One Kernel, Three Payoffs

A Payoff-Agnostic, Gate-First Architecture for Prediction-Market, Perpetual & Defined-Risk-Option Calibration — Q7 AEGIS AI S7

Haneef R. Haqq, MBA[1]

[1]Quant7 Alpha, LLC

Technical Whitepaper · v1.0 · Strategy S7 (Prediction Markets / Perpetuals / Options)

June 30, 2026

Abstract

We describe S7, the prediction-market member of the Q7 AEGIS AI family, which operates leveraged event and perpetual contracts and defined-risk options across three venues (Kalshi, a regulated options broker, and a decentralized prediction venue). S7 departs from its sibling strategies in two ways. First, it is *payoff-agnostic*: a single execution kernel sizes and risk-governs three distinct payoff classes—stop-bounded leveraged perpetuals, full-premium long options, and binary [0,1] event contracts—through a common per-trade risk unit, so the same capital-protection law applies whether maximum loss is set by a protective stop, an option premium, or a contract that settles against the position. Second, and more consequentially, S7 is *gate-first* rather than signal-first: unlike the directional Q7 strategies, which posit a microstructural regime edge a priori, S7 assumes no tradeable directional edge and admits a configuration only when it clears a fail-closed structural-integrity cornerstone, a multiple-testing-corrected charter gate, and an out-of-sample protocol evaluated under a fair sign-preserving permutation null. We detail the payoff-agnostic risk unit, the pinned capital-protection spine (a per-trade clamp, a soft entry-block, and a non-latching hard-flatten backstop held under the venue auto-liquidation line), the train-equals-serve hybrid-AI gate with a distribution-anchored reachable threshold, the event/correlation cap, and the calibration acceptance criteria. We then report—in the interest of intellectual honesty—the current state of the edge search: on the 364-day underlying that S7 expresses through the Kalshi perpetual, a weak but non-null directional signal survives net of the venue's 6× cap, fees, and funding (strongest on ETH, profit factor 1.21, fair-null $p = 0.035$), yet *no* configuration clears the program's multiple-testing-corrected gate—including under the full calibrated, HAI-gated, out-of-sample test, where ETH and XRP hold a persistent out-of-sample profit factor near 1.9 but miss the deflated-Sharpe and Bonferroni bars on thin samples, while BTC and SOL overfit and collapse. On the native Kalshi instrument, only ~27 days old, the 90-day out-of-sample floor is not yet meetable, so confirmation there is pending data, not refuted. The maturing instrument is the one remaining open confirmation. We state the architecture's limitations explicitly, and no realized-performance claims are made.

1. Edge Thesis

The Q7 family's one independently validated edge (S2) is not a price-prediction edge: it is a *structural econometric relationship*—a DV01-hedged Treasury spread whose cointegration tethers it to a stationary process, so fundamentals *force* reversion—admitted only by a cointegration-health gate that stands aside when the relationship breaks. The directional members (S1, S3–S5) make, by design, *no realized-performance claim*; they ship an architecture with mechanistic priors and a battery of calibration gates. S7 takes the strict reading of that evidentiary posture. Its premise is **conditional and falsifiable**: *if* a structural mispricing exists in an event or perpetual contract—a no-arbitrage tether between complementary or cross-venue contracts, or a durable information asymmetry—the harness will discover, calibrate, and admit it; *absent* such structure, the harness ships nothing. The thesis is therefore not "S7 predicts direction" but "S7 deploys capital only on configurations that survive a fail-closed gate, and the gate is adversarial to search-induced artifacts." This is the honest analogue of S2's discipline applied to a venue class where, as Section 10 documents, the directional signals that the futures-complex strategies rely on do not reproduce.

2. Architecture

S7 is organized in three independently testable layers, identical in shape to the family reference but re-tailored to event/perp/option payoffs.

Signal. Per (asset, horizon), a producer emits a candidate `(direction, confidence)` each bar—a momentum-confluence construction or the hybrid-AI overlay—and the kernel is deliberately signal-agnostic: it consumes the resolved signal and never generates one, so the producer and the risk kernel are separately testable.

