AURA-PUI

A Real-Time, Market-Implied Index Constructed from Prediction Market Prices

Alex Byerly
Founder, Aurelian
support@aurelianhq.com • Aurelianhq.com
Working Paper | April 2026 | Version 1.0

Disclaimer: The views expressed in this paper are solely those of the author. The author is the founder of Aurelian, which publishes the AURA-PUI.

Abstract

This paper presents the methodology for the AURA Policy Uncertainty Index (AURA-PUI), a high-frequency, daily-published measure of the dispersion and repricing intensity of currently tradable, policy-relevant event probabilities on U.S. prediction markets. Unlike text-based measures that infer uncertainty from newspaper keyword counts (such as the Baker-Bloom-Davis Economic Policy Uncertainty Index), the AURA-PUI derives its signal from the prices, volatility, and microstructure of binary event contracts traded on regulated exchanges and established decentralized platforms. The index is intended as a forward-looking, market-implied complement to existing text- and survey-based uncertainty measures, not a general-purpose proxy for macroeconomic policy uncertainty in the broadest sense. To the author's knowledge, AURA-PUI is the first published, rules-based policy uncertainty index constructed from prediction market microstructure.

The index is constructed from a dynamically managed basket of prediction market contracts spanning five economic policy domains: monetary policy, trade and geopolitical economics, fiscal and regulatory affairs, macroeconomic outcomes, and sector-specific economic policy. Each contract is scored using a two-layer composite framework that captures both the static dispersion of the contract's market-implied probability (via the Probability Dispersion Score, or PDS, and logit-space volatility) and abnormal per-contract price movement, used as a proxy for repricing intensity (via cross-sectionally normalized momentum). Contracts are weighted by observable liquidity metrics: trading volume and executable bid-ask spreads. Once the classification framework, exclusion filters, and parameters are fixed, no day-to-day discretionary overrides or salience reweightings enter the published index.

The methodology employs continuous reconstitution with asymmetric entry/exit buffers, score-proportional per-contract concentration caps, a 50% domain concentration ceiling, and a domain persistence layer that prevents artificial index collapse during contract expiration clusters. Multi-platform sourcing from Kalshi (CFTC-regulated) and Polymarket (decentralized CLOB) provides breadth across the prediction market ecosystem while accommodating platform-specific microstructure differences in spread estimation and volume reporting. An architectural decomposition finds that in the current backtest sample, the concentration caps have negligible effect on the headline series, while persistence contributes modest smoothing; the headline is predominantly determined by the scoring and liquidity-weighted aggregation rather than by the stabilizing rules.

The AURA-PUI is computed at hourly intervals with daily publication, providing a high-frequency, market-implied complement to existing monthly uncertainty measures.

Keywords: Policy uncertainty, prediction markets, event contracts, Probability Dispersion Score, index methodology, high-frequency risk measurement

JEL Codes: G14, G17, D84, E66

1. Introduction

1.1 Motivation

Political and policy risk is among the most consequential yet poorly measured factors in institutional portfolio management. Regulatory shifts, fiscal policy changes, trade actions, and geopolitical events create material uncertainty that affects asset valuations, capital allocation decisions, and risk budgets across every major asset class.
The dominant quantitative measure, the Baker-Bloom-Davis (BBD) Economic Policy Uncertainty Index (Baker, Bloom, and Davis, 2016), hereafter the BBD EPU Index, or EPU Index where the author attribution is unambiguous, has been referenced in thousands of academic papers, Federal Reserve communications, and institutional research reports. Despite its widespread adoption, the EPU Index has well-documented limitations: it is backward-looking (reflecting media coverage after events), subject to media framing effects, available only at monthly frequency, and measures attention rather than genuine economic uncertainty.

The emergence of liquid, regulated prediction markets, particularly following the expansion of CFTC-regulated event contract exchanges and the demonstrated forecasting accuracy of decentralized prediction platforms, creates the opportunity for a fundamentally different approach: measuring policy-event uncertainty directly from the prices at which informed participants are willing to transact. Prediction markets apply the information aggregation properties of traditional financial markets, where participants bear the financial consequences of their probabilistic assessments (Arrow et al., 2008; Wolfers and Zitzewitz, 2004), to discrete policy events that other asset classes price only indirectly. The interpretation of prediction market prices as probabilities requires care (Manski, 2006), but the index does not rely on a strict probabilistic interpretation: PDS and the liquidity-weighted aggregation operate on market-implied prices regardless of whether those prices are well-calibrated probability estimates.

What this index measures, and what it does not. The AURA-PUI measures the dispersion and repricing intensity of currently tradable, policy-relevant event probabilities. It is *not* a measure of the macroeconomic severity of policy actions, nor of attention paid to policy in the broader information environment. When a policy event is realized, contracts tracking *whether the event will occur* move toward resolution, and the index can decline even as the economic shock from the realized policy intensifies. This distinction is a design feature, not a flaw, but it bounds the interpretive claims that can be made about the index. Section 8.4 documents one such episode (the April 2, 2025 tariff announcements) explicitly. The index is described throughout this paper as “forward-looking” in the sense that the underlying contract claims resolve at future dates; this refers to the nature of the instruments, not to the index’s role as a predictor of subsequent uncertainty or market outcomes.

1.2 Contribution

This paper presents the complete methodology for the AURA Policy Uncertainty Index (AURA-PUI). The index transforms raw prediction market pricing into a rules-based measure of market-implied U.S. policy-event uncertainty. The core contributions are:

1. A two-layer composite scoring framework that captures both the static dispersion of contract prices (Layer 1: Probability Dispersion Score and logit-space volatility) and abnormal per-contract price movement as a proxy for repricing intensity (Layer 2: cross-sectionally normalized per-contract absolute momentum).

2. A purely liquidity-weighted aggregation pipeline under which, once the classification framework, exclusion filters, and parameters are fixed, no day-to-day discretionary overrides or salience-based reweightings modify the published index. Contracts are weighted by observable volume and executable bid-ask spreads, with score-proportional concentration caps.

3. A continuous reconstitution mechanism with asymmetric entry/exit buffers that reflects the event-driven nature of prediction market contract lifecycles, enabling the index to respond to major policy shocks within 24 hours.

4. Multi-platform data sourcing from Kalshi (CFTC-regulated) and Polymarket (decentralized CLOB), with explicit handling of platform-specific microstructure differences.

5. A domain persistence layer that prevents artificial index collapse during contract expiration clusters by carrying forward the last known domain contribution with exponential decay.

Existing quantitative measures of policy uncertainty are derived from text (BBD EPU and its extensions; Manela and Moreira, 2017), forecast-error volatility (Jurado, Ludvigson, and Ng, 2015), or option-implied equity volatility (VIX). Academic prediction-market research (Wolfers and Zitzewitz, 2004; Arrow et al., 2008) studies prediction markets as objects of inquiry. The closest published use of prediction-market pricing as a policy-uncertainty signal is Goodell, McGee, and McGroarty (2020), who employ the incumbent party's re-election probability from daily prediction markets to measure election-period policy uncertainty across seven U.S. presidential campaigns. Their work uses a single probability series targeted at one class of policy event; it does not construct a standing, multi-domain, rules-based, aggregated index from prediction market microstructure. To the author's knowledge, the AURA-PUI is the first index occupying this category: a market-implied policy uncertainty measure sourced from prediction market pricing, with formal rules, multi-platform coverage, domain decomposability, and no discretionary overrides.

1.3 Scope and Limitations

The AURA-PUI is intended as a complement to, not a replacement for, existing uncertainty measures. Its primary advantages (high frequency, contemporaneous market-implied signal, money-weighted expectations) come with corresponding limitations: a short history (from early 2024), dependence on prediction market liquidity that is still maturing, and sensitivity to the available contract universe at any point in time.

This paper documents the index methodology and presents a core architectural decomposition (§8.6) that isolates how much of the headline series is attributable to live market signal versus the deterministic stabilizing rules (caps and persistence) layered on top. A companion paper currently in preparation presents the full empirical validation: event-study analysis across policy events from 2024–2026, Granger causality tests against the BBD EPU Index and the VIX, and correlation structure with realized market outcomes.

The companion paper will also include a pre-registered out-of-sample evaluation with the methodology frozen as of the present publication date and evaluated on subsequent data without retuning.

A persistence decay-variant sensitivity analysis is presented in Appendix D. Additional ablations (including platform-composition robustness) and a formal event-study framework with pre-specified event sets, window lengths, sign rules, and randomization-based inference are deferred to that paper.

2. Index Universe and Contract Selection

2.1 Eligible Platforms

The AURA-PUI sources data from prediction market platforms meeting three criteria: (1) the platform is regulated by either the CFTC or an equivalent national authority, or is a decentralized platform with demonstrated daily trading volume exceeding $5 million and a functioning central limit order book; (2) the platform provides programmatic API access to real-time and historical contract pricing, volume, and order book data; and (3) the platform has published resolution rules and a documented dispute resolution mechanism.
As of the publication date, the eligible platform universe comprises Kalshi (a CFTC-regulated centralized exchange providing OHLC prices, bid/ask OHLC, volume, and open interest via candlestick API) and the offshore decentralized Polymarket CLOB (Polygon-based), providing price history via the Gamma and CLOB APIs and live executable spread and midpoint data. A separate CFTC-regulated U.S. venue, Polymarket US, was authorized by the CFTC on November 25, 2025 (operated through the QCX LLC DCM license that Polymarket acquired in July 2025) and began a limited, invite-only rollout for U.S. users in December 2025, initially restricted to sports-outcome contracts. Polymarket US is not currently integrated into the index pipeline; its admission would be treated as a methodology change under §10.2.

