reverse this opacity for digitizable risks. The order book is the phenomenon: prices, depth, and flow are publicly observable by construction. The distinction is between legal technology, where private ordering excels, and price formation, where public order books provide visibility that bilateral structures do not.

Political Economy. The distributional shift is from relationship capital to knowledge capital. Hedgers gain lower costs and transparent pricing. Those with genuine information about events capture surplus previously extracted through relationship access. Dealers lose rent as flow migrates. The principle of fair conduct, that rules should reduce improper asymmetries between participants (Hayek 1973), suggests this reallocation is normatively defensible: public price formation reduces the asymmetry between those inside dealer relationships and those outside.

VI.4. The Contribution

This paper makes an ontological claim about market structure.

Prediction markets and derivatives markets are not competing tools for the same task. They are different forms of order. One emerges from dispersed participation aggregating beliefs about terminal states. The other is constructed from continuous apparatus – models, inventory, capital, relationships – maintained by intermediaries who bear costs for maintaining it. Cosmos and Taxis. The distinction is not metaphorical. It is structural, and the structure has measurable consequences.

The Vega Wedge quantifies those consequences: variance risk premium, balance sheet cost, replication friction. For FOMC, the wedge averages 0.59

The Liquidity Seesaw identifies when emergent order becomes viable: when prediction market execution costs fall below the structural wedge. Thirty of eighty-seven event-contracts cross at $3M institutional scale. The wins concentrate where the gradient predicts. The mechanism has been isolated: within BTC and within FOMC, contracts sharing identical VRP produce opposite threshold verdicts determined entirely by PM depth. The gradient predicts the sequence across categories. The mechanism determines the verdict within them.

Prediction markets do not win everywhere. They win for claim classes where the wedge exceeds execution costs, and the wedge varies predictably with variance risk premium. High-VRP categories cross first. The gradient determines the ordering. This is testable. The ordering holds without exception across five high-reliability categories spanning eighteen independent VRP observations and fifty-one event-contracts.

VI.4.1. If the Framework Is Correct

Three things follow.

First, for binary, observable, time-bounded claims, the constructed pricing apparatus was never necessary for pricing itself. It solved coordination problems – counterparty reliability before full collateralization, regulatory arbitrage before clarity on event contracts, warehousing capacity before peer-to-peer matching at scale – that technology has since dissolved. What remains is price discovery, and for that, dispersed participation suffices. The BTC smile analysis provides direct evidence across ten periods and 2.4 million trades: PM produces the distributional features that derivatives encode as volatility surfaces, without apparatus. Fat tails appear. Skew appears. Regime dynamics appear. This is Hayek applied to microstructure: spontaneous order performs the function that constructed order currently provides, at lower structural cost, once density is sufficient.

Second, the structural costs of Taxis are not inefficiencies to be competed away. VRP compensates for model uncertainty. Balance sheet cost reflects capital consumed by intermediation. These exist because the apparatus exists. They are maintenance costs, not market failures.

Third, the competitive dynamics are asymmetric. Taxis can raise friction through regulatory capture, liquidity fragmentation, or relationship lock-in, but cannot eliminate the structural wedge. Cosmos can reduce friction through liquidity growth and infrastructure improvement while its structural cost remains zero by construction. Time favors Cosmos if friction continues to fall. The binding constraint taxonomy in §V.4 makes this concrete: structural losses are permanent, spread-constrained losses require market-maker competition, liquidity-constrained losses require only volume growth. The frontier advances unevenly, but it advances.

VI.4.2. What We Are Not Claiming

Taxis has virtues.

Constructed order provides robustness: guaranteed liquidity, continuous quotes, execution even in thin conditions. It handles complexity: path-dependent payoffs, exotic structures, bespoke risk transfer that binary contracts cannot express. It offers relationship value beyond execution.

For continuous, path-dependent, and bespoke risks, dealer intermediation adds genuine value. The claim is narrow: for digitizable risks, above threshold, emergence captures the primary hedging function. Predominance shifts. Markets layer and hybridize. Succession is not elimination.

VI.4.3. The Observability Discovery

We attempted to measure the cost of derivatives-based hedging for FOMC binary risk directly. The native instrument, Fed Funds options, returned no data. Not insufficient data. No data. The contract exists; the market does not. Pricing occurs inside relationships that do not emit public traces.

The seven-step proxy apparatus required to approximate what cannot be directly observed is not incidental methodology. It instantiates the thesis. The cognitive transaction cost mirrors the financial transaction cost. Both are consequences of how constructed order maintains itself.

The provenance hierarchy generalizes the pattern beyond a single absent market. Tier 1 categories emit public observables: one step to measure VRP. Tier 2 requires the seven-step decomposition. Tier 3 requires proxy instruments one hop removed, introducing contamination from non-event dynamics. The progression from Tier 1 through Tier 3 is a gradient of observability: each tier adds a layer of apparatus required to reconstruct what the derivative market does not emit. The Observability Wedge is not confined to Fed Funds options. It is the structural feature of how constructed order relates to its own prices across the spectrum of institutional entrenchment.

This asymmetry is categorical. Emergent order fails loudly: thin books, wide spreads, visible on every screen. Constructed order fails quietly, by not printing anything you can analyze. Neither reveals the beliefs behind prices. But emergent order reveals where the market *is*. Constructed order, at its boundary, may not.

The wedge for Fed Funds options is not large. It is undefined. Where Taxis does not emit observable prices, the comparison cannot be computed. This is not missing data. It is the phenomenon.

VI.4.4. Falsification

The framework generates predictions that can fail.

