Itô | MARKETS

METHODOLOGY

# Thematic Prediction-Market Basket Methodology

*Construction, Rebalancing & NAV Measurement for Binary-Claim Portfolios*

A mathematical research note formalizing thematic portfolios of binary prediction-market claims: admissible information set, oriented claims, selection, weighting, lifecycle, and the daily NAV recursion.

**Produced by Itô Markets**

Thematic Baskets · Quantitative Research · Version 1.0 · Q2 2026

# Contents

# Abstract

This note formalizes a family of thematic portfolios built from binary prediction-market claims. Each theme is treated as a tradable economic exposure to a latent event class. The methodology specifies the admissible information set, the transformation of YES/NO contracts into oriented thematic claims, the relevance and maturity scores used to select claims, the constrained portfolio-weighting problem, the cash-reserve rule, the treatment of settlement and missing marks, and the daily net-asset-value recursion. The empirical figures at the end show the resulting index paths on an inception-normalized basis.


**Empirical window used for figures.** 2022-01-01 through 2026-03-07. The diagrams cover 10 themes. Index levels are normalized to 100 at inception.

# 1 Economic Object

Let there be a sequence of rebalance dates $\{\tau_m\}_{m\geq 0}$. At each $\tau_m$, an investor observes a set of binary claims. A binary claim has a market price $p_i(t)$ in $[0, 1]$ and a terminal resolution $R_i$ in $\{0, 1\}$. Economically, $p_i(t)$ is the market-implied price of one dollar paid if event $i$ resolves YES. The basket problem is to create theme-level portfolios whose payoffs increase when a specified latent theme becomes more likely or is realized.

$$B_b(\tau_m) \;=\; \{\, i \in U(\tau_m) : i \text{ economically admissible for theme } b \text{ at } \tau_m \,\} \tag{1}$$

Formula (1) defines the candidate set for theme $b$. $U(\tau_m)$ is the observable universe at the rebalance date; membership requires that the claim is known, economically active, sufficiently interpretable, and inside the horizon/risk constraints. The set is ex-ante: it cannot depend on information revealed after $\tau_m$.

$$V_b(t) \;=\; C_b(t) \;+\; \sum_{i\in H_b(t)} n_{i,b}(t)\, x_i(t) \tag{2}$$

Formula (2) defines the marked value of theme basket $b$. $C_b(t)$ is cash, $H_b(t)$ is the held claim set, $n_{i,b}(t)$ is the number of units of oriented claim $i$, and $x_i(t)$ is the side-adjusted price. The basket is therefore a portfolio of bounded binary-risk assets plus cash.

# 2 Information Set and No-Lookahead Constraint

Denote by $\mathcal{F}_t$ the information available at time $t$. Every selection variable, orientation variable, weight, cash reserve, and rebalance decision must be $\mathcal{F}_{\tau_m}$-measurable. This is the central econometric restriction: the methodology may exploit contemporaneous prices, text, maturity, classification, and lifecycle status, but not later resolutions or later revisions.

$$w_{i,b,m},\; d_{i,b,m},\; c_{b,m},\; H_b(\tau_m) \quad \text{are } \mathcal{F}_{\tau_m}\text{-measurable} \tag{3}$$

Formula (3) states the no-lookahead condition. The selected side $d$, portfolio weight $w$, cash reserve $c$, and held set $H$ are all functions of the information set at the rebalance date. This makes the resulting NAV a historical reconstruction of a feasible rule rather than an ex-post fit.

$$I_{i,m} \;=\; \mathbf{1}\{a_i \leq \tau_m < e_i\} \cdot \mathbf{1}\{\text{unresolved}_i(\tau_m)\} \cdot \mathbf{1}\{\text{markable}_i(\tau_m)\} \tag{4}$$

Formula (4) is the admissibility indicator. The symbol $a_i$ is the first date on which claim $i$ is economically observable, $e_i$ is its terminal date, $\text{unresolved}_i(\tau_m)$ requires that the claim has not already resolved, and $\text{markable}_i(\tau_m)$ requires that it can be valued or carried under the stated marking convention.

## 3 Oriented Binary Claims

Prediction-market contracts are quoted as YES prices, but a theme can require either a positive YES exposure or a positive NO exposure. For example, a governance-stability theme may prefer the NO side of a crisis contract, while a conflict-escalation theme may prefer the YES side of an escalation contract. The methodology therefore converts every selected claim into an oriented payoff.

$$d_i \in \{+1,\, -1\}; \qquad d_i = +1 \text{ for YES exposure}, \quad d_i = -1 \text{ for NO exposure} \tag{5}$$

Formula (5) introduces the side variable. A positive orientation means the basket benefits when the event resolves YES. A negative orientation means the basket benefits when the event resolves NO.