The inclusion of both platforms is motivated by complementary coverage: Kalshi provides comprehensive domestic policy and economic indicator contracts with high-quality order book data, while Polymarket contributes contracts in trade, geopolitical, and international policy domains where Kalshi coverage is thinner. Multi-platform sourcing also reduces single-platform dependency risk.

2.2 Domain Definition

The index covers contracts related to U.S. political and policy uncertainty across five thematic domains:

Domain	Scope
Monetary Policy	Federal Reserve decisions, interest rate expectations, central bank leadership
Trade & Geopolitical Economics	International trade policy, sanctions, tariffs, geopolitical events with economic transmission channels
Fiscal & Regulatory	Government operations, fiscal policy, regulatory agency actions, executive orders
Macroeconomic Outcomes	Key economic indicators: CPI/inflation, GDP, recession probability, employment

Sector-Specific Economic	Industry-specific policy: energy, crypto regulation, housing, antitrust

Contracts must have a demonstrable nexus to economic or financial policy outcomes. Pure political horse-race contracts, entertainment markets, sports markets, and word-count novelty contracts (e.g., contracts resolving on how many times a speaker uses a given word in a State of the Union address) are excluded.

Automated exclusion filters. Polymarket's open market creation model requires systematic filtering. The index applies an extensive set of keyword and pattern filters to screen out non-policy contracts. Per-event market caps prevent single-topic bracket floods from dominating the universe.

2.3 Contract Eligibility Criteria

Individual contracts must satisfy all of the following criteria on a rolling basis. The set of contracts that passes all criteria at time t constitutes the *basket* used for index calculation at that snapshot; the term is used throughout this paper to refer to the active constituent set. Persisted domains (§5.3), which contribute to the headline from contracts that have expired or dropped below eligibility thresholds, are not part of the basket; they operate as a separate overlay on top of the basket-derived index value.

Criterion	Threshold	Rationale
Minimum Volume	\$500/day trailing 7-day avg daily USD notional; immediate removal on breach	Ensures meaningful price discovery. Same threshold governs entry and exit; bypasses the 3-day exit buffer (§5.1), which applies only to the PDS floor, time-to-expiration, and price-data availability criteria.
Maximum Spread	5% bid-ask spread, 80% session compliance	Ensures executable pricing; real order book data only
Time to Expiration	≥ 7 calendar days	Excludes contracts in resolution convergence zone
PDS Floor	PDS ≥ 0.10 (price within ~2.5%–97.5%)	Excludes near-certain outcomes
Basket Minimum	≥ 5 contracts total	Publication gate for sufficient diversification

Domain Diversity	≥ 2 domains with ≥ 1 contract each	Prevents single-domain index

Volume Floor Calibration. The volume threshold is the single most consequential eligibility parameter. The \$500/day floor is calibrated against current prediction market depth: thresholds high enough to ensure deep liquidity at maturity (e.g., \$50,000/day) fail the publication gate on a substantial fraction of historical days under present market conditions, while a \$500 floor publishes on more than 99% of days while still excluding contracts with no meaningful price discovery. The threshold is documented as subject to tightening as prediction market depth matures.

2.4 Duplicate and Overlap Control

Prediction market platforms may list overlapping contracts on the same underlying event. The index manages concentration through three mechanisms:

1. Score-proportional per-contract caps (8%–22%) limit the weight of any single contract regardless of its liquidity.

2. Domain concentration cap (50%) limits the total weight of all contracts within a single domain.

3. Metadata deduplication. Polymarket contracts are deduplicated by conditionId to prevent duplicate entries. Per-event market caps limit bracket floods.

The methodology does not employ a pairwise correlation screen for basket exclusion. The cap stack described above achieves the concentration-control objective through weight limits rather than basket exclusion, and is fully implemented and auditable.

3. Per-Contract Scoring Methodology

3.1 Probability Dispersion Score (PDS)

The Probability Dispersion Score quantifies the static dispersion of a contract's market-implied probability distribution. For a binary contract with market-implied probability p:
$PDS(p) = 4p(1 - p)$

The PDS achieves its maximum of 1.0 at p = 0.50 (maximum dispersion) and approaches zero as p approaches 0 or 1 (near-certainty). It is symmetric: PDS(0.30) = PDS(0.70) = 0.84.

For multi-outcome contracts with K mutually exclusive brackets and outcome probabilities p_1, ..., p_k:

$PDS_m = [1 - \Sigma p_k^2] / [1 - 1/K]$

This is the normalized Gini-Simpson diversity index. It reduces to 4p(1−p) when K = 2.

Multi-outcome implementation. Multi-outcome contracts (e.g., "GDP growth in Q3: <1%, 1–2%, 2–3%, >3%") are treated as bracket families sharing a single underlying event. Brackets are grouped by their platform-provided event identifier (Kalshi event ticker or Polymarket conditionId). Bracket probabilities are taken as the platform-reported midpoint prices; if these do not sum to 1.0 (due to bid-ask spreads or missing brackets), they are renormalized to sum to 1.0 before PDS computation. Brackets that individually fail the volume or spread eligibility criteria are excluded from the family; PDS is computed over the remaining K brackets with renormalized probabilities. If fewer than 2 brackets survive eligibility filtering, the family is excluded entirely. Per-event market caps (§2.4) prevent single-topic bracket floods from dominating the basket: the index admits at most a fixed number of brackets per underlying event, with excess brackets excluded by lowest volume.

Relationship to standard statistical quantities. For a Bernoulli random variable with parameter p, 4p(1−p) is exactly four times the variance, and is therefore a bounded, monotone transformation of the Bernoulli standard deviation on the unit interval. Under the natural (logit) parameterization θ = ln[p/(1−p)], the Fisher information about θ is I(θ) = p(1−p), so PDS is proportional to Fisher information in that parameterization. Under the canonical p-parameterization, Fisher information is 1/[p(1−p)], which is unbounded at the endpoints and not the quantity used here. This paper does not rely on any information-theoretic interpretation of PDS; the score is used purely as a bounded, symmetric measure of probability dispersion that admits a natural multi-outcome generalization. Earlier drafts of this methodology used the label "Fisher Information Score" for this quantity; the renaming in this version reflects the dispersion-based, not information-theoretic, role the score plays in the index.

Design choice. PDS is preferred over Shannon entropy H(p) = −p log p − (1−p) log(1−p) because it is algebraically simpler, avoids the 0·log(0) boundary convention, has a clean multi-outcome generalization (Gini-Simpson), and produces identical ordinal rankings for binary contracts.

Eligibility floor. Contracts with PDS below 0.10 are excluded. This corresponds to approximately $p < 0.025$ or $p > 0.975$ in the binary case. The floor excludes near-certain outcomes that contribute negligible dispersion and would otherwise be artificially included by liquidity alone. This PDS floor is the binding constraint on basket membership in the probability dimension; the price clips documented in §7.3 operate upstream (ingestion) and downstream (logit numerical safety) of this gate, and do not themselves determine eligibility.

3.2 Logit-Space Volatility

To estimate the rate of information arrival, raw probabilities are transformed into logit space:
$\lambda(t) = \ln[p(t) / (1 - p(t))]$

The logit transformation maps the bounded probability space (0, 1) to the unbounded real line. In logit space, a movement from 0.50 to 0.55 produces a comparable magnitude to a movement from 0.80 to 0.86, reflecting the similar information content. Logit-space volatility: $\sigma\lambda$ = StdDev[$\Delta\lambda(t)$] over a trailing 14-day window (336 hourly observations). Minimum history: 7 days. Probabilities clipped to [0.001, 0.999] before transformation. Contracts with fewer than 7 days of history (168 hourly observations) receive a volatility estimate equal to the cross-sectional median across the current basket, flagged internally as data-imputed. The 7-day minimum is a reliability gate: below it, $\sigma\lambda$ is imputed from the cross-section rather than computed from the contract's own history, since standard-deviation estimation on shorter windows is unstable.

3.3 Bid-Ask Bounce Adjustment

Following the intuition of Roll (1984), a microstructure-motivated attenuation is applied to remove the mechanical bid-ask bounce component of observed volatility:
$\sigma_{adj} = \sqrt{\max[0, \sigma^2 - \gamma \cdot s^2]}, \gamma = 0.25$

Roll's classical result links the effective spread to the negative serial covariance of price changes under simplifying assumptions. The form used here is a heuristic, fixed-attenuation adjustment inspired by that intuition rather than a direct implementation of Roll's estimator: γ is held constant at 0.25 across platforms and contracts rather than estimated per-contract from observed return autocovariance. This choice trades estimation noise (per-contract serial-covariance estimation is unstable on the short windows available for many contracts in the universe) for a small, consistent attenuation of bid-ask induced variance.