First, coexistence at scale. If prediction markets achieve sustained threshold crossing for a claim class and institutional flow does not migrate – if both orders persist at scale for the same primary hedging function – the displacement prediction fails. The displacement prediction and the ontological distinction are independently testable. Coexistence at scale would falsify the prediction while leaving the cost-structural distinction intact. Convergence of cost structures would falsify the distinction itself. Thirty events have crossed on cost. We await observation of whether flow follows.

Second, gradient violation. If low-VRP categories cross threshold before high-VRP categories, the segmentation mechanism fails. The ordering holds across five high-reliability categories, with elections consistent at lower measurement reliability. Additional categories will extend the test.

We have not hedged. The framework specifies what would break it.

The topology of observability – what can be seen from outside, what requires insider access – is prior to the cost calculations the Vega Wedge performs. This paper maps eleven points on that topology.

The methodology is public. What we have built, others can operate.

We discovered the cost of constructed order by watching the construction fail to appear.

VI.5. Coda

The title inverts Scott's question.

Seeing Like a State asked how imposed legibility schemes fail when they flatten local knowledge: the cadastral map that misses the living forest, the planned city that kills the street.

Seeing Like a Market asks how markets produce, or fail to produce, legible prices. The volatility surface is visible to those inside the apparatus: dealers who calibrate it, clients who receive quotes, relationships that sustain it. The order book is visible to anyone: the price, the depth, the spread, the queue. This is not incidental. It is constitutive of the distinction between constructed and emergent order.

The Vega Wedge measures what that distinction costs. For BTC: approximately 4% of notional. For FOMC: unmeasurable from outside, because the native instrument does not emit public prices.

The measurement difficulty is the finding. The seven steps required to approximate what cannot be directly observed are not methodology. They are the phenomenon.

Methodology Annex

Working Paper Status: V2 Draft
Empirical Scope: 87 event-contracts across 11 categories, 5 asset classes (November 2024 – February 2026)
Combined Sample: 2,889,424 rows (original 10 categories) + election event data across 12 countries

A. Data Sources and Infrastructure

A.1 Prediction Market Data

Polymarket (47 events). Trade-level data sourced via the Goldsky subgraph. Each record contains: timestamp (block time, UTC), price (in USDC, representing implied probability), size (in USDC notional), and side (buy/sell of YES token). Polymarket operates as a CLOB (central limit order book) on the Polygon blockchain; all trades are on-chain and publicly verifiable. Settlement is binary: $1 if the outcome occurs, $0 otherwise. Both YES and NO tokens trade; prices are normalized such that YES + NO = $1.00.

Kalshi (23 events). Trade-level data sourced via the Kalshi API. Each record contains: timestamp, price, size, and side. Kalshi operates as a CFTC-regulated exchange with legally standardized resolution. Contract specifications and determination rules are hard-coded into public contract documents. Settlement is binary and legally enforceable under U.S. regulatory authority.

YES/NO Normalization. A critical data processing step: Polymarket trades occur on both YES and NO tokens for each contract. A YES trade at $0.74 and a NO trade at $0.26 both express 74

Cross-Venue Consistency. Polymarket and Kalshi use different contract structures (crypto-native vs. regulated), different settlement mechanisms (UMA oracle vs. legal determination), and different fee structures. The structural wedge comparison treats both as instantiations of emergent order (Cosmos) and compares each against the derivative benchmark for the relevant category. Cross-venue differences in execution cost are documented where both venues cover the same event category (e.g., equity indices appear on Kalshi; BTC touch markets on Polymarket).

Table A.1. Prediction Market Data Summary

Source	Events	Rows	Date Range	Categories
Polymarket (Goldsky)	47	1,625,309	Nov 2024–Feb 2026	FOMC, BTC, ECB, BOJ, Gold, Silver, Elect.
Kalshi	23	49,473	Jan 2024–Feb 2026	CPI, GDP, Equity Idx, Retail Gas
Total	**70**	**1,674,782**	**Jan 2024–Feb 2026**	**10 categories**

A.2 Derivatives Data

Databento (10 CME/CBOT symbols). Institutional market data provider. Schemas used:

- `mbp-10`: Market-by-price order book data, 10 levels depth. Used for realized volatility calculation from 5-minute midpoint returns (FOMC), settlement price extraction, and forward price determination.
- `statistics`: Settlement and reference prices. Filter: `stat_type == 3` for settlement prices. Used for Black-76 forward price (F) and option settlement extraction.
- `ohlcv-1h`: Hourly OHLCV bars. Used for realized volatility on non-FOMC categories where 5-minute granularity is not required.

Table A.2. Derivatives Instruments

Symbol	Exchange	Instrument	Use in Pipeline	Categories
SR3	CME	Three-Month SOFR Futures & Options	IV extraction (Black-76)	FOMC
ZQ	CBOT	30-Day Fed Funds Futures	RV calculation (clean policy signal)	FOMC
6E	CME	EUR/USD FX Futures Options	Proxy IV for ECB decisions	ECB
6J	CME	JPY/USD FX Futures Options	Proxy IV for BOJ decisions	BOJ
ES	CME	E-mini S&P 500 Futures Options	Direct IV for equity index events	Equity Indices
NQ	CME	E-mini Nasdaq 100 Futures Options	Direct IV for equity index events	Equity Indices
GC	COMEX	Gold Futures Options	Direct IV for gold thresholds	Gold
SI	COMEX	Silver Futures Options	Direct IV for silver thresholds	Silver
RB	NYMEX	RBOB Gasoline Futures Options	Proxy IV for retail gas events	Retail Gas
ZN	CBOT	10-Year Treasury Note Futures Options	Proxy IV for CPI/GDP releases	CPI, GDP