$$x_i(t) \;=\; \tfrac{1}{2}(1+d_i)\,p_i(t) \;+\; \tfrac{1}{2}(1-d_i)\left[\,1-p_i(t)\,\right] \tag{6}$$

Formula (6) converts quoted YES prices into side-adjusted prices. If $d_i = +1$, then $x_i(t) = p_i(t)$. If $d_i = -1$, then $x_i(t) = 1 - p_i(t)$. This lets YES and NO positions be valued with one common price variable.

$$Y_i \;=\; \tfrac{1}{2}(1+d_i)\,R_i \;+\; \tfrac{1}{2}(1-d_i)\left[\,1-R_i\,\right] \tag{7}$$

Formula (7) is the corresponding terminal payoff. It equals one when the selected side wins and zero when the selected side loses. This is why the entire portfolio can be interpreted as a basket of one-dollar binary claims after orientation.

$$r_i^{\text{hold}}(t,\, t{+}1) \;=\; \frac{x_i(t{+}1) - x_i(t)}{x_i(t)} \tag{8}$$

Formula (8) is the single-claim holding-period return before transaction costs and lifecycle events. Because $x_i(t)$ is bounded, percentage returns can be large when the oriented price is very small; this motivates minimum-price, maximum-price, and concentration controls.

## 4 Horizon Normalization

Binary-event prices with different expiries are not directly comparable. A 20 percent probability over one month and a 20 percent probability over one year express different intensities. The methodology maps a claim price into a constant-hazard approximation and then evaluates that hazard over a target horizon.

$$\lambda_i(t) \;=\; \frac{-\log\!\left[\,1 - x_i(t)\,\right]}{D_i(t)} \tag{9}$$

Formula (9) defines the implied daily event intensity under a constant-hazard approximation. $D_i(t)$ is the positive number of days remaining to expiry. If the oriented event probability is high, the implied hazard is high; if it is near zero, the hazard is near zero.

$$x_i^H(t) \;=\; 1 - \exp\!\left[\,-\lambda_i(t)\,H\,\right] \tag{10}$$

Formula (10) converts the observed claim into a horizon-$H$ probability. This gives a common-horizon signal for ranking claims whose contractual maturities differ. The transformation is used for comparability, not as a structural model of event arrivals.

$$M_i(t;\, D^*) \;=\; \exp\!\left(-\frac{|D_i(t) - D^*|}{D^*}\right) \tag{11}$$

Formula (11) is a smooth maturity-proximity score around the target tenor $D^*$. It equals one at the target and decays as maturity moves away from the target. The purpose is to prefer claims that express the intended horizon while still allowing scarce themes to hold imperfect maturities.

## 5 Theme Relevance and Direction

Each theme $b$ has an economic description: a latent event class, an intended exposure direction, preferred maturities, probability ranges, and diversification slots. For each claim $i$, the method computes separate evidence for direct exposure and inverse exposure. The direction is not purely textual; it combines semantic fit, category fit, maturity fit, and side-specific economic interpretation.

$$E^{+}_{i,b} \;=\; a_b'\, z(A_i) \;+\; k_b'\, z(K_i) \;+\; q_b\, Q_i \;+\; m_b\, M_i \tag{12}$$

Formula (12) is the direct-exposure evidence score. $A_i$ represents theme attributes, $K_i$ represents text or event descriptors, $Q_i$ is a claim-quality term, and $M_i$ is maturity fit. The vectors $a_b$ and $k_b$ are theme-specific loadings. The $z$ operator means variables are standardized within the contemporaneous candidate universe so that heterogeneous signals can be combined.

$$E^{-}_{i,b} \;=\; a_b^{-\prime}\, z(A_i) \;+\; k_b^{-\prime}\, z(K_i) \;+\; q_b^{-}\, Q_i \;+\; m_b^{-}\, M_i \tag{13}$$

Formula (13) is the inverse-exposure evidence score. It asks whether the opposite side of the claim is the cleaner expression of the theme. This is essential for themes where resilience, de-escalation, or non-occurrence is the economically desired exposure.

$$d_{i,b} = +1 \;\text{ if }\; E^{+}_{i,b} \ge E^{-}_{i,b}; \qquad \text{otherwise }\; d_{i,b} = -1 \tag{14}$$

Formula (14) chooses the side of the contract. The YES side is held only when direct evidence is at least as strong as inverse evidence; otherwise the NO side is held. The rule makes direction endogenous to the economic meaning of the claim relative to the theme.

$$S_{i,b} \;=\; z(E_{i,b}) + \beta_b\, P_i + \gamma_b\, T_i + \delta_b\, Q_i - \rho_b\, R_i - \eta_b\, Z_i \tag{15}$$

Formula (15) is the total selection score. $E_{i,b}$ is the chosen-side evidence, $P_i$ is a probability-band adjustment, $T_i$ is a tenor-band adjustment, $Q_i$ is quality, $R_i$ penalizes very short-dated claims, and $Z_i$ penalizes nearly certain or economically saturated claims. The coefficients let different themes prefer different tenors, confidence levels, and risk profiles.