Dimensional approximation. The formula above applies σ in logit space and s in probability space. A dimensionally exact correction would first transform the spread into logit space via $s_{logit} \approx s / [p(1 - p)]$, which can be substantially larger than s (e.g., at $p = 0.50$, $s_{logit} = 4s$; at $p = 0.90$, $s_{logit} \approx 11s$). The current form therefore *understates* the bid-ask bounce correction, making the adjustment conservative: more bid-ask variance remains in σ_{adj} than would survive a logit-space-consistent correction. Given the maximum eligible spread of $s = 0.05$ (§2.3) and $\gamma = 0.25$, the subtracted term $\gamma \cdot s^2$ is at most 0.000625, a small fraction of typical logit-space variance values; for most Polymarket contracts the adjustment is skipped entirely due to the absence of historical BBO data. The adjustment is retained as a conservative noise floor, and the downstream CDF normalization (§3.4) absorbs cross-sectional level differences in volatility regardless.

Platform-specific application. Kalshi provides real bid/ask OHLC data from its matching engine; the adjustment is always applied. Polymarket's historical API returns midpoint prices only. The adjustment is skipped for synthetic-spread observations. Live BBO data is captured on each daily refresh and cached, progressively enabling the correction.

3.4 CDF Normalization

The adjusted volatility is mapped to [0, 1] using the CDF of the log-normal distribution fitted to the trailing 60-day cross-sectional volatility pool:
$\hat{V} = \Phi[\ln(\sigma_{adj}); \mu_{60}, \sigma_{60}]$

This normalization ensures the volatility score reflects relative positioning within the current market environment. Default value: 0.50 when insufficient data.

3.5 Layer 1: Static Dispersion

$L_1(i) = \alpha \cdot PDS(i) + (1 - \alpha) \cdot \hat{V}(i), \alpha = 0.70$
Layer 1 captures the static dispersion of the contract's market-implied probability (PDS) and the dynamic rate of belief revision (normalized volatility). The $\alpha = 0.70$ weighting emphasizes the dispersion component. Concretely: a maximally dispersed contract at $p = 0.50$ with no recent volatility ($PDS = 1.0$, $\hat{V} = 0$) scores $L_1 = 0.70$, while a near-resolved contract at $p = 0.90$ with maximum normalized volatility ($PDS = 0.36$, $\hat{V} = 1.0$) scores $L_1 = 0.55$. The design principle is that high static dispersion dominates high dynamic volatility in the Layer 1 score, because a contract near $p = 0.50$ represents genuinely unresolved uncertainty regardless of recent price stability.

3.6 Layer 2: Abnormal Repricing Intensity

Layer 2 captures whether individual contracts are experiencing abnormal price movement relative to the historical cross-section. It is a measure of *repricing intensity* (the magnitude of recent probability change relative to typical cross-sectional behavior) and provides per-contract differentiation that Layer 1 alone cannot.
Repricing intensity is not synonymous with uncertainty. Large price moves can reflect new uncertainty arrival, uncertainty resolution, one-sided information flow, liquidity shocks, event-proximity effects, or migration of trading activity between related contracts. Layer 2 treats all of these symmetrically as evidence that the contract is an active locus of policy-relevant price discovery, and weights it accordingly within the composite. The interpretation should remain "this contract is moving more than peers" rather than "this contract is becoming more uncertain."

Computation pipeline:

1. Per-contract momentum. Compute $|\Delta p(7d)|$ for each eligible contract: the absolute 7-day price change in probability space.

2. Cross-sectional normalization. At each snapshot, the median of non-zero $|\Delta p|$ values across the basket is computed (one scalar per snapshot). Using the median of non-zero moves makes the normalizer robust to both basket sparsity (contracts with no trades produce zero moves that would deflate the mean) and isolated large repricings that would inflate it; the EWM variance then tracks how the typical cross-sectional repricing level changes over time. This cross-sectional median is fed to an EWM normalizer that independently tracks two running state variables via West-style online updates: a running mean (μ_{ewm}) and a running variance (σ^2_{ewm}) of the sequence of daily medians. The resulting $\sigma_{ewm} = \sqrt{\sigma^2_{ewm}}$ captures how much the typical cross-sectional repricing level varies over time. Warmup: the first 3 daily snapshots return a neutral score of 0.50 (insufficient data); daily snapshots 4–90 use expanding-window mean and standard deviation; at daily snapshot 91 the EWM initializes from warmup-period statistics and applies the 60-day halflife thereafter.

3. Per-contract z-scoring. Each contract's $|\Delta p|$ is z-scored against the normalizer: $z = (|\Delta p(i)| - \mu_{ewm}) / \sigma_{ewm}$. This compares each contract's absolute price movement to the EWM-smoothed central tendency of the cross-section, normalized by the historical variability of that central tendency. Contracts with $|\Delta p|$ well above the running median receive positive z-scores; those near or below the median receive $z \approx 0$ or negative z-scores.

Why temporal normalization rather than same-snapshot z-scoring. An alternative design would z-score each contract against the cross-sectional mean and standard deviation computed at the same snapshot. The EWM-based temporal approach is preferred for two reasons. First, the basket size varies substantially (4 to 46 contracts); a same-snapshot z-score computed on a basket of 6 contracts is statistically unstable, while the EWM accumulates information across hundreds of prior snapshots. Second, the temporal approach captures *regime-relative* abnormality: a contract moving 5 cents in a quiet regime receives a higher z-score than the same move in a volatile regime, because the EWM variance reflects the historical norm. A same-snapshot z-score would miss this regime conditioning entirely. Strictly, z is a ratio of a contract-level move to the EWM-estimated historical variability of the cross-sectional median, not a distributionally valid z-score; it is used as a regime-relative intensity measure, and the downstream sigmoid (Step 4) maps it to [0, 1] without requiring a distributional interpretation.

4. Sigmoid mapping. $z \to [0, 1]$ via $L_2(i) = 1 / (1 + \exp(-2.0 \cdot z))$. At $z = 0$: 0.50; at $z = +2$: ~0.98; at $z = -2$: ~0.02.

5. Output smoothing. Per-contract EWM with 2-day halflife prevents single-observation whipsaw. Contracts with fewer than 7 days of price history receive the cross-sectional median L2 value at each snapshot, flagged internally as data-imputed.

Probability space vs. logit space. Layer 2 momentum is computed in raw probability space, while Layer 1 volatility operates in logit space. This is a deliberate design choice. Layer 1 measures belief revision (logit-space equalizes information content). Layer 2 captures capital reallocation magnitude (a 10-cent move represents roughly the same notional reallocation regardless of the starting probability). For a liquidity-weighted index, the capital-flow interpretation is the relevant signal.

3.7 Composite Score

$CS(i) = 0.50 \cdot L_1(i) + 0.50 \cdot L_2(i)$

Expanding the Layer 1 definition, the effective decomposition is:

$CS(i) = 0.35 \cdot PDS(i) + 0.15 \cdot \hat{V}(i) + 0.50 \cdot L_2(i)$

Design history and justification. The original architecture of the v3 calculation engine (see Appendix C) included a Layer 3 (cross-domain contagion) that produced a constant 0.50 value with zero differentiation. It was removed, and its 10% weight redistributed equally to Layers 1 and 2. The 50/50 weight is therefore a historical artifact of the Layer 3 removal. To assess whether this matters, a sensitivity sweep evaluated six weight configurations. The published AURA-PUI is the 50/50 configuration; correlations and mean absolute deviations in the table below are computed relative to this baseline, which is why the 50/50 row reports reference values rather than comparative statistics.

L1/L2 Split	Level Correlation	Change Correlation	Mean Abs. Deviation
40/60	> 0.99	> 0.97	~2.3 pts
50/50	**(baseline)**	**(baseline)**	**(reference)**
60/40	> 0.99	> 0.97	~2.3 pts
70/30	0.98	0.94	~4.5 pts
100/0 (L1 only)	0.92	0.71	~12 pts
30/70 (L2 heavy)	0.97	0.88	~7 pts

The key findings: (1) adjacent splits (±10 percentage points) produce level correlations exceeding 0.99 (the index is not fragile); (2) extreme splits diverge substantially, confirming L1 and L2 capture genuinely different signals; (3) the 50/50 split is the neutral point within a stable plateau. Any split between 40/60 and 60/40 would produce a materially equivalent index.

4. Weighting and Aggregation

4.1 Liquidity Weights

The AURA-PUI weights contracts using only observable market microstructure data. For contracts with real order book spread data:
$LW(i) = 0.60 \cdot V_{norm}(i) + 0.40 \cdot S^{-1}_{norm}(i)$

Definitions of normalized inputs. Both $V_{norm}(i)$ and $S^{-1}_{norm}(i)$ are share-of-total normalizations computed across the active basket at each snapshot. Specifically, $V_{norm}(i) = V(i) / \Sigma V(j)$, the share of contract i's trailing 7-day average daily USD notional in the basket total; and $S^{-1}_{norm}(i) = S^{-1}(i) / \Sigma S^{-1}(j)$, the share of contract i's inverse trailing 7-day average bid-ask spread in the basket total. Share-of-total normalization ensures weights sum to 1.0 before cap application and preserves the proportional interpretation: a contract with twice the volume of another receives twice the volume component.

Volume activation gateway. The spread component activates only when a contract's volume exceeds 50% of the cross-sectional median. Below this threshold: $LW(i) = V_{norm}(i)$. For contracts with synthetic spread data (Polymarket historical observations without real BBO), weighting is volume-only.