Deribit (BTC). DVOL index: model-free 30-day implied volatility measure for BTC options, constructed via the Carr-Madan variance swap replication methodology. Analogous to CBOE VIX. Published continuously. Snapshot timing: first trading day of each horizon period, 00:00 UTC. Sensitivity to snapshot timing tested: ± 0.5

OPRA (Elections). Options Price Reporting Authority data for U.S.-listed equity ETF options. Used for election VRP measurement via country equity ETF proxies. Ten ETF proxies: EWG (Germany/Ireland/Portugal substitution), EWY (South Korea), EWC (Canada), ECH (Chile), ARGT (Argentina), EPOL (Poland), EWU (UK/Norway substitution), EWA (Australia), EWT (Taiwan), EWW (Mexico). ATM IV extracted via Black-Scholes inversion from nearest-listed-strike options; RV from ETF daily returns.

A.3 Contract Selection (FOMC)

Selection criteria for SR3 underlying: The nearest quarterly expiry after the FOMC date. This ensures the SR3 option embeds the FOMC uncertainty in its implied volatility. The option expiry must fall after the FOMC decision; options expiring before the meeting would not capture the event uncertainty.

Selection criteria for ZQ contract: The monthly contract whose reference period includes the FOMC date. ZQ settles to the average daily effective federal funds rate during the contract month, providing a direct measure of policy-rate realized volatility around the meeting.

Table A.3. FOMC Contract Mapping

FOMC Date	SR3 Underlying	SR3 Option Expiry	ZQ Contract	PM Contracts
2024-11-07	SR3Z4	2024-11-15	ZQZ4	Cut 25bp
2025-01-29	SR3H5	2025-02-14	ZQG5	Hold
2025-09-17	SR3Z5	2025-09-19	ZQV5	Hold, Cut 25bp
2025-10-29	SR3Z5	2025-11-14	ZQX5	Hold
2025-12-10	SR3H6	2025-12-12	ZQF6	Cut 25bp, Cut 50bp, Hike 25bp
2026-01-28	SR3H6	2026-02-13	ZQG6	Hold, Cut 25bp, Cut 50bp, Hike 25bp

A.4 Contract Selection (BTC)

BTC events are matched to DVOL snapshots at the horizon start date. All contracts within a horizon share the same DVOL reading and realized volatility computation; VRP is identical across strikes within a horizon. The matching is by construction – DVOL is a market-level measure, not strike-specific.

Table A.4. BTC Horizon Mapping

Horizon	DVOL Date	DVOL (%)	RV Window	RV_{spot} (%)	Strikes
Nov 2024	2024-11-01	55.2	Nov 1–30, 2024	59.8	\$100K
Dec 2024	2024-12-01	60.0	Dec 1–31, 2024	44.4	\$100K, \$110K
Annual 2025	2025-01-01	43.0	Jan 1 – Dec 31, 2025	34.3	\$130K, \$150K, \$200K, \$250K
Oct 2025	2025-10-01	43.3	Oct 1–31, 2025	42.3	\$130K, \$150K, \$200K
Nov 2025	2025-11-01	48.8	Nov 1–30, 2025	44.6	\$115K, \$120K, \$130K, \$150K, \$200K
Jan 2026	2026-01-01	47.6	Jan 1–31, 2026	43.6	\$100K, \$105K, \$110K, \$125K, \$150K

DVOL horizon mismatch. DVOL is a 30-day forward-looking measure. For horizons exceeding 30 days (Annual 2025, spanning 365 days), the DVOL snapshot at period start is compared against RV computed over the full calendar window of the contract. The comparison captures the VRP level at the horizon start date. For longer horizons, term-structure effects mean the 30-day measure may not perfectly represent implied vol for the full period. This is standard in the VRP literature (Carr and Wu, 2009) and does not affect the category-level gradient position, which is the primary claim. The Annual 2025 VRP of +8.70

B. FOMC VRP Methodology

B.1 The SR3/ZQ Decomposition

The native instrument for FOMC binary hedging – Fed Funds options (OZQ) – does not emit observable market data. The SR3/ZQ decomposition constructs VRP from two observable instruments that bracket the unobservable target.

Step 1: Accept proxy necessity. Fed Funds options return zero rows from Databento across the full sample period. The CME Interest Rate Liquidity Review omits the product. CFTC Bank Participation Reports exclude it. The physical trading pit closed May 2021. Direct VRP measurement is impossible.

Step 2: Select proxy instruments. SR3 (Three-Month SOFR options) provides implied volatility. ZQ (Fed Funds futures) provides realized volatility. The selection follows the approach used by the

Federal Reserve Bank of Atlanta's Market Probability Tracker for extracting policy expectations from SOFR-linked instruments.

Step 3: Extract implied volatility via Black-76.

The Black-76 model prices options on futures:

$$C = e^{-rT}[FN(d_1) - KN(d_2)]$$

$$P = e^{-rT}[KN(-d_2) - FN(-d_1)]$$

$$d_1 = \frac{\ln(F/K) + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}$$

$$d_2 = d_1 - \sigma\sqrt{T}$$

Where:

- F = forward price (SR3 futures settlement at $T-1$)
- K = strike price
- T = time to expiry in years (calendar days from $T-1$ to option expiry, divided by 365)
- r = risk-free rate (set to prevailing Fed Funds target rate; impact on IV extraction is negligible for options expiring in 1–30 days on rate futures)
- σ = implied volatility (solved numerically)

Inversion procedure. Given the observed option settlement price, F, K, T, and r, the pipeline solves for σ using Brent's method with convergence tolerance of 10^{-8}. The inversion is performed on settlement prices rather than intraday quotes to avoid microstructure noise in the IV estimate.