## 6 Probability and Tenor Bands

The probability band is an economic prior about where a binary claim is most useful. Very cheap claims can be lottery-like and noisy; very expensive claims can be nearly resolved and offer little convexity. The tenor band is an economic prior about the event horizon the theme is meant to represent.

$$P_i = \varphi_p\left(x_i^H\right), \quad \text{positive in the preferred probability interval, negative outside admissible intervals} \tag{16}$$

Formula (16) describes the probability adjustment. It is not a forecast by itself; it is a ranking correction that favors claims whose market-implied probability is in the region where the theme wants exposure.

$$T_i = \varphi_D(D_i), \quad \text{highest near the target tenor, lower for too-short or too-long claims} \tag{17}$$

Formula (17) describes the tenor adjustment. A claim expiring too soon can behave like realized-event settlement risk rather than thematic exposure. A claim expiring too late can dilute the intended horizon. The adjustment balances relevance and tradability.

$$i \in H_b(\tau_m) \quad \text{only if} \quad S_{i,b} \geq s_b^{\min} \quad \text{and} \quad I_{i,m} = 1 \tag{18}$$

Formula (18) is the threshold rule. A claim must both pass the economic admissibility filter and exceed the minimum score for the theme. This prevents weakly related contracts from entering a basket only because a theme is sparse.

## 7 Deduplication and Diversification

Many prediction-market contracts are close substitutes: several contracts can ask the same question with different wording, or multiple claims can reference the same event family. If all are held simultaneously, the basket can become a leveraged bet on one event. The methodology therefore deduplicates event-equivalent claims and caps exposure by economic slot.

$$G_{j,b} \; = \; \{\, i : \text{claim } i \text{ belongs to event-equivalence class } j \text{ for theme } b \,\} \tag{19}$$

Formula (19) defines equivalence classes. Claims in the same class are economically redundant enough that they should compete with one another rather than all enter at full weight.

$$i_j^* \; = \; \arg\max_{i \in G_{j,b}} S_{i,b} \tag{20}$$

Formula (20) keeps the highest-scoring representative of each equivalent event family. The selected representative is the one with the best combination of relevance, maturity, quality, and probability-band fit.

$$\sum_{i \in g} w_{i,b,m} \; \leq \; u_{g,b} \tag{21}$$

Formula (21) is the slot-cap constraint. A slot $g$ can be a geopolitical region, policy topic, event mechanism, or other economic bucket. The cap $u_{g,b}$ limits the weight that any one bucket can contribute to a theme basket.

$$w_{i,b,m}^{\text{new}} \; = \; w_{i,b,m}^{\text{old}} \times \min\left(1, \; \frac{u_{g,b}}{\sum_{j \in g} w_{j,b,m}^{\text{old}}}\right) \tag{22}$$

Formula (22) shows the proportional scaling used when a slot breaches its cap. Excess weight is reallocated to uncapped claims where possible. This preserves relative ranking within a constrained group while reducing concentration.

# 8 Raw Weights, Caps, and Cash

After selection, the methodology assigns nonnegative raw weights, then normalizes them subject to single-name and slot constraints. Weights are not proportional only to theme score; they also reflect tail relevance, slot completion, and claim quality. A cash sleeve is kept deliberately because sparse or illiquid themes should not be forced to be fully invested.

$$r_{i,b,m} \;=\; \max\!\big(s_{\min},\, S_{i,b,m}\big)\,\big[1+\alpha\,\Pi_{i,b,m}\big]\,\big[1+\gamma\,L_{i,b,m}\big]\,\big[1+\delta\,Q_{i,m}\big] \tag{23}$$

Formula (23) defines the raw pre-cap score. $\Pi$ is a tail-probability relevance term, $L$ is a slot-completion term, and $Q$ is claim quality. The multiplicative form means that a claim must first be relevant; boosts refine a relevant claim rather than rescue an irrelevant one.

$$\hat{w}_{i,b,m} \;=\; \frac{r_{i,b,m}}{\sum_{j\in H_b(\tau_m)} r_{j,b,m}} \tag{24}$$

Formula (24) converts raw scores into preliminary fully-invested weights. At this stage the weights sum to one across non-cash claims before final caps and the cash reserve are applied.

$$0 \;\le\; \hat{w}_{i,b,m} \;\le\; u_{i,b}; \qquad \sum_i \hat{w}_{i,b,m} = 1 \tag{25}$$

Formula (25) is the single-name cap problem. If a preliminary weight exceeds the name cap $u_{i,b}$, it is clipped and the remaining mass is redistributed over uncapped names. The iteration continues until all weights satisfy the cap and total non-cash weight equals one.