4.2 Entry Blending

New contracts ramp from score-proportional to liquidity weighting over 7 days:
$w(i) = b \cdot LW(i) + (1 - b) \cdot CS_{norm}(i), b = min(1.0, age/7)$

4.3 Expiration Ramp-Down

Contracts within 7 calendar days of resolution have their weight linearly reduced:
$\omega(i) = min(1.0, hours\_remaining / 168)$

4.4 Per-Contract Concentration Caps

$cap(i) = 0.08 + (0.22 - 0.08) \cdot CS_{percentile}(i)$
where $CS_{percentile}(i)$ is the rank percentile of contract i's composite score CS(i) within the current basket, mapped to [0, 1] (0 = lowest CS in the basket, 1 = highest). Ties receive averaged ranks.

CS Percentile	Cap
0th (lowest)	8%
50th (median)	15%
100th (highest)	22%

High-uncertainty contracts receive more generous caps; low-uncertainty contracts are constrained. Excess weight is redistributed proportionally. The algorithm iterates up to 10 rounds until stable. If convergence is not reached within 10 rounds (extremely rare; no occurrence in the backtest), weights are hard-clipped to their per-contract caps and any residual mass is distributed pro-rata across uncapped contracts in a single final pass. Additionally, within each domain's sub-index computation, no single contract may exceed 40% of the domain's sub-index weight (INTRA_DOMAIN_CONTRACT_CAP in Appendix A).

4.5 Domain Concentration Cap

No single domain may exceed 50% of total headline weight. A post-cap compliance check re-runs if any domain exceeds 52%.

4.6 Headline Index Value

$AURA\text{-}PUI(t) = S \cdot \Sigma\, W(i) \cdot CS(i, t)$

Calibration. S was calibrated on an earlier snapshot of the backtest so the mean weighted composite score maps to approximately 100. Subsequent updates (primarily the progressive accumulation of Polymarket BBO cache data replacing synthetic spread estimates, and minor contract-classification refinements) have modestly shifted the realized mean, yielding a mean index level of 98.35 over the final backtest sample (see §8.2). S is permanently fixed at its initial calibration value and is not subject to periodic recalibration (see §10.2); the index level should be read in relative terms (changes from a baseline of approximately 100), not as a measurement on an absolute uncertainty scale. An alternative inception-anchored normalization (rebasing the index to 100 at the first valid publication date) would produce an identical change-based time series and is internally documented but not used in the published series.

5. Reconstitution, Lifecycle Management, and Domain Persistence

5.1 Continuous Reconstitution

The AURA-PUI employs continuous reconstitution with asymmetric entry/exit buffers:

- **Entry buffer:** 1 calendar day (24 hourly observations). The short buffer ensures response to major policy shocks within one day.
- **Exit buffer:** 3 calendar days (72 hourly observations). Prevents flapping during single quiet sessions. The 3-day exit buffer applies to the PDS floor, time-to-expiration, and price-data availability criteria. Volume is gated by a same-day hard floor at the eligibility threshold (§2.3) and does not transit the exit buffer.
- **Hard floor override:** Immediate removal if trailing 7-day volume falls below $500/day.
- **First-day seeding:** On initialization, all eligible contracts enter immediately.

5.2 Contract Lifecycle

Prediction market contracts have finite lifespans. The index handles lifecycle through ramp-down (7-day linear weight reduction), successor transition (standard eligibility pipeline during overlap), and domain persistence (smooth fade for entire-domain gaps).

5.3 Domain Persistence Layer

Contract expiration clusters can temporarily eliminate entire domain coverage. The persistence layer activates when a domain drops below 2 live constituents (with ≥ 7 days of prior history):

$decay(t) = exp(-ln(2) \cdot h_{dark} / 720)$

where h_{dark} is hours since the domain dropped below 2 live contracts. Halflife = 30 days (720 hours).

What decays. Both the persisted *weight* and the persisted *sub-index value* decay independently with the same 720-hour halflife. The domain's contribution to the headline at time t is therefore:

$C(t) = [w_0 \cdot decay(t)] \times [V_0 \cdot decay(t)] = w_0 \cdot V_0 \cdot decay(t)^2$

where w_0 is the floor weight (10% of pre-depletion weight), V_0 is the sub-index value at the moment of depletion, and decay(t) = exp(−ln(2) · h_{dark} / 720). Because the contribution is the product of two independently decaying terms, the effective halflife of the *contribution* is approximately 360 hours (15 days), not 720. This means persisted domains fade from the headline roughly twice as fast as the component halflife alone would suggest: at h_{dark} = 720 (30 days), the contribution has fallen to 25% of its initial value (0.5 × 0.5), not 50%. This aggressive fade is consistent with the design intent of preventing stale values from persisting indefinitely, and the dropout threshold (contribution weight below 0.5%) is reached correspondingly sooner.

The floor weight is the initial persisted weight assigned to a newly depleted domain at the moment it falls below two live contracts, set to 10% of its pre-depletion weight (PERSISTENCE_FLOOR_WEIGHT in Appendix A). At h_{dark} = 0, the domain contributes its pre-depletion value at 10% weight; at h_{dark} = 360, the contribution has halved; at h_{dark} = 720, the contribution has fallen to one-quarter of its initial level; below the 0.5% dropout threshold, the domain exits persistence entirely.

Safeguards. Maximum total persisted weight: 30% across all persisted domains combined, enforced before the final weight renormalization step of the per-snapshot pipeline. Because live weight share and persisted weight share are constructed to sum to 1.0 (live_share = 1.0 − total_persisted_weight), renormalization is effectively a no-op and the 30% cap binds on the published weights as well as on the pre-normalization weights. Dropout: persisted contribution exits when weight falls below 0.5%. Minimum domain history: 7 days of valid sub-index observations prior to depletion are required for a domain to qualify for persistence (otherwise the domain is simply absent until new live constituents enter).

Conceptual rationale. The persistence layer exists because prediction market contract expiration is an operational event, not an information event. When all contracts in a domain expire simultaneously, as frequently happens around FOMC dates or quarterly indicator releases, the underlying policy uncertainty does not vanish; it simply becomes temporarily unobservable through the prediction market channel. A hard drop-to-zero would attribute the expiration cluster to a genuine uncertainty resolution, producing a mechanical headline decline unrelated to any change in the policy environment. Exponential decay is preferred over a flat carry-forward because stale values should lose influence over time: the longer a domain goes without live contracts, the less confidence the index places in the last-known sub-index value. The double-decay design (effective contribution halflife ≈ 360 hours) implements an aggressive fade that balances continuity against staleness: persisted domains lose three-quarters of their contribution within 30 days, and exit persistence entirely shortly thereafter. The 30% total persisted weight cap provides an additional safeguard: even during extended coverage gaps, the headline remains majority-determined by live market data.

Methodological caveat. The persistence layer described above, and the per-contract and domain concentration caps specified in §4.4 and §4.5, are deterministic stabilizing rules layered on top of live market inputs. In the backtest sample documented in §8, the domain concentration cap binds on 85.4% of valid days and the persistence layer is active on 51.9% of valid days. The headline index path therefore reflects both market pricing and the index architecture imposed on a sparse, expiring contract universe. This is a deliberate trade-off (without these safeguards, contract expiration clusters and short-horizon concentration shifts would produce mechanical headline movements unrelated to underlying uncertainty conditions), but readers should interpret all empirical results in the remainder of the paper with this operational structure in mind. Persisted domains do not count toward the MIN_DOMAINS_FOR_PUBLICATION = 2 gate, and persisted contracts do not count toward MIN_BASKET_SIZE = 5 (§7.2); publication always requires a minimal live-market core, with persistence acting as a smoothing overlay rather than a substitute for live data. Section 8.6 provides a core architectural decomposition that quantifies the contribution of each rule to the headline series; the companion paper described in §1.3 will report extended ablation analyses.

5.4 Reconstitution Attribution

Each daily index change is decomposed into three additive components. Let C denote contracts present in both today's and yesterday's basket (continuing), E denote contracts entering today, and X denote contracts exiting today.
$\Delta I(t) = \text{Signal} + \text{Weight Evolution} + \text{Composition}$

where:

$\text{Signal} = S \cdot \Sigma_{i} \in C\ w_i(t-1) \cdot \Delta CS_i(t)$

$\text{Weight Evolution} = S \cdot \Sigma_{i} \in C\ \Delta w_i(t) \cdot CS_i(t)$

$\text{Composition} = S \cdot [\Sigma_{i} \in E\ w_i(t) \cdot CS_i(t) - \Sigma_{i} \in X\ w_i(t-1) \cdot CS_i(t-1)]$

Signal isolates pure market repricing of continuing contracts at prior-day weights. Weight evolution captures the effect of weight changes (from blend-age ramp, volume/spread shifts, and cap redistribution) applied to current-day scores. Composition captures the net effect of basket turnover. The decomposition is exact: Signal + Weight Evolution + Composition = $I(t) - I(t-1)$ by construction, with no residual term, because the weight-evolution term uses current-day scores (an end-of-period convention that absorbs the cross-term $\Delta w \cdot \Delta CS$ into weight evolution rather than creating a separate interaction component).

6. Data Sources and Microstructure

6.1 Kalshi

Candlestick endpoints provide per-period OHLC prices, yes_{bid} and yes_{ask} OHLC (real order book snapshots from the matching engine), volume, and open interest. The bid/ask fields are real order book snapshots, not synthetic. When no trades occur, bid/ask remain populated.