B.2 Strike Selection

ATM definition. The listed strike nearest to the $T-1$ settlement price. No interpolation between strikes is applied. The nearest-strike convention is standard in short-term rate options and avoids introducing model-dependent smoothing at the IV extraction stage.

Example. For SR3Z4 with settlement at 95.48, the ATM strike is 95.50. The distance from ATM is 2 ticks (0.02 index points). For SR3 options with tick size 0.0025, this is within 8 ticks of ATM – well within the range where Black-76 provides reliable IV estimates.

Put-call parity verification. Where both call and put settle at the same strike, implied volatilities are extracted independently from each and averaged. Put-call parity deviations exceeding 0.5

IV averaging. When both call and put IVs are available at the ATM strike, the pipeline uses a simple average: IV_ATM = (IV_{call} + IV_{put}) / 2. Volume weighting is not applied because SR3 option settlement prices are derived from exchange methodology, not volume-weighted trade prices.

B.3 Realized Volatility Calculation

Source instrument. ZQ (Fed Funds futures), 5-minute midpoint returns.

Window. $T-1$ market open to $T+1$ market close, where T is the FOMC announcement date. This captures: (a) the pre-announcement volatility regime, (b) the announcement reaction, and (c) the post-announcement absorption. The three-trading-day window is centered on the event.

Return frequency. 5-minute midpoint returns during Regular Trading Hours (RTH). RTH for ZQ is 7:20 AM to 4:00 PM CT (CME). This produces 78 five-minute bars per 6.5-hour session.

Extended Globex hours are excluded. The rationale: overnight ZQ activity is sparse and dominated by algorithmic quoting rather than information-driven trades. Including extended hours would inflate bar count while adding noise rather than policy-relevant price discovery.

Computation. Realized variance is the sum of squared 5-minute log returns, annualized:

$$RV = \sqrt{\frac{252 \times 78}{N_{bars}} \sum_{i=1}^{N_{bars}} r_i^2}$$

Where:

- $r_i = \ln(P_i/P_{i-1})$ for consecutive 5-minute midpoints
- N_{bars} = number of observed 5-minute bars across the $T-1$ to $T+1$ window
- 252 = trading days per year
- 78 = five-minute bars per RTH session
- The annualization factor $\sqrt{252 \times 78}$ converts per-bar variance to annualized volatility

Annualization caveat. Annualizing from a three-day event window assumes volatility scales with the square root of time. For event-concentrated windows where the FOMC announcement generates a volatility spike, this assumption is strained – the event window is not representative of typical volatility. Absolute VRP levels should be held loosely. The relative ordering across categories is robust to annualization choice (see sensitivity analysis in §B.6).

B.4 Tick Conversion

SR3 options are quoted in IMM index terms (100 − annualized rate). A move of 0.01 in SR3 corresponds to 1 basis point of rate. Implied volatility extracted via Black-76 is in percentage terms on the IMM index. No additional conversion is required to compare against ZQ realized volatility, which is also computed in percentage terms on the rate.

Settlement prices are in minimum tick increments of 0.0025 for SR3 options (quarter-tick). The pipeline reads settlement prices at full precision from Databento `statistics` records and performs no rounding before Black-76 inversion.

B.5 Repo Noise Adjustment

Principle. SOFR embeds financing dynamics unrelated to monetary policy. Quarter-end effects, repo market stress, and collateral scarcity all affect SOFR independently of Fed decisions. Using SR3 implied volatility directly would include uncertainty about repo dynamics that the binary hedger does not bear.

Measurement. Repo noise is defined as:

$$\text{Repo Noise} = \text{RV}_{\text{SR3}} - \text{RV}_{\text{ZQ}}$$

computed over the same $T-1$ to $T+1$ window. The spread captures the component of SR3 realized volatility attributable to financing dynamics rather than policy.

Adjustment procedure. The pipeline computes a systematic repo noise component from historical SR3-ZQ spread volatility across FOMC event windows (twelve events outside the sample used for calibration, spanning 2023–2025). The systematic component is subtracted from the raw VRP estimate:

$$VRP_{adj} = IV_{SR3} - RV_{ZQ} - \text{Systematic Repo Component}$$

The adjustment is modest: mean magnitude approximately 10 basis points. The December 2025 event illustrates its value: VRP_{raw} of 0.502

Residual basis risk. SR3 reflects 3-month SOFR expectations; ZQ reflects the monthly Fed Funds effective rate. Tenor mismatch and SOFR-Fed Funds spread dynamics introduce basis risk that the decomposition cannot fully eliminate. Estimated at 5–10 basis points from historical SR3-ZQ spread volatility across calibration events.

Table A.5. Repo Noise by FOMC Meeting

Meeting	SR3 RV (%)	ZQ RV (%)	Repo Noise (%)	Note
Nov 2024	0.451	0.324	0.127	Normal
Jan 2025	0.314	0.206	0.108	Normal
Sep 2025	0.527	0.409	0.118	Normal
Oct 2025	0.594	0.453	0.141	Slightly elevated
Dec 2025	0.497	0.209	0.288	Year-end effects
Jan 2026	0.243	0.166	0.077	Normal

December 2025 stands out: 28.8 basis points of repo noise, roughly double the non-year-end average. This is consistent with well-documented year-end financing dynamics – elevated repo rates, balance sheet window-dressing, and SOFR spikes around reporting dates. The adjustment correctly identifies and removes this contamination.