$$w_{i,b,m} \;=\; (1-c_{b,m})\,\hat{w}_{i,b,m} \tag{26}$$

Formula (26) applies the cash sleeve. The investable claim weights sum to $1 - c_{b,m}$; the remaining $c_{b,m}$ is held in cash. Cash is not a residual error. It is an explicit risk-control variable for sparse, unstable, or high-turnover baskets.

$$c_{b,m} \;=\; \min\!\Big(c_{\max},\; \max\big(c_{\min},\; c_b^0 + f_T(\mathrm{Turn}_{b,m}) + f_R(\mathrm{Conc}_{b,m}) + f_N(\mathrm{Scarcity}_{b,m})\big)\Big) \tag{27}$$

Formula (27) defines dynamic cash. The baseline cash $c_b^0$ is increased when turnover is high, risk concentration is high, or too few clean claims are available. The outer min/max keeps cash inside a pre-specified economically acceptable interval.

# 9 Rebalancing, Turnover, and Transaction Friction

At each rebalance date the basket compares current holdings to target weights. Turnover is a measure of the notional adjustment required to move from the old portfolio to the new target. Even when explicit transaction costs are set to zero for an index reconstruction, turnover is economically important because it identifies dates when the strategy is relying on a large portfolio change.

$$\mathrm{Turn}_{b,m} \;=\; \tfrac{1}{2}\sum_i \big|w^{\text{target}}_{i,b,m} - w^{\text{post}}_{i,b,m-1}\big| \tag{28}$$

Formula (28) is one-way portfolio turnover. The factor one-half avoids double-counting a sale and the corresponding purchase when the portfolio remains fully invested. If the cash sleeve changes, turnover captures the movement into or out of cash as well.

$$\mathrm{TC}_{b,m} \;=\; \kappa_b\,\mathrm{Turn}_{b,m}\,V_b(\tau_m^-) \tag{29}$$

Formula (29) is the generic proportional transaction-cost model. $\kappa_b$ is the cost per unit of one-way notional turnover. When $\kappa_b = 0$, the index is a frictionless mark-to-market reconstruction; when $\kappa_b > 0$, costs lower NAV at rebalance.

$$n_{i,b}(\tau_m) \;=\; \frac{V_b(\tau_m^+)\,w_{i,b,m}}{x_i(\tau_m)} \tag{30}$$

Formula (30) converts target weights into units. The basket buys enough units of each oriented claim so that its dollar exposure equals the target weight times post-rebalance basket value. This unit representation is what allows daily NAV to be computed between rebalances.

## 10 Lifecycle, Settlement, and Missing Marks

Binary claims can stop trading, resolve, become stale, or become economically certain before formal settlement. These lifecycle states matter more than in ordinary equity indexing because terminal payoff is discontinuous. The methodology treats settlement as a payoff event, certainty as a forced exit condition, and missing observations as a valuation-risk state.

$$x_i^*(t) \;=\; Y_i \;\text{ if claim } i \text{ has resolved by } t; \quad \text{otherwise } \; x_i(t) \tag{31}$$

Formula (31) defines the lifecycle-adjusted mark. Once resolution is known, the claim is valued at its oriented payoff. Before resolution, it is marked at the oriented market price subject to the observability convention.

$$\mathrm{Exit}_i(t) \;=\; \mathbf{1}\{x_i(t) \le \text{lower}\} + \mathbf{1}\{x_i(t) \ge \text{upper}\} + \mathbf{1}\{D_i(t) \le \text{roll\_floor}\} \tag{32}$$

Formula (32) identifies forced-exit states. Extremely cheap, extremely expensive, or nearly expiring claims cease to be clean thematic exposures; they are either close to dead, close to certain, or dominated by settlement mechanics.

$$\mathrm{Coverage}_b(t) \;=\; \sum_{i \in H_b(t)} w_{i,b}(t)\,\mathbf{1}\{x_i(t) \text{ observable}\} \tag{33}$$

Formula (33) measures price observability. A low value means the daily NAV is being supported by carry-forward, cash, or reduced investability rather than fresh marks. It is a diagnostic for interpreting the reliability of short-horizon index moves.

## 11 Daily NAV Recursion

Between rebalance dates, the basket is a fixed-units portfolio except for lifecycle exits and reinvestment rules. The daily NAV recursion is therefore a standard self-financing portfolio equation adapted to oriented binary claims.

$$C_b(t) \;=\; C_b(t{-}1)\,\big[1 + r_f(t)\,\Delta t\big] \;+\; \text{cash flows from exits and rebalances} \tag{34}$$

Formula (34) tracks the cash account. $r_f(t)$ is the cash rate over the daily interval $\Delta t$. If the cash rate is zero, cash is carried at par. Settlement proceeds and forced exits move through this account before redeployment.

$$V_b(t) \;=\; C_b(t) \;+\; \sum_{i \in H_b(t)} n_{i,b}(t)\, x_i^*(t) \tag{35}$$

Formula (35) is the daily basket value. The star on $x$ indicates lifecycle-adjusted marks: unresolved claims use oriented prices; resolved claims use oriented payoffs.

$$R_b(t) \;=\; \frac{V_b(t)}{V_b(t-1)} - 1 \tag{36}$$

Formula (36) is the daily basket return. It is the return of the whole theme portfolio, including cash and all active claim marks.

$$L_b(t) \;=\; 100 \cdot \prod_{u \le t} \big[1 + R_b(u)\big] \tag{37}$$

Formula (37) converts returns into an index level with inception equal to 100. This is the NAV time series shown in the theme-level diagrams. A level of 150 means the basket has compounded to 1.5 times inception value; a level of 50 means it has halved.