6.2 Polymarket

The Polymarket data ingested by the index is sourced from the offshore decentralized CLOB (Polygon-based, geoblocked for U.S. retail), not from the CFTC-regulated Polymarket US venue referenced in §2.1. The microstructure description below refers throughout to the offshore venue.
Market discovery is automated via the Gamma API with policy-relevant tag filtering, conditionId deduplication, volume gating, and an extensive set of keyword and pattern exclusion filters.
Price history: CLOB API returns midpoint prices only. No historical bid/ask available.

Live BBO capture: On each daily refresh, midpoint and spread endpoints are queried for each active token. Snapshots are cached to disk, progressively replacing synthetic estimates with real data.

6.3 Volume Computation

Kalshi: USD notional = shares traded × yes-price, where each contract has $1 notional per share and yes-price $\in [0, 1]$ is the market-implied probability expressed in dollars. Polymarket: total USD notional from Gamma API. Eligibility uses trailing 7-day average daily volume, aggregated from hourly candles.

7. Publication and Quality Controls

7.1 Calculation Schedule

Hourly intervals; daily publication uses the most recent hourly snapshot. ~17,000 snapshots per year at hourly granularity. Timestamps from both platforms are normalized to UTC at ingestion; the daily publication snapshot is taken at 23:00 UTC.

7.2 Quality Flags

Flag	Condition	Effect
INSUFFICIENT_CONSTITUENTS	< 5 eligible contracts	Index not published
INSUFFICIENT_DOMAINS	< 2 active domains	Index not published
REDUCED_CONFIDENCE	< 10 contracts	Published with advisory flag
SYNTHETIC_SPREADS	> 50% synthetic spread data	Published with advisory flag
DOMAIN_PERSISTENCE	Carry-forward values active	Published with advisory flag

7.3 Staleness Controls

72-hour flag for review. 96-hour exclusion. Forward-fill within staleness window. Raw prices from both platforms are clipped to [0.01, 0.99] on ingestion as a data-quality step matching exchange tick-size conventions. This is not the eligibility gate: the binding constraint on basket membership is the PDS floor (§2.3, §3.1), which excludes contracts with market-implied probability outside approximately [0.025, 0.975]. A separate, wider clip to [0.001, 0.999] is applied inside the logit transformation (§3.2) as a numerical safeguard against ±∞; because upstream stages already constrain inputs to a tighter band, this clip is not expected to bind on data that reaches the aggregation pipeline.

8. Backtest Results and Initial Empirical Evidence

This section presents the AURA-PUI backtest results and initial empirical evidence regarding the index's behavior. The objective is to assess plausibility, internal coherence, and descriptive distinctness from related measures, not to establish definitive econometric validation. The index is designed to measure market-implied policy-event uncertainty, not to forecast asset returns; the empirical framework therefore emphasizes construct plausibility (does the index behave as a measure of policy-event uncertainty should?) over predictive power for financial assets. Although no ex-post parameter tuning was performed against the metrics reported here, the parameter space was developed in the same historical environment used for descriptive evaluation; this limits the independence of the empirical results from the construction process. Stronger validation, including pre-registered out-of-sample evaluation and ablation analyses, is deferred to the companion paper described in §1.3.

8.1 Sample Period and Data

The backtest covers January 4, 2024 to April 13, 2026, comprising 815 daily observations with valid index values. Data is sourced from 582 unique contracts across Kalshi (231) and Polymarket (351) drawn from a curated universe of classified policy-relevant event contracts. The backtest uses the identical calculation engine, parameters, and eligibility criteria documented in this paper; no ex-post parameter tuning was performed.
At backtest inception on January 4, 2024, all contracts meeting eligibility criteria as of that date entered the basket under the first-day seeding rule (§5.1); the 1-day entry buffer applies only to contracts becoming eligible after inception. Early-sample values therefore do not reflect steady-state reconstitution dynamics and should be interpreted as transitional.

Retrospective data enrichment. The backtest is computed using the final-state contract classification dictionary and the final-state Polymarket BBO cache as of April 13, 2026; it therefore reflects retrospective data enrichment (contracts originally scored with synthetic spreads were later re-scored with real BBO data as the cache accumulated) rather than what the index would have published in real time. The companion paper's pre-registered out-of-sample evaluation, which freezes all inputs as of the methodology publication date, is not subject to this concern.

8.2 Summary Statistics

Statistic	Value
Sample period	January 4, 2024 – April 13, 2026
Daily observations (valid index)	815
Index level	
Mean	98.35

Median	99.48
Minimum	53.01 (March 23, 2024)
Maximum	143.24 (May 29, 2025)
Standard deviation	20.02
1-day autocorrelation	0.967
Constituent count	
Mean / Median	22.9 / 23
Min / Max	4 / 46
Domain coverage	
Mean active domains	3.4
Days with all 5 domains	8.0% (65 days)
Days with 4+ domains	58.7% (478 days)
Domain persistence active	51.9% (423 days)
Quality	
REDUCED_CONFIDENCE flag	8.2% (67 days)
Domain cap binding	85.4% (696 days)
Constituent–level correlation	$r = 0.12$ ($r^2 = 0.015$)

Note on minimum constituent count. The reported minimum of 4 constituents corresponds to a single day (January 15, 2024) during the early backtest period. That observation was flagged INSUFFICIENT_CONSTITUENTS in the quality control pipeline and would not have been externally published under the production publication gate (MIN_BASKET_SIZE = 5, §7.2); it appears in the summary statistics because the backtest computes index values for analytical continuity even on flagged days. On all subsequent days, the constituent count is ≥ 5. Excluding this single flagged observation does not materially affect any reported summary statistic

8.3 Cross-Domain Independence

The five domain sub-indices exhibit low pairwise correlations, consistent with their capturing distinct dimensions of policy uncertainty:

	Monetary	**Trade**	**Fiscal**	**Macro**	**Sector**
Monetary	1.000	0.198	0.257	0.249	0.051
Trade	0.198	1.000	−0.141	−0.205	0.303
Fiscal	0.257	−0.141	1.000	0.285	−0.015

Macro	0.249	−0.205	0.285	1.000	−0.098
Sector	0.051	0.303	−0.015	−0.098	1.000

All pairwise correlations are below 0.31 in absolute value. Correlations are computed using pairwise deletion: each pair uses only the days on which both sub-indices have at least one live contract (i.e., are not both in persistence). Pairwise sample sizes vary substantially across pairs, reflecting the uneven domain coverage documented in §8.5; pairs involving the Monetary and Fiscal domains have the most observations, while pairs involving Macro or Sector have the fewest. The persistence layer introduces autocorrelated smoothing into persisted sub-indices; correlations involving frequently persisted domains (Macro, Sector) should be interpreted with this in mind. The negative correlations (Trade–Fiscal at −0.14, Trade–Macro at −0.21) are small but directionally consistent with potential substitution in attention or capital across domains; with this sample size they should not be read as identifying such an effect. The weak correlation between constituent count and headline level ($r = 0.12$) is consistent with the index not being mechanically driven by basket size. Readers should note, however, that the domain concentration cap binds on 85.4% of valid days and the persistence layer is active on 51.9% of valid days; the headline series therefore reflects a deterministic transformation of underlying market data rather than a raw aggregation, and findings throughout this section should be interpreted with that operational structure in mind.

8.4 Event Response Analysis

The AURA-PUI was evaluated against 19 scoreable policy events across monetary, fiscal, and composite categories. All event scores in this section are assigned using the domain sub-index corresponding to the event type, under a uniform rule applied across the full sample. Domain sub-indices are used rather than the headline composite because they target the policy-specific signal more directly than the headline, which blends across all domains. Directional accuracy, counting any move in the expected direction, was 63% (12/19), comprising 8 strong passes (greater than 3 point move in the expected direction), 4 weak passes (any move in the expected direction below the 3-point threshold), and 7 fails. Restricted to strong passes only, the figure is 42% (8/19). For reference, 12/19 under a binomial null of 50% gives $p \approx 0.18$, but this should not be interpreted as a hypothesis test: the events were hand-selected, sign assignments reflect researcher judgment, and the T+3 window was chosen descriptively. No formal test of directional accuracy is appropriate given these degrees of freedom; the binomial p-value is reported solely as a statistical anchor for readers assessing the descriptive result. An additional 7 macroeconomic events were excluded from scoring due to thin constituent coverage (1–2 contracts) in that domain; these are listed below. The exercise reported here is descriptive rather than inferential: event selection, sign assignment, window choice, and the absence of a formal null distribution all preclude treating these results as a statistical test. A pre-specified event-study framework with placebo events and randomization-based inference is deferred to the companion paper described in §1.3.

Selected event responses (5 of 19 scored events, chosen for illustrative diversity across event types; this is not a pass-rate summary). The full 12/19 directional accuracy cited above is computed across all 19 scored events, which include both passes and fails not shown here. The table reports headline composite values for comparability across events; directional scoring uses the relevant domain sub-index as stated above, and for all events shown, the sub-index moved in the same direction as the headline with comparable or larger magnitude.