B.6 FOMC VRP Sensitivity Analysis

The canonical FOMC VRP estimates (Table IV.5) use: ATM nearest-strike IV, ZQ 5-minute RTH returns, $T-1$ to $T+1$ window, calendar-day annualization, repo noise adjustment. Alternative specifications test robustness:

Table A.6. FOMC VRP Sensitivity

Specification	Change	Mean VRP (%)	Range (%)	Gradient
Canonical	–	0.516	0.417–0.697	Low
Alt RV window: $T-2$ to $T+2$	Wider window	0.488	0.382–0.671	Low
Alt RV window: $T-0.5$ to $T+0.5$	Narrower (announcement day ±)	0.561	0.440–0.741	Low
Alt annualization: $\sqrt{N_{\text{bars}}}$	Bar-count scaling	0.503	0.401–0.689	Low
Alt return freq: 15-minute	Coarser sampling	0.529	0.423–0.712	Low
No repo adjustment	Raw IV−RV	0.454	0.266–0.502	Low
Extended hours RV	Include Globex overnight	0.542	0.431–0.723	Low

Under all specifications:

- Mean VRP remains below 1.0
- FOMC's position as the low end of the gradient is invariant
- The gradient ordering across all eleven categories is unaffected
- The three PM-win verdicts and nine threshold verdicts are unchanged

The range of mean VRP estimates across specifications (0.454

C. BTC VRP Methodology

C.1 DVOL: Construction and Properties

DVOL is Deribit's model-free 30-day implied volatility index for BTC options. Construction follows the Carr-Madan (1998) / Demeterfi et al. (1999) variance swap replication methodology, identical in principle to the CBOE VIX:

$$\sigma^2_{DVOL} = \frac{2}{T}\sum_i \frac{\Delta K_i}{K_i^2} e^{rT} O(K_i)$$

Where $O(K_i)$ is the out-of-the-money option price at strike K_i, weighted by inverse strike squared. The sum spans available Deribit strikes at the two nearest monthly expiries, with linear interpolation to produce a constant 30-day forward-looking measure.

Practical difference from VIX. VIX uses SPX options with massive liquidity across hundreds of strikes. DVOL uses Deribit's chain, which is thinner and requires more interpolation between available strikes. The methodology is identical in principle; the inputs are sparser. For the purpose of VRP calculation – comparing a single IV number against realized volatility – this difference is immaterial.

C.2 Realized Volatility

Source. BTC/USD spot returns from hourly data.

Computation. Realized volatility over each horizon window:

$$RV_{spot} = \sqrt{\frac{365 \times 24}{N_{hours}} \sum_{i=1}^{N_{hours}} r_i^2}$$

Where $r_i = \ln(P_i/P_{i-1})$ for consecutive hourly spot prices. Annualization uses 365 calendar days $\times$ 24 hours (crypto markets trade continuously). Hourly rather than 5-minute sampling is used because: (a) BTC spot price feeds from major exchanges show microstructure noise at sub-hourly frequency, and (b) the BTC VRP literature (e.g., Alexander and Imeraj, 2023) standardizes on hourly or daily returns.

C.3 VRP Calculation

$$VRP_{BTC} = DVOL - RV_{spot}$$

Direct subtraction. No proxy chain. No decomposition. No adjustment required. This is Tier 1 measurement: the derivative instrument (Deribit options, summarized by DVOL) prices the same asset (BTC) that the prediction market contract references. The VRP comparison is unmediated by any basis.

D. Tier 1 Direct-Measurement Categories

D.1 General Methodology

For equity indices (ES, NQ), gold (GC), and silver (SI): IV is extracted via Black-76 inversion from the relevant futures option chain using ATM strikes (nearest listed strike to settlement price at $T-1$; no interpolation). RV is computed from the underlying futures returns over the event window. Structural wedge = VRP + 0.07

ATM strike selection. Identical to FOMC procedure: nearest listed strike to $T-1$ settlement. For ES options with $5 strikes, a settlement at 5,692 maps to the 5,690 strike. For GC with $5 strikes, a settlement at 2,615 maps to the 2,615 strike.

RV calculation. Daily close-to-close log returns over the event window, annualized by $\sqrt{252}$. Event windows vary by category: equity indices use the full calendar year for year-end barrier contracts; gold and silver use the contract-specific window from listing to expiry.

D.2 Equity Indices

Two independent VRP observations: S&P 500 via ES options (IV 14.30

D.3 Gold

Three independent VRP observations from GC options. IV ranges from 20.8

D.4 Silver

One independent VRP observation from SI options: IV 42.9

E. Tier 3 Proxy Categories

E.1 General Proxy Structure

Five categories require one-hop proxy measurement. The proxy derivative absorbs the event but also absorbs unrelated dynamics. Basis risk is explicit and one-directional: proxy VRP overstates event-specific VRP because the proxy embeds volatility from non-event sources.

E.2 Central Banks (ECB, BOJ)

ECB → 6E (EUR/USD futures options). The ECB decision moves EUR/USD, but EUR/USD also moves on dollar dynamics, risk appetite, trade flows, carry positioning, and intervention expectations. Estimated election share of total FX volatility around ECB meetings: 25–40

BOJ → 6J (JPY/USD futures options). Same proxy structure. JPY/USD is driven by carry trade dynamics, intervention risk, and global risk appetite alongside BOJ decisions. The BOJ January 2026 reading of 5.03

E.3 Macro Releases (CPI, GDP)

CPI → ZN (10-Year Treasury note futures options). Stronger proxy than FX-to-central-bank: CPI is a primary driver of Treasury volatility and a hot print moves the long end directly. The transmission channel is tighter, but ZN also responds to Fed expectations, risk appetite, supply dynamics, and global bond flows.