## 12 Aggregate Thematic Index

A multi-theme aggregate index combines theme-level baskets with ex-ante theme weights. Active weights are renormalized over themes that have a valid level on a given date. This produces an aggregate thematic barometer while keeping each underlying theme interpretable on its own.

$$\bar{\Omega}_{b,t} \;=\; \frac{\Omega_b\, \mathbf{1}\{L_b(t)\text{ observed}\}}{\sum_k \Omega_k\, \mathbf{1}\{L_k(t)\text{ observed}\}} \tag{38}$$

Formula (38) renormalizes theme weights over observable themes. $\Omega_b$ is the intended strategic weight of theme $b$. If every theme is observed, $\overline{\Omega}_{b,t}$ equals the normalized strategic weight.

$$L_A(t) \;=\; \sum_b \bar{\Omega}_{b,t}\, L_b(t) \tag{39}$$

Formula (39) defines the aggregate index level as a weighted combination of theme levels. This is a level aggregation rather than a claim-level portfolio; its purpose is to summarize the family of thematic baskets.

$$\text{Coverage}_A(t) \;=\; \sum_b \bar{\Omega}_{b,t}\, \text{Coverage}_b(t) \tag{40}$$

Formula (40) gives aggregate observability. It is the weighted fraction of the aggregate index supported by observable claim marks, and it should be read together with the NAV path whenever the index moves sharply.

## 13 Empirical NAV Diagrams

The following figures illustrate the index mechanics after applying the methodology to a historical sample. They are not a separate model; they are diagnostic views of formulas (34) through (40). The diagrams should be read as normalized index paths, not as dollar profit-and-loss for a particular investor.

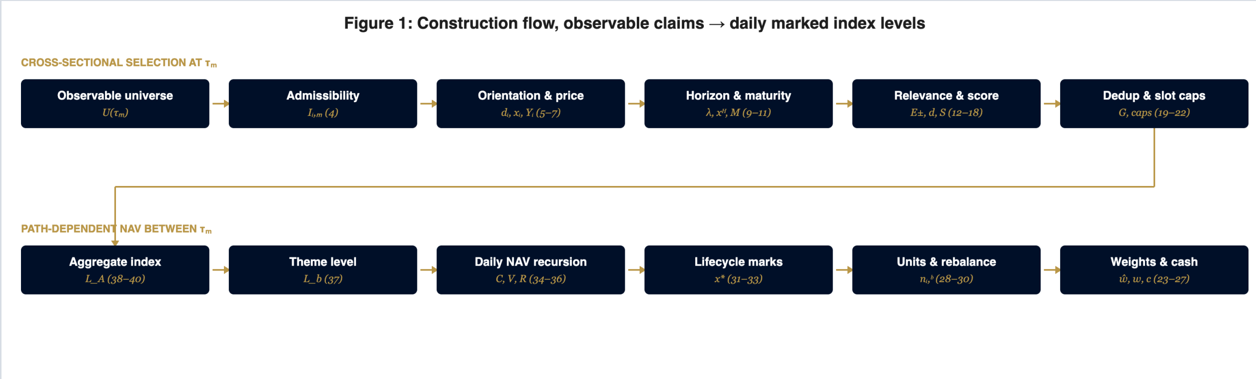


*Figure 1. Mathematical construction flow from observable binary claims to daily marked index levels.*