Event	Date	Headline T Value	Headline T+3 Change	Domain Sub-Index Result
Election Day (compress)	Nov 5, 2024	88.33	−6.08	Pass
Election result (spike)	Nov 6, 2024	84.26	+8.27	Pass
First tariff announcements	Feb 1, 2025	105.35	+14.11	Pass
Liberation Day tariffs	Apr 2, 2025	127.47	−17.01	See below
Government shutdown / CR	Mar 14, 2025	102.03	+13.28	Pass

Macroeconomic domain coverage limitations. The macroeconomic outcomes domain is structurally thin throughout the backtest period, typically carrying only 1–2 eligible contracts (see §8.5 and §11). This materially limits the event-response analysis: seven macroeconomic events (four GDP advance releases, two jobs reports, and one CPI release between February 2025 and February 2026) could not be scored because the domain carried fewer than 3 eligible contracts at event time. The exclusion reflects a universe limitation, not an analytical choice: the index is not designed to produce reliable domain-level signals with fewer than 3 eligible contracts.

Uncertainty vs. Severity: the Liberation Day response. The index declined 28 points over the week following the April 2, 2025 tariff announcement, a counterintuitive result if one equates policy uncertainty with policy severity. The explanation is structural: the AURA-PUI measures uncertainty about whether policy actions will occur, not their economic consequences once enacted. When tariffs were announced, contracts tracking "will tariffs happen" moved toward resolution (PDS declining as probabilities approached 0 or 1), while contracts tracking downstream consequences (recession probability, Fed rate path) absorbed the uncertainty transfer. The recession-in-2025 contract moved from $p = 0.42$ to $p = 0.60$ over T+3, and two new tariff-specific contracts entered the basket. The net effect was a decline in aggregate PDS that outweighed the new-contract additions. The fact that uncertainty resolves even when severity increases is a structural consequence of the PDS-based scoring methodology. Users interpreting short-horizon index movements should account for resolution dynamics, decomposable via the three-way attribution framework (§5.4).

8.5 Constituent Count and Domain Coverage

The constituent count ranges from 4 to 46, with a mean of 23 contracts. Key observations:

1. Expiration clusters. Periodic drawdowns occur around FOMC dates and quarterly indicators. The domain persistence layer (§5.3) activates on 51.9% of valid days.

2. Platform contribution. Polymarket contracts contribute primarily to trade and geopolitical coverage. Multi-platform sourcing expanded peak constituent count from approximately 25 (Kalshi-only) to over 46.

3. Domain breadth. Full five-domain coverage on 8.0% of days. Four or more domains on 58.7% of days. The Fiscal & Regulatory domain maintains the most consistent coverage; the Macroeconomic Outcomes and Sector-Specific Economic domains are the most frequently persisted.

4. Macroeconomic coverage gap. The macroeconomic outcomes domain frequently operates with 1–2 eligible contracts, insufficient for reliable sub-index scoring. This reflects limited prediction market depth in GDP, CPI, and employment bracket contracts and is a known coverage limitation, not a methodology failure. The domain is expected to deepen as prediction market volumes mature.

8.6 Architectural Decomposition: How Much of the Headline Is Architecture?

The combination of an 85.4% domain-cap binding rate and a 51.9% persistence-active rate raises a natural question: how much of the headline series is attributable to live market signal, and how much to the deterministic stabilizing rules layered on top? This section addresses that question directly by computing the index under four nested architectural configurations on the same backtest sample (January 2024 – April 2026, 815 daily observations):

1. **Live-only.** Raw aggregation with no caps, no persistence, no smoothing overlays. Per-contract liquidity weights are renormalized to sum to 1.0 across the eligible basket each day.
2. **Live + per-contract caps.** Adds the score-proportional per-contract concentration caps (§4.4) but no domain cap and no persistence.
3. **Live + caps (per-contract and domain).** Adds the 50% domain concentration cap (§4.5).
4. **Headline (full architecture).** Adds the domain persistence layer (§5.3). This is the published AURA-PUI.

The four configurations are computed using identical eligibility filtering, scoring, liquidity weights, entry blending, and ramp-down logic; only the cap and persistence stages differ.

Central finding: in this sample, the concentration caps have negligible effect on the headline. Configurations 1 through 3 produce identical index series to floating-point precision. The per-contract caps and domain concentration cap redistribute weight across contracts, but because contracts within and across domains have similar composite scores, the weighted sum $\Sigma(w \times CS)$ is insensitive to that redistribution. Only persistence produces a measurably different series. The table below therefore reports two columns: the uncapped live-only series (representing all three cap configurations identically) and the published headline with persistence.

Statistic	Live-only (= +Caps)	Headline
Mean index level	99.75	98.35
Standard deviation	21.09	20.02
1-day autocorrelation	0.966	0.967
Correlation with headline	0.997	1.000
Correlation with BBD EPU (monthly levels)	0.67	0.73

Notes. Live-only = configurations 1–3 (identical). Headline reproduces the published series (§8.2). BBD EPU correlation computed over 25 months of overlapping monthly data (February 2024 – February 2026). Minor boundary-condition differences between the ablation-pipeline recomputation and the canonical published series (~0.06 points on the mean) are documented in Appendix D.

Interpretation.

(i) In this sample, caps do not reshape the index. The 85.4% domain-cap binding rate reported in §8.2 is a weight-allocation statistic, not an index-impact statistic. The cap *binds* (meaning it constrains the maximum weight a single domain can receive), but this constraint does not alter the headline value because the contracts whose weights are being redistributed have comparable composite scores to the contracts receiving the redistributed weight. This claim is empirically verifiable: the mean cross-sectional standard deviation of CS(i) across the full basket is 0.212 on days when the domain cap binds versus 0.213 on non-binding days (Welch t, $p = 0.69$), statistically indistinguishable. The weighted dot product is stable under permutations among similarly-scored contracts. This is an empirical property of the current contract universe, not a mathematical inevitability: in a future environment where domain-level composite scores diverge more sharply, the caps would become active. They exist as structural safeguards against concentration scenarios that have not yet materialized.

(ii) Persistence is the only active architectural component. On persistence-active days (51.9% of the sample, per §8.2), the mean total persisted floor weight is 7.6% of the headline (median 7.3%, maximum 19.6%); the 30% cap is never reached, because persistence typically activates for 1–2 domains at a time with moderate decay. Adding persistence lowers the mean by 1.4 points (99.75 → 98.35) and the standard deviation by 1.1 points (21.09 → 20.02). It slightly raises the 1-day autocorrelation (0.966 → 0.967). These effects are modest: the live-only and headline series correlate at $r = 0.997$. Persistence's primary role in the current sample is to smooth domain transitions during expiration clusters, not to fundamentally reshape the signal. The BBD EPU correlation is marginally higher for the headline (0.73 vs. 0.67), suggesting persistence modestly improves alignment with the monthly newspaper-based measure, likely because the smoothing effect of carrying forward depleted-domain contributions reduces high-frequency noise that does not correspond to monthly EPU movements. A sensitivity analysis across four persistence decay-variant configurations (Appendix D) confirms that the headline is insensitive to the specific decay specification: excluding one degenerate variant (value-only

decay, which lacks weight disengagement), all configurations produce pairwise correlations above 0.998 and differ by at most 1 point in mean level.

(iii) In this sample, the headline is predominantly signal-driven. In the present backtest, the headline AURA-PUI is predominantly determined by the scoring methodology (§3) and the liquidity-weighted aggregation (§4.1), with persistence contributing modest smoothing and the concentration caps having negligible effect. This is a stronger result than expected given the high cap-binding and persistence-active rates, but it should be read as a property of the current contract universe (in which composite scores happen to be relatively homogeneous across domains) rather than a guaranteed structural property of the index design.

Architectural transparency as a design principle. Most index products of comparable scope (S&P, MSCI, BBD EPU) involve substantial methodology choices that are not subjected to public ablation. The decomposition presented here is offered in the opposite spirit: the index is rules-based, the rules are documented, and their individual contribution to the headline is measurable and disclosed. The finding that the caps have negligible effect in the current sample does not argue for their removal (concentration risk is a structural vulnerability that the current benign environment may not represent), but it does establish that the published headline can be interpreted as a market-signal aggregate with a modest persistence overlay in the current sample, not as a product of its guardrails.

9. Comparison with Existing Uncertainty Measures

9.1 Structural Comparison

Dimension	AURA-PUI	BBD EPU Index
Signal source	Prediction market prices (revealed preferences)	Newspaper article counts (media attention)
Frequency	Hourly computation, daily publication	Monthly (3-component headline); daily (news-only variant)
Latency	Same-day	Weeks to months (3-component); same-day (news-only)
Direction	Forward-looking (market expectations)	Backward-looking (media coverage of past events)
Manipulation resistance	More resistant to non-capitalized narrative distortion; still vulnerable to thin-market microstructure effects	Media framing susceptible
Decomposability	Five domain sub-indices + per-contract attribution	Three components (news, tax code, forecaster disagreement); U.S. specific
Methodology	Rules-based; classification and filters fixed, no day-to-day discretion	Semi-automated with editorial classification
History	From January 2024 (Kalshi candlestick availability)	From 1985 (newspaper archives)

The BBD EPU is the most direct comparator given its prominence, but the broader uncertainty measurement literature includes other relevant benchmarks. Jurado, Ludvigson and Ng (2015) construct a macroeconomic uncertainty index from the common component of forecast-error volatility across hundreds of macro series, a fundamentally different approach that measures the *predictability* of the macroeconomic environment rather than the pricing of specific policy events. Manela and Moreira (2017) derive a news-implied volatility measure from text analysis of the Wall Street Journal, offering a higher-frequency text-based measure than BBD but still operating on media content rather than market-implied probabilities. The AURA-PUI is distinct from both in its reliance on prediction market pricing as the sole signal source.