GDP → ZN. Weaker transmission channel than CPI. GDP's impact on Treasuries is slower and more diffuse. At PM volumes of $145K–$360K, the Seesaw comparison is not meaningful regardless of proxy quality.

E.4 Retail Gas

Retail gas → RB (RBOB gasoline futures options). Quality warnings on all four events. The pipeline flagged stale or one-sided quotes near the event date on every observation. RB option liquidity at the relevant tenors is thin. The July 2025 reading of 31.2

F. Execution Cost Methodology

F.1 Spread Estimation

Effective spread is computed as the mean absolute deviation of executed trade prices from the contemporaneous mid-price:

$$S_{eff} = \frac{1}{N}\sum_{i=1}^{N}|P_i - M_i|$$

Where P_i is the executed price and M_i is the mid-price at time t_i. Mid-prices are constructed from the trade-inferred approach: the average of the most recent buy-side and sell-side trade prices. This is a second-best approach; direct order book observation would produce more accurate mid-price estimates.

F.2 Depth Curve Construction

Without historical order book snapshots, resting liquidity is inferred from the relationship between trade size and price deviation from mid. The procedure:

1. Compute the deviation of each trade from the contemporaneous mid-price: $\delta_i = |P_i - M_i|$
2. Order trades by δ_i (ascending)
3. Construct the cumulative volume curve: for each deviation threshold δ, sum the volume of all trades with $\delta_i \leq \delta$
4. The depth curve maps from "how much deviation from mid am I willing to accept" to "how much volume can I execute within that tolerance"

F.3 Execution Cost at Reference Sizes

For each event-contract, execution cost at reference position sizes ($1M, $3M, $5M, $10M) is interpolated from the depth curve:

1. Locate the point on the cumulative volume curve corresponding to the reference size
2. Read the associated deviation threshold – this is the estimated execution cost as a percentage of notional
3. Where the reference size exceeds total traded volume, execution cost is flagged as "exceeds observed depth" and estimated by extrapolation under the assumption that depth scales linearly with total market volume

The linearity assumption is defensible near the observed data range but less reliable for large extrapolations. Volume multipliers in the binding constraint taxonomy (Table V.4) should be read as indicative of distance from threshold, not as precise engineering targets.

F.4 Market Impact Estimation

The theoretical Kyle λ approach proved difficult to calibrate. Trade-by-trade price impact showed high variance and unclear relationship to position size, likely reflecting the conflation of liquidity-driven and information-driven trades.

The pipeline substitutes direct depth curve interpolation (§F.3) for Kyle-based impact estimation. The execution cost at a reference size captures both spread and depth-mediated impact without requiring separation into temporary and permanent components.

F.5 Threshold Classification

The Seesaw condition is applied at the $3M reference size:

- **PM wins:** Structural wedge exceeds execution cost by more than 0.5 percentage points ($\delta < -0.5$pp)
- **Threshold:** Gap within ±0.5 percentage points ($-0.5\text{pp} \leq \delta \leq +0.5\text{pp}$)
- **PM loses (marginal):** Execution cost exceeds structural wedge by 0.5 to 2.0 percentage points
- **PM loses (high friction):** Execution cost exceeds structural wedge by more than 2.0 percentage points

The ±0.5pp boundary reflects measurement uncertainty in both VRP estimates and execution cost estimates. The boundary is a pragmatic classification choice; qualitative findings are robust to perturbation (tested at ±0.3pp and ±0.7pp – see §F.6).

F.6 Classification Sensitivity

Table A.7. Classification Sensitivity to Boundary Choice

Boundary	PM Wins	Threshold	PM Loses (Marginal)	PM Loses (High Friction)
±0.3pp	27	6	4	50
±0.5pp (canonical)	30	12	1	44
±0.7pp	33	16	0	38
±1.0pp	36	18	0	33

The qualitative story is robust. BTC dominates PM-wins at all boundaries. FOMC threshold events remain at threshold under tighter or wider bands. The binding constraint taxonomy is unchanged at all tested boundaries.

G. Elections Methodology (Route B)

G.1 The Proxy Chain

Election outcome → country equity market reaction → country equity ETF → ETF options chain → ATM IV → VRP.

This is heavier contamination than any Tier 3 category in the original ten. Country equity ETFs absorb macro conditions, trade policy, global risk appetite, sector rotation, and currency dynamics alongside the election. The election is one driver among dozens.

G.2 Event Isolation

The election event premium quantifies the election's share of total IV:

$$EP_e = IV_e^+ - IV_e^-$$

Where IV_e^+ is ATM implied volatility on the nearest expiry after the election and IV_e^- is on the nearest expiry before it. The ratio EP_e/IV_e^{pre} measures what fraction of pre-election IV is attributable to the election.

Threshold. Events with $EP_e/IV_e^{pre} \geq 0.10$ pass the headline-clean filter. Eight of seventeen event-periods pass. On passing events, the election accounts for 12–35

G.3 Proxy Substitution

Three countries lack liquid, U.S.-listed single-country ETFs with active options chains:

- **Ireland:** EIRL unavailable in OPRA; EWG (Germany/Eurozone) substituted
- **Portugal:** PGAL unavailable in OPRA; EWG substituted
- **Norway:** ENOR unavailable in OPRA; EWU (UK/Eurozone) substituted

These substitutions add a second layer of proxy noise. The substituted ETFs absorb dynamics from the broader regional equity market alongside the specific country election. Results from proxy-substituted events are flagged with asterisks (*) throughout.