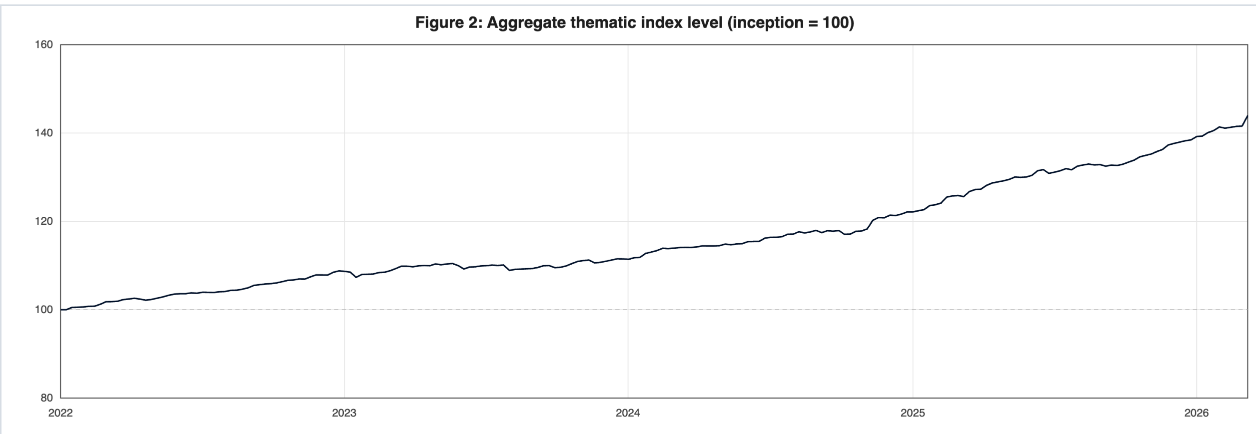


*Figure 2. Aggregate thematic index level, inception normalized to 100.*

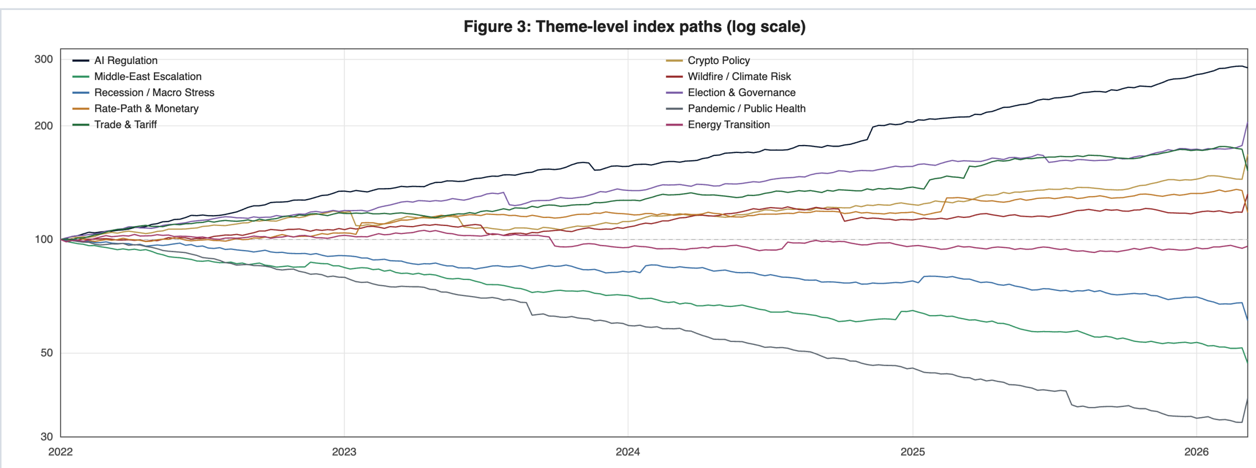


*Figure 3. Theme-level index paths on a logarithmic scale. The log axis is used because compounded outcomes diverge materially across themes.*

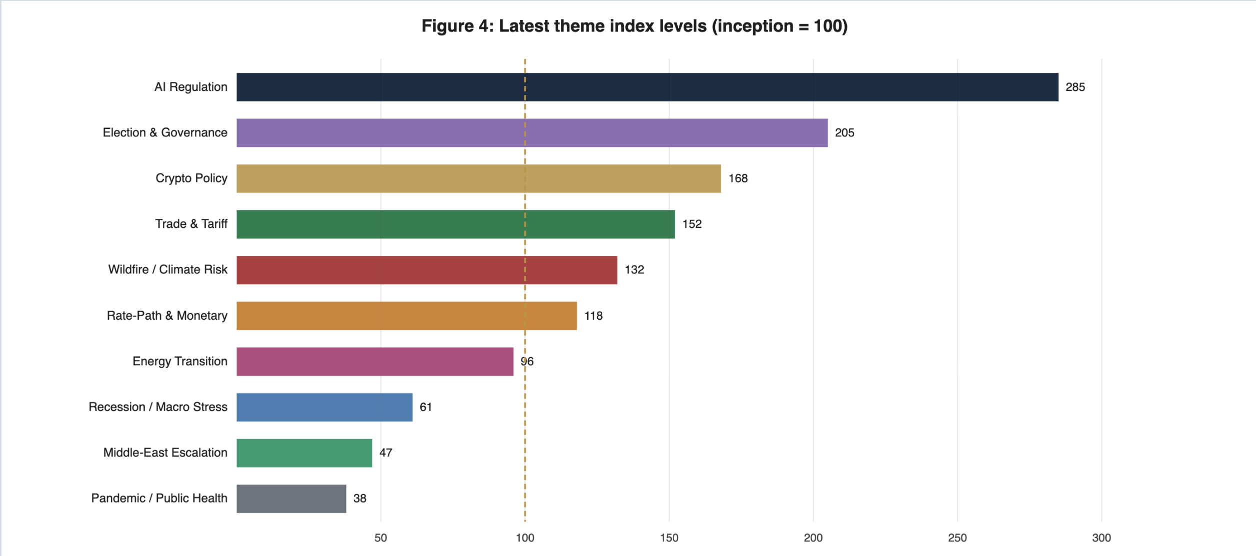


*Figure 4. Latest theme index levels. The vertical reference at 100 is the inception-normalized base level.*