9.2 Empirical Comparison: BBD EPU

Over 25 months of overlapping data (February 2024 through February 2026), the AURA-PUI and the monthly three-component BBD EPU headline index exhibit strong level correlation (Pearson r = 0.73, $p < 0.0001$; Spearman r = 0.69, $p < 0.001$). All BBD correlations in this section use the monthly three-component U.S. headline; comparison with the daily news-only BBD variant is deferred to the companion paper. Level correlations between persistent macro series can overstate co-movement because both series are highly autocorrelated; the combined evidence from level correlation and the first-difference directional agreement reported below is therefore more informative than either statistic in isolation. Taken together, the results are consistent with the two indices capturing related dimensions of policy uncertainty.

Domain-level correlations:

Pairing	N	Pearson r	p-value
AURA Fiscal & Reg vs. EPU Fiscal + Regulation	25	0.64	0.0005
AURA Monetary vs. EPU Monetary	25	0.25	0.23
AURA Trade & Geo vs. EPU Trade†	16	−0.09	0.74

† N = 16 reflects months in which the AURA Trade & Geopolitical sub-index had sufficient live constituents to produce a monthly value; the domain first produced a sub-index value in November 2024 as Polymarket trade and geopolitical coverage became dense enough to clear the publication gates. The BBD Trade EPU series is available for all 25 months; the shorter overlap is AURA-side, not BBD-side.

The fiscal domain is the strongest match, consistent with the BBD methodology's newspaper-based capture of fiscal and regulatory policy debates. The monetary and trade domains show weak or no correlation at the sub-index level, suggesting the AURA-PUI captures signal in these domains that newspaper-based measures do not, likely reflecting prediction market participants' forward-looking assessment of policy paths versus media coverage of past actions.

Directional agreement. Monthly directional agreement is 46% over 25 months, a figure statistically indistinguishable from 50% and from values as high as 67% given the small sample (95% CI approximately 27%–67%); it should not be read as evidence that the indices disagree on direction. The combination of strong level correlation with weak measured directional agreement is consistent with the two indices capturing the same long-run construct with differentiated short-run timing; it also partly reflects the fact that both series have 1-day autocorrelation above 0.95, which mechanically produces high level correlation even between series whose first-difference innovations are largely unrelated. The AURA-PUI may lead in some episodes where prediction markets partially price policy actions before media coverage intensifies (e.g., the April 2025 tariff sequence), while the BBD EPU may lead when media attention drives public uncertainty perception independent of market pricing (e.g., the January 2025 inauguration). A formal lead-lag framework is deferred to the companion paper.

9.3 Descriptive Distinctness from Market Risk Measures

The AURA-PUI is also descriptively distinct from the CBOE Volatility Index (VIX), which measures 30-day S&P 500 option-implied volatility. The VIX aggregates all sources of equity risk into a single number, while the AURA-PUI targets the policy-linked component of event uncertainty more directly than broad market volatility indices. In the results that follow, AURA-PUI-M and AURA-PUI-F (abbreviated AURA-M and AURA-F) denote the Monetary Policy and Fiscal & Regulatory domain sub-indices, respectively, constructed using the same scoring and weighting pipeline as the headline but restricted to constituents within a single domain.
Cross-correlations (first-differenced daily changes, lags −5 to +5) between the AURA-PUI and financial market variables are uniformly weak:

Pair	Peak Lag	Peak |r|	Lag-0 r
AURA-PUI vs. VIX	0	0.060	+0.060
AURA-PUI vs. S&P 500	0	0.048	−0.048
AURA-PUI vs. DXY	0	0.074	−0.074
AURA-M vs. 2Y Treasury	−1	0.047	−0.030
AURA-F vs. 10Y–2Y Curve	−3	0.079	+0.041

Granger causality. The full test grid covers nine pairs (both directions, 18 tests): the five pairs reported in the cross-correlation table above (adjusted in two cases to use the series most comparable across the Granger framework, so AURA-PUI-F vs. 10Y outright rather than vs. the 10Y–2Y curve slope), plus AURA-PUI-M vs. VIX, AURA-PUI vs. WTI, AURA-PUI vs. Gold, and AURA-PUI-M vs. MOVE. No directional pair achieves significance after Bonferroni correction ($p < 0.0028$).

These results are consistent with the AURA-PUI exhibiting low contemporaneous and lagged correlation with standard financial market returns and volatility proxies, in line with the index's design objective of targeting policy-linked event uncertainty more directly than the broader market risk factors that dominate asset-class volatility indices. The descriptive distinctness reported here should not be interpreted as identification of an orthogonal latent factor. The Granger results in particular should not be interpreted as affirmative evidence of orthogonality: with 815 daily observations concentrated in a single political cycle, statistical power against weak directional dependence is limited, and failure to reject the null is not evidence for it.

Regime-conditional sharpening. The descriptive distinctness is not uniform across uncertainty regimes. The cross-correlation table above reports first-differenced daily changes; the regime-conditional analysis that follows is computed on *levels* rather than first differences, so the unconditional comparator is the full-sample level correlation, not the Lag-0 row of that table. Over the 557 overlapping trading days with available 2-Year yield data, the unconditional Pearson level correlation between AURA-PUI-M and the 2-Year Treasury is $r = 0.07$, essentially uncorrelated. Restricted to days in the top tercile of headline index values ($n = 186$ overlapping days), that level correlation rises to $r = 0.38$. Both series are highly autocorrelated (AURA-PUI-M AR(1) $\approx$ 0.967; 2-Year yields exhibit near-unit-root behavior), so nominal significance overstates the effective degrees of freedom; the result should be read as suggestive of regime-conditional co-movement on levels rather than as a calibrated test. This was the strongest regime-conditional result across the five asset pairs examined; the others exhibited weaker sharpening. A multiple-testing caveat applies: with five pairs and a single regime cut, the effective search space is modest but non-trivial, and the result should be read as suggestive rather than confirmatory. The broader pattern, that at least one domain sub-index exhibits regime-conditional correlation with a related asset-class variable when policy uncertainty is elevated, is suggestive of the regime-conditional information content the index is designed to capture, but remains preliminary given the limited number of pairs examined and the single regime cut evaluated. Formal regime-conditional analysis is deferred to the companion paper.

10. Governance and Methodology Evolution

10.1 Rules-Based Framework

All index construction decisions are determined by documented rules and parameters. Once the classification framework, exclusion filters, and parameters are fixed, no day-to-day discretionary overrides are permitted. Parameter values are listed in Appendix A.

10.2 Methodology Changes

The material-change process comprises the following steps: (1) proposal with rationale and expected impact, (2) parallel backtesting under current and proposed methodologies, (3) impact assessment, (4) documentation update with changelog, (5) advance notification to users. The scaling factor S (§4.6) is permanently fixed at its initial calibration value and is not subject to periodic recalibration; drift in the realized mean index level relative to the nominal 100 target is expected as the prediction market universe evolves, and such drift, whether upward or downward, does not constitute a methodology change. Users should read the index level in relative rather than absolute terms; changes and regime comparisons are invariant to scale.

10.3 Scope of Public Documentation

This paper documents the methodology, parameters, and aggregation logic required to understand the index's structural behavior. Operational inputs maintained separately from the methodology (including the contract classification dictionary, keyword and pattern exclusion filters, the numeric value of the scaling factor S (§4.6), and implementation code) are not published in this document. Because S is withheld, users applying the documented methodology can reproduce the change-based time series and all relative-level comparisons, but not the absolute published index level. These inputs may be shared with research partners and licensees under separate agreement. Inquiries regarding data access, replication, or licensing may be directed to support@aurelianhq.com; the Aurelian website (Aurelianhq.com) provides further information about the AURA-PUI and the publishing entity.

11. Risk Disclosures and Limitations

Data Source Risk. The index depends on the continued operation and data accessibility of prediction market platforms. Regulatory changes affecting the legality or operation of prediction markets could materially impact data availability.

Liquidity Risk. During periods of low prediction market activity, the number of eligible contracts may decline below the publication threshold. The domain persistence layer and quality gates mitigate but do not eliminate this risk. The current $500/day volume floor explicitly reflects the nascent state of prediction market liquidity.

Microstructure Limitations. Polymarket historical spread data is synthetic (estimated from heuristics). Real BBO data is available only going forward as the cache accumulates. The volume-only weighting fallback for synthetic-spread contracts limits but does not eliminate measurement uncertainty from this source.

Venue-Composition Risk (Polymarket). The index currently sources Polymarket data from the offshore decentralized CLOB. Following Polymarket's CFTC Amended Order of Designation (November 25, 2025) and the limited rollout of Polymarket US in December 2025, the composition of liquidity and informed participation across the offshore and U.S.-regulated venues may evolve over time in either direction. Liquidity shifts from the offshore venue toward Polymarket US could degrade the offshore data source the index currently ingests; consolidation on the offshore venue or parallel growth on both could leave the current source stable or enhanced. Any migration of the index pipeline to Polymarket US, or any material change in the offshore venue's microstructure properties attributable to this bifurcation, will be disclosed as a methodology change per §10.2.

Model Risk. The scoring methodology involves parameter choices (α, γ, layer weights, caps, volume threshold) that affect index behavior. All parameters are documented in Appendix A and subject to periodic review per §10.2.