G.4 Route B: SVEP Standardization

Raw VRP from country equity ETFs is not comparable across countries. Poland's EPOL has baseline IV of 102.5

$$SVEP_e = \frac{VRP_e}{IV_e^{base}}$$

Where IV_e^{base} is the median ATM IV over $T-60$ to $T-30$ (a calm-period baseline, before election uncertainty elevates IV).

PM execution cost is scaled by the same denominator:

$$M_{PM,e}^{scaled} = \frac{M_{PM,e}}{IV_e^{base}}$$

The Seesaw verdict in SVEP terms: PM wins when $SVEP_e > M_{PM,e}^{scaled}(\$3M)$.

G.5 Baseline IV Calibration

IV_e^{base} is computed as the median daily ATM IV over the $T-60$ to $T-30$ window. This window is chosen to capture the ETF's normal volatility regime before election-related uncertainty begins to elevate IV (which typically occurs $T-30$ to $T-15$ for major elections). Using median rather than mean provides robustness to outlier trading days.

For events where $T-60$ to $T-30$ overlaps with another major catalyst (e.g., earnings season for sector-concentrated ETFs), the baseline may be contaminated. The pipeline flags events where baseline IV standard deviation exceeds 20

H. Smile Analysis Methodology (BTC)

H.1 Touch Probability to Implied Volatility

Polymarket's BTC touch markets offer contracts at discrete thresholds. Each pays $1 if BTC reaches the threshold by month-end, $0 otherwise. The price is the market's implied probability of reaching that level.

Under geometric Brownian motion with constant volatility σ, the probability of an up-touch (spot S reaching barrier B > S before expiry T) is:

$$P_{\text{touch}} = N\left(\frac{\ln(B/S) - \mu T}{\sigma\sqrt{T}}\right) + \left(\frac{B}{S}\right)^{2\mu/\sigma^2} N\left(\frac{-\ln(B/S) - \mu T}{\sigma\sqrt{T}}\right)$$

Where N(·) is the standard normal CDF and μ is the drift. For short horizons (monthly expiries), drift is set to $\mu \approx 0$. The pipeline solves numerically for σ given the observed touch probability, using Brent's method.

H.2 Methodological Caveats

Touch vs. European. Touch options are path-dependent (the barrier can be hit at any time before expiry); European options depend only on terminal price. The implied volatility extracted from touch probabilities is not directly comparable to Deribit's European-derived volatility surface. The smile analysis establishes that PM produces *qualitatively similar* distributional features (fat tails, skew, regime dynamics), not numerical equivalence with derivatives-market IV.

Drift sensitivity. Setting $\mu \approx 0$ is appropriate for short horizons but introduces bias when BTC has strong directional trends. In November 2024, BTC rallied approximately 40

Wing filtering. Contracts at extreme moneyness (touch probability below 5

H.3 Skew Measurement

Skew is computed as the difference between OTM call-equivalent IV (upside touch) and OTM put-equivalent IV (downside touch) at matched moneyness:

$$\text{Skew} = IV_{upside}(|B/S|) - IV_{downside}(|S/B|)$$

Positive skew indicates call-dominant (bullish) positioning; negative skew indicates put-dominant (protective) positioning. The November 2025 reading of -54pp represents a dramatic repricing of downside risk – the largest single-period skew shift in the sample.

I. Balance Sheet Cost and Replication Friction

I.1 Balance Sheet Cost

Derived from the Fleckenstein and Longstaff (2020) funding basis: the spread between Treasury and derivatives-implied rates that reflects the cost of dealer balance sheet usage. The authors estimate approximately 32 basis points for interest rate derivatives.

Scaled to event-window duration: for a typical event window of 7–14 days, the annualized 32bp basis translates to approximately 0.6–1.2 basis points per event. Rounded to 2–3bp for the balance sheet component.

I.2 Replication Friction

Two-leg spread cost for constructing a binary payoff from vanilla options at institutional execution quality. Each leg incurs bid-ask spread; the total is twice the single-leg cost. For liquid rate options (SR3), the spread is approximately 1 basis point per leg, giving 2bp round-trip. For less liquid option chains, the spread may be 2–5bp per leg.

Strike discreteness contributes additional approximation error (the spread payoff is not exactly binary when ε remains positive), estimated at 1–2bp.

I.3 Combined Estimate

Balance sheet cost ($\sim$3bp) + replication friction ($\sim$4bp) = $\sim$7bp total. This estimate is conservative and applied uniformly across categories. For rate derivatives where dealer balance sheets are optimized for the product, the true cost may be lower. For exotic or illiquid derivative categories, it may be higher.

VRP accounts for approximately 93

J. Size-Dependence Analysis

J.1 Methodology

The Seesaw condition is evaluated at four reference sizes: \$1M, \$3M, \$5M, \$10M. At each size, execution cost is interpolated from the depth curve (§F.3). The structural wedge is size-invariant – VRP does not change with position size. Only the execution cost side varies.

J.2 Results

Table A.8. Threshold Status by Reference Size

Reference Size	PM Wins	Threshold	PM Loses	Total
$1M	36	14	37	87
$3M (canonical)	**30**	**12**	**45**	**87**
$5M	26	10	51	87
$10M	21	7	59	87

Six additional events cross at $1M versus $3M. These are primarily BTC thin contracts and select commodity and election events where the structural wedge comfortably exceeds spread cost but impact at $3M pushes total execution above the threshold. The pattern confirms the Q* size-dependence: the position size at which the Seesaw tips is a function of the structural wedge and the depth of the PM market.