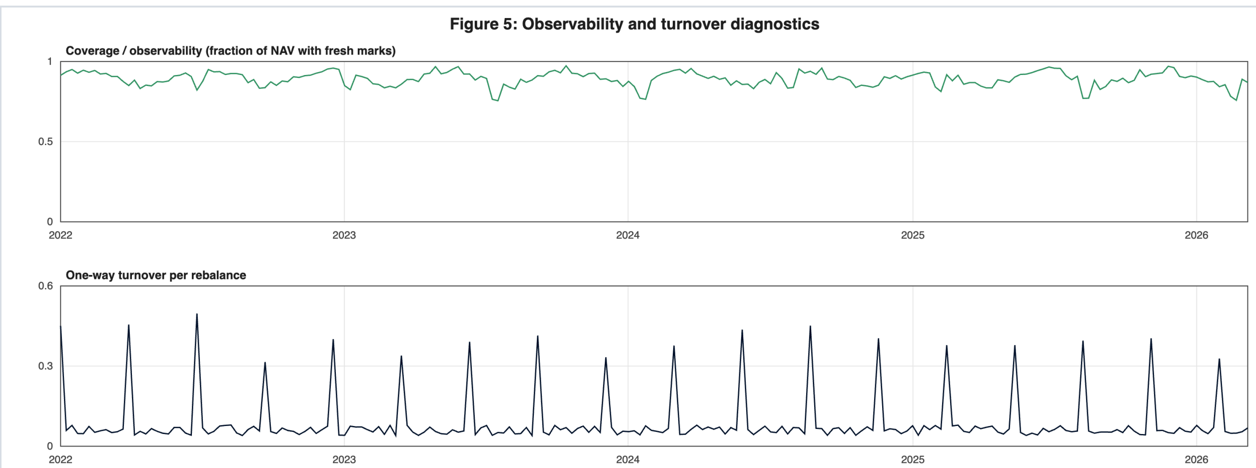


*Figure 5. Observability and turnover diagnostics. Low observability weakens confidence in daily marks; high turnover indicates large changes in target exposures.*

# 14 Economic Interpretation of the NAV Paths

A theme-level NAV rises when the oriented claims held by the theme appreciate or settle favorably. Because the claims are binary, sharp changes usually mean one of three things: an event probability repriced, a settlement became known, or the basket rolled into a materially different set of claims. The observability and turnover diagnostics separate these mechanisms. A smooth rise with high observability is a marked repricing. A jump around lifecycle events is a settlement or forced-exit effect. A jump with high turnover is partly a construction effect from changing the claim set.

The methodology is intentionally cross-sectional at rebalance and path-dependent between rebalances. Cross-sectionally, claims compete to enter a theme through equations (12) to (18). Path-dependently, once a claim is held, its contribution to NAV depends on units purchased in equation (30), lifecycle marks in equation (31), and the daily portfolio recursion in equations (34) to (37). This distinction matters for interpretation: the selection rule explains why a claim enters, while the NAV recursion explains how that claim affects realized index performance.

## 15 Econometric Caveats

- Sparse markets create selection uncertainty: a theme with few claims can be economically correct but statistically fragile.
- Prices are not pure probabilities when liquidity, collateral, fees, or market-microstructure frictions are large.
- Textual and categorical relevance is an imperfect proxy for economic exposure, especially when contracts contain compound events.
- Settlement jumps can dominate ordinary mark-to-market variation because payoffs are binary.
- Observed historical NAV is conditional on the available claim universe; it is not evidence that the same opportunity set will exist in future months.