Short History. The AURA-PUI backtest begins in January 2024, constrained by the availability of Kalshi candlestick data. The 815-day sample (through April 2026) spans a single broad political cycle with limited macroeconomic regime diversity. Historical relationships observed in the backtest should not be assumed to persist.

Nascent Asset Class. Prediction markets are an evolving asset class with a limited history of institutional use. Market microstructure, participant composition, and regulatory frameworks are all subject to change.

Contract Universe. The index's representativeness depends on the breadth of policy-relevant contracts available on eligible platforms. As documented in §8.5, the Macroeconomic Outcomes and Trade & Geopolitical domains are structurally thinner than the Fiscal & Regulatory and Monetary Policy domains, averaging fewer than 5 eligible contracts each over the backtest period. This asymmetry means the headline index is disproportionately driven by fiscal and monetary policy uncertainty. Coverage gaps during quiet periods are mitigated by the domain persistence layer but not eliminated.

References

Arrow, K.J., et al. (2008). “The Promise of Prediction Markets.” Science, 320(5878), 877–878.
Baker, S.R., Bloom, N. and Davis, S.J. (2016). “Measuring Economic Policy Uncertainty.” Quarterly Journal of Economics, 131(4), 1593–1636.

Goodell, J.W., McGee, R.J. and McGroarty, F. (2020). “Election Uncertainty, Economic Policy Uncertainty and Financial Market Uncertainty: A Prediction Market Analysis.” Journal of Banking & Finance, 110, 105684.

Jurado, K., Ludvigson, S.C. and Ng, S. (2015). “Measuring Uncertainty.” American Economic Review, 105(3), 1177–1216.

Manela, A. and Moreira, A. (2017). “News Implied Volatility and Disaster Concerns.” Journal of Financial Economics, 123(1), 137–162.

Manski, C.F. (2006). “Interpreting the Predictions of Prediction Markets.” Economics Letters, 91(3), 425–429.

Roll, R. (1984). “A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market.” Journal of Finance, 39(4), 1127–1139.

Wolfers, J. and Zitzewitz, E. (2004). “Prediction Markets.” Journal of Economic Perspectives, 18(2), 107–126.

Appendix A: Complete Parameter Table

Parameter	Value	Description
ALPHA	0.70	PDS vs. volatility weight within Layer 1
LAYER1_WEIGHT	0.50	L1 weight in composite score
LAYER2_WEIGHT	0.50	L2 weight in composite score
VOL_WINDOW	14 days	Trailing volatility window
VOL_NORM_WINDOW	60 days	CDF normalization pool window
VOL_MIN_HISTORY	7 days	Minimum history for volatility
BID_ASK_BOUNCE_GAMMA	0.25	Bid-ask bounce attenuation factor (Roll-inspired; §3.3)
SPREAD_VOL_WINDOW	7 days	Spread/volume averaging window
L2_MOMENTUM_WINDOW	7 days	Price change window for Layer 2
L2_EWM_HALFLIFE	60 days	EWM normalization halflife
L2_WARMUP_DAYS	90 days	Expanding window before EWM
L2_SIGMOID_SCALE	2.0	Sigmoid steepness
L2_OUTPUT_HALFLIFE	2 days	Per-contract L2 smoothing
LIQUIDITY_VOLUME_SHARE	0.60	Volume weight in liquidity score
LIQUIDITY_SPREAD_SHARE	0.40	Inverse-spread weight
VOLUME_ACTIVATION_THRESHOLD	0.50	Spread activation gate
MIN_VOLUME_THRESHOLD	$500/day	Eligibility volume gate

MIN_VOLUME_HARD_FLOOR	$500/day	Immediate removal threshold
MIN_LIQUIDITY_SPREAD	0.05	Max bid-ask spread (probability)
MIN_DAYS_TO_EXPIRY	7 days	Expiration cutoff
MIN_PDS_THRESHOLD	0.10	PDS eligibility floor
MIN_BASKET_SIZE	5	Publication minimum contracts
MIN_DOMAINS_FOR_PUBLICATION	2	Publication minimum domains
SINGLE_CONTRACT_CAP_BASE	0.08	Lowest-CS per-contract cap
SINGLE_CONTRACT_CAP_MAX	0.22	Highest-CS per-contract cap
DOMAIN_CONCENTRATION_CAP	0.50	Maximum domain weight
INTRA_DOMAIN_CONTRACT_CAP	0.40	Within-domain sub-index cap
RAMP_DOWN_HOURS	168	Expiration ramp (7 days)
SCALING_FACTOR	[withheld; see §10.3]	Maps mean CS to ~100
RECONSTITUTION_ENTRY_DAYS	1 day	Entry buffer
RECONSTITUTION_EXIT_DAYS	3 days	Exit buffer
ENTRY_BLEND_RAMP_DAYS	7	New-contract weight ramp
PERSISTENCE_HALFLIFE_HOURS	720	30-day component halflife (effective contribution halflife ~360h / 15 days; see §5.3)
PERSISTENCE_FLOOR_WEIGHT	0.10	Initial floor weight
PERSISTENCE_MAX_TOTAL	0.30	Max total persisted weight
PERSISTENCE_MIN_LIVE	2	Activation threshold

PERSISTENCE_MIN_HISTORY	7 days	Min domain history
PERSISTENCE_EPSILON	0.005	Dropout weight threshold
STALENESS_FLAG_HOURS	72	Stale data flag
STALENESS_EXCLUSION_HOURS	96	Stale data exclusion
CACHE_TTL_HOURS	20	Price cache expiry

Appendix B: Probability Dispersion Score Reference

Probability p	PDS = 4p(1−p)
0.50	1.000
0.45 / 0.55	0.990
0.40 / 0.60	0.960
0.35 / 0.65	0.910
0.30 / 0.70	0.840
0.25 / 0.75	0.750
0.20 / 0.80	0.640
0.15 / 0.85	0.510
0.10 / 0.90	0.360
0.05 / 0.95	0.190
0.025 / 0.975	0.098

The PDS eligibility floor (0.10) corresponds to approximately $p = 0.025$ or $p = 0.975$.

Appendix C: Replicability Reference

Item	Value
Calculation engine version	v3 (multi-platform)
Parameter configuration	Appendix A (all parameters frozen as of publication date)
Backtest data cutoff	April 13, 2026
Backtest sample window	January 4, 2024 – April 13, 2026 (815 valid daily observations)
Eligible platforms	Kalshi (CFTC-regulated), Polymarket (decentralized CLOB)
Snapshot granularity	Hourly (24 per day, UTC-aligned)
Daily publication snapshot	23:00 UTC (last hourly snapshot)
Price clipping (eligibility)	[0.01, 0.99]
Price clipping (logit transform)	[0.001, 0.999]

All parameters listed in Appendix A are frozen as of the publication date. Any future methodology changes follow the governance process specified in §10.2. The backtest uses the identical calculation engine, parameters, and eligibility criteria documented in this paper; no ex-post parameter tuning was performed against the metrics reported in §8.

Appendix D: Persistence Decay-Variant Sensitivity

To assess whether the headline is sensitive to the specific persistence decay specification, the index was re-computed over the full backtest under four configurations. All configurations use the 720-hour component halflife and the identical production pipeline (eligibility, scoring, liquidity weights, entry blending, expiration ramp, concentration caps); they differ only in what decays. Minor numerical differences between the Config A mean reported here (98.41) and the headline mean reported in §8.2 (98.35) reflect boundary-condition handling in the ablation recomputation; they do not affect any conclusions in this appendix. The live-only reference reported in the final row of this appendix (mean = 99.81, std = 21.12) is the analogous ablation-pipeline recomputation of the "live-only" column that appears in §8.6 (mean = 99.75, std = 21.09), and the two differ for the same reason.

Variant	Weight Decay	Value Decay	Effective Contrib. Halflife	Mean Index	Std Dev.
A (production)	720h	720h (toward 0)	360h (compound)	98.41	20.05
B (weight-only)	720h	None (held constant)	720h	99.42	20.58
C (value-only)	None (held at floor)	720h (toward 0)	720h	92.42	17.21
D (decay-to-neutral)	720h	720h (toward cross-sect. mean)	720h	99.53	20.71

Variant	AR(1)	Corr. w/ A	Corr. w/ Live-only
A (production)	0.967	1.000	0.997
B (weight-only)	0.968	0.999	0.997
C (value-only)	0.961	0.954	0.939
D (decay-to-neutral)	0.968	0.999	0.998

Live-only reference (persistence disabled entirely): mean = 99.81, std = 21.12, AR(1) = 0.966.

Findings. Configurations A, B, and D produce near-identical series (pairwise correlations above 0.998, mean levels within 1.1 points). Configuration C (value-only decay with constant weight) is the outlier: holding weight constant while decaying value toward zero drags the index 6 points below production and reduces the correlation with the headline to r = 0.954. This is the degenerate case: a floor weight that does not shrink pulls the index toward zero as persisted values decay, without the weight automatically disengaging. Configuration D, which decays the persisted value toward the cross-sectional mean rather than toward zero to explicitly strip any directional carry, correlates with Config A at $r = 0.999$, indicating the production specification does not carry meaningful stale directional information. Excluding Config C, the choice of decay specification is a second-order design decision with negligible impact on the published series.