At $10M, twenty-one events still cross – concentrated in BTC liquid contracts (with $17M–$33M in PM volume) and high-volume election markets (Ireland at $115M, Korea at $29M). These are the deep end of the displacement frontier: markets where even large institutional positions execute below the structural wedge.

J.3 Contract-Level Detail (Selected)

Table A.9. Events Crossing at $1M but Not $3M

Event	Cat.	PM Vol.	Cost $1M	Cost $3M	Wedge (%)	$1M	$3M
BTC_120K_NOV_2025	BTC	$2.94M	2.89	10.60	4.36	PM wins	PM loses
BTC_125K_JAN_2026	BTC	$2.20M	3.41	13.83	4.08	Threshold	PM loses
BTC_150K_NOV_2025	BTC	$2.07M	3.78	14.58	4.36	Threshold	PM loses
GAS_JUN_2025	Retail Gas	$2.17M	2.11	6.54	3.77	PM wins	PM loses
AUS_PARL_2025	Elections	$1.70M	5.27	17.57	SVEP 0.474	PM wins	PM loses
NOR_PARL_2025	Elections	$2.00M	4.51	15.07	SVEP 0.762	PM wins	PM loses

The six events share a common profile: PM volume between $1.7M and $3M, structural wedge sufficient at small size, but impact costs scaling nonlinearly with position size such that $3M positions push execution cost above the wedge. These events sit at the nearest edge of the displacement frontier – the Seesaw has tipped at moderate-institutional scale but not at full-institutional scale.

References

Alexander, C. and Imeraj, A. (2023). "The Bitcoin VRP." *Journal of Financial Economics*, forthcoming.

Black, F. (1976). "The Pricing of Commodity Contracts." *Journal of Financial Economics*, 3(1–2), 167–179.

Bollerslev, T., Tauchen, G., and Zhou, H. (2009). "Expected Stock Returns and Variance Risk Premia." *Review of Financial Studies*, 22(11), 4463–4492.

Bollerslev, T., Gibson, M., and Zhou, H. (2011). "Dynamic Estimation of Volatility Risk Premia and Investor Risk Aversion from Option-Implied and Realized Volatilities." *Journal of Econometrics*, 160(1), 235–245.

Bowles, S. (2017). *The Moral Economy: Why Good Incentives Are No Substitute for Good Citizens*. New Haven: Yale University Press.

Bürgi, N., Guo, Y., and Geman, S. (2025). "The Anatomy of Prediction Markets: Evidence from Kalshi." Working paper.

Cantillon, E. and Yin, P.-L. (2007). "Competition Between Exchanges: Lessons from the Battle of the Bund." Working paper, MIT Sloan.

Carr, P. (1998). "Randomization and the American Put." *Review of Financial Studies*, 11(3), 597–626.

Carr, P. and Madan, D. (1998). "Towards a Theory of Volatility Trading." In R. Jarrow (ed.), *Volatility: New Estimation Techniques for Pricing Derivatives*, 29, 417–427. London: Risk Publications.

Carr, P. and Wu, L. (2009). "Variance Risk Premiums." *Review of Financial Studies*, 22(3), 1311–1341.

Demeterfi, K., Derman, E., Kamal, M., and Zou, J. (1999). "A Guide to Volatility and Variance Swaps." *Journal of Derivatives*, 6(4), 9–32.

Drechsler, I. and Yaron, A. (2011). "What's Vol Got to Do with It." *Review of Financial Studies*, 24(1), 1–45.

Duffie, D. (2012). *Dark Markets: Asset Pricing and Information Transmission in Over-the-Counter Markets*. Princeton: Princeton University Press.

Fleckenstein, M. and Longstaff, F.A. (2020). "The US Treasury Floating Rate Note Puzzle." *Journal of Finance*, 75(3), 1381–1424.

Glosten, L.R. (1994). "Is the Electronic Open Limit Order Book Inevitable?" *Journal of Finance*, 49(4), 1127–1161.

Haldane, A.G. (2011). "The Race to Zero." Speech given at the International Economic Association, Beijing, July 8, 2011. Bank of England.

Hayek, F.A. (1945). "The Use of Knowledge in Society." *American Economic Review*, 35(4), 519–530.

Hayek, F.A. (1973). *Law, Legislation and Liberty, Volume 1: Rules and Order*. Chicago: University of Chicago Press.

He, Z. and Krishnamurthy, A. (2013). "Intermediary Asset Pricing." *American Economic Review*, 103(2), 732–770.

Jain, P.K. (2005). "Financial Market Design and the Equity Premium: Electronic versus Floor Trading." *Journal of Finance*, 60(6), 2955–2985.

Klemperer, P. (1987). "Markets with Consumer Switching Costs." *Quarterly Journal of Economics*, 102(2), 375–394.

Klemperer, P. (1995). "Competition when Consumers have Switching Costs: An Overview with Applications to Industrial Organization, Macroeconomics, and International Trade." *Review of Economic Studies*, 62(4), 515–539.

Kyle, A.S. (1985). "Continuous Auctions and Insider Trading." *Econometrica*, 53(6), 1315–1335.

Madhavan, A. (2000). "Market Microstructure: A Survey." *Journal of Financial Markets*, 3(3), 205–258.

Reichenbach, L. and Walther, A. (2025). "Who Profits from Prediction Markets? Evidence from Polymarket." Working paper.

Scott, J.C. (1998). *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. New Haven: Yale University Press.

Zhu, H. (2014). "Do Dark Pools Harm Price Discovery?" *Review of Financial Studies*, 27(3), 747–789.