## 16 Theme Set Used in the Empirical Diagrams

The empirical diagrams cover ten themes spanning geopolitical, macroeconomic, policy, climate, and technology exposures. Each theme is a latent event class with an intended exposure direction; the exact constituent claims are deployment-specific and are re-selected at every rebalance under equations (12)–(18). The representative set is:

| **Theme** | **Latent event class** | **Intended exposure** |
|---|---|---|
| AI Regulation | AI policy and enforcement actions | YES on regulatory milestones |
| Crypto Policy | Stablecoin, enforcement, reserve rules | Side per claim |
| Middle-East Escalation | Cross-contract conflict escalation | YES on escalation |
| Wildfire / Climate Risk | Regional wildfire and severity | YES on severity |
| Recession / Macro Stress | GDP, labor, default, slowdown | YES on slowdown |
| Election & Governance | Election and governance stability | Side per theme |
| Rate-Path & Monetary | Policy-rate path | Side per claim |
| Pandemic / Public Health | Outbreak and public-health emergency | YES on emergence |
| Trade & Tariff | Tariff and trade actions | YES on imposition |
| Energy Transition | Energy and transition milestones | Side per claim |

# Appendix A: Symbol Glossary

| Symbol | Meaning |
|---|---|
| $\tau_m$ | Rebalance date *m* |
| $p_i(t)$ | Quoted YES price of claim *i* (market-implied probability) |
| $R_i$ | Terminal resolution of claim *i*, in $\{0,1\}$ |
| $d_i$ | Orientation / side variable, +1 (YES) or −1 (NO) |
| $x_i(t)$ | Side-adjusted price (oriented mark) |
| $x_i^*(t)$ | Lifecycle-adjusted mark (payoff if resolved) |
| $Y_i$ | Oriented terminal payoff, in $\{0,1\}$ |
| $D_i(t)$ | Days remaining to expiry |
| $\lambda_i(t)$ | Implied daily event intensity (constant-hazard) |
| $x_i^H(t)$ | Horizon-*H* normalized probability |
| $M_i$ | Maturity-proximity score around target tenor $D^*$ |
| $E_{i,b}^+,\ E_{i,b}^-$ | Direct- and inverse-exposure evidence scores |
| $S_{i,b}$ | Total selection score of claim *i* for theme *b* |
| $P_i,\ T_i$ | Probability-band and tenor-band adjustments |
| $Q_i,\ R_i,\ Z_i$ | Quality term, short-dated penalty, saturation penalty |
| $G_{j,b}$ | Event-equivalence class *j* for theme *b* |
| $u_{i,b},\ u_{g,b}$ | Single-name cap and slot cap |
| $\hat{w}_{i,b,m},\ w_{i,b,m}$ | Preliminary and final (post-cash) claim weights |
| $c_{b,m}$ | Cash sleeve fraction for theme *b* at rebalance *m* |
| $\mathrm{Turn}_{b,m}$ | One-way portfolio turnover |
| $\kappa_b$ | Proportional transaction-cost coefficient |
| $n_{i,b}(t)$ | Units of oriented claim *i* held by theme *b* |
| $C_b(t),\ V_b(t)$ | Cash account and marked basket value |
| $R_b(t),\ L_b(t)$ | Daily basket return and inception-normalized index level |
| $\mathrm{Coverage}_b(t)$ | Weight-fraction of basket with observable marks |
| $\Omega_b,\ \bar{\Omega}_{b,t}$ | Strategic theme weight and renormalized active weight |
| $L_A(t)$ | Aggregate thematic index level |
| $\mathcal{F}_t$ | Information set available at time *t* |

# Appendix B: Formula Map

| Eq. | Name | Eq. | Name |
|---|---|---|---|
| (1) | Candidate set | (21) | Slot-cap constraint |
| (2) | Marked basket value | (22) | Slot-cap rescaling |
| (3) | No-lookahead condition | (23) | Raw pre-cap score |
| (4) | Admissibility indicator | (24) | Preliminary weights |
| (5) | Side variable | (25) | Single-name cap problem |
| (6) | Side-adjusted price | (26) | Cash-sleeve weights |
| (7) | Oriented terminal payoff | (27) | Dynamic cash rule |
| (8) | Single-claim return | (28) | One-way turnover |
| (9) | Implied event intensity | (29) | Transaction-cost model |
| (10) | Horizon-normalized probability | (30) | Target weights to units |
| (11) | Maturity-proximity score | (31) | Lifecycle-adjusted mark |
| (12) | Direct-exposure evidence | (32) | Forced-exit states |
| (13) | Inverse-exposure evidence | (33) | Theme observability |
| (14) | Side selection rule | (34) | Cash-account recursion |
| (15) | Total selection score | (35) | Daily basket value |
| (16) | Probability-band adjustment | (36) | Daily basket return |
| (17) | Tenor-band adjustment | (37) | Index-level recursion |
| (18) | Threshold rule | (38) | Renormalized theme weight |
| (19) | Event-equivalence class | (39) | Aggregate index level |
| (20) | Class representative | (40) | Aggregate observability |

# Changelog

| Version | Date | Changes |
|---|---|---|
| 1.0 | Q2 2026 | Initial publication. Mathematical methodology for thematic prediction-market baskets: admissible information set, oriented claims, selection and weighting, lifecycle marking, and the daily NAV recursion, with the aggregate thematic index. |

# Approvals

| Role | Name | Signature | Date |
|---|---|---|---|
| Head of Research | | | |
| Chief Executive Officer | | | |