Intelligence. Subsystems run in parallel: the confluence/conviction scorer with a gate-reachability guard (Section 5); the hybrid-AI overlay (Section 6); the intra-trade exit manager (Section 7); and the event/correlation cap (Section 8).

Execution. Signal → confidence gate → hybrid-AI gate → payoff-agnostic sizing under the per-trade clamp → entry, with the soft entry-block and the non-latching hard-flatten backstop enforced every bar. The kernel imports *only* the pinned risk spine: every Q7 strategy owns its execution core, and S7 borrows no sibling's kernel.

Venue	Payoff class	Max-loss unit	Role
Kalshi	Leveraged perpetual	stake × stop-distance	BTC·ETH·SOL·XRP perps
Options broker	Long defined-risk option	premium (full stake)	S1–S6 underlyings
Prediction venue	Binary YES/NO event	stake (settles 0/1)	event contracts

Each (asset, horizon) configuration is calibrated independently; venue and payoff class fix the *risk unit*, not an operating point.

3. The Payoff-Agnostic Risk Unit

S7's distinctive engineering choice is that one kernel risk-governs three payoff classes whose maximum loss is set by entirely different mechanisms. The unifying abstraction is the *fraction of stake at risk*, $f \in (0,1]$: for a long option or a binary YES, the full stake can be lost ($f = 1$); for a stop-bounded perpetual, only the stop (or liquidation) distance is at risk ($f =$ stop fraction). Sizing then solves a single clamp—the largest stake whose *defined* maximum loss does not exceed the per-trade ceiling:

$$\text{stake} \le (P_{\text{trade}} \cdot Q) \,/\, f, \quad P_{\text{trade}} = 2\% \text{ of equity.}$$

The construction is fail-safe by default: an unknown or malformed instrument collapses to the most conservative full-loss assumption ($f = 1$), so a mis-specified payoff can only *reduce* exposure, never inflate it. This is the S7 counterpart to the order-flow leverage map of the directional strategies—but where those scale *up* on conviction, S7's primary sizing act is to bound *down* to a defined loss, because on binary and premium payoffs there is no protective-stop continuum to lean on.

4. The Capital-Protection Risk Spine

Risk is enforced preventively at the sizing layer, not reactively at a kill-switch, because an event contract can gap discretely at resolution and a perpetual can gap through a stop. The spine is a pinned capital ladder, identical in invariant to S1–S6 and adapted only in its per-trade unit:

- **2% per-trade clamp** (Section 3): defined max loss per position ≤ 2% of equity.
- **20% soft entry-block**: at or above a 20% drawdown from the working equity peak, no *new* positions open; existing positions manage to their stops.
- **23% non-latching hard-flatten backstop**: an intrabar mark-to-market on *unrealized* drawdown (at the bid/liquidation price, never mid) force-closes all positions at 23%—two points under the venue auto-liquidation line—then *re-arms* by resetting the working peak to post-flatten equity.
- **25% venue auto-flatten line**: the broker/venue line the backstop exists to keep us under; it is never reached by construction.

The backstop is deliberately a *circuit breaker, not a latch*. A permanent latch would convert a transient excursion into a dead account and would mask a weak-edge configuration that should instead be rejected by the drawdown pass-gate. By re-arming on a working-peak reset, the spine holds the realized ceiling under the venue line on every excursion while letting the calibration gate—not a one-way lock—adjudicate whether the configuration deserves capital. Every constant in the ladder is **pinned and non-calibratable**; a safety cap is never a search dimension, and a Stage-1 gate fails closed if any drawdown or ruin constant leaks into the search space—searchability is precisely how such a cap once drifted off its pin in an earlier strategy.

5. Signal Construction and the Confluence Gate

The reference producer is a momentum-confluence construction on the perpetual leg: a directional candidate with a 0–100 conviction score, admitted only when conviction clears a per-configuration floor. The architecture imposes one discipline absent from naive gating—a *gate-reachability guard*. A confidence (or model-probability) cutoff is admissible only if the producer's own observed score distribution actually reaches it; a floor the signal can never clear silently starves the configuration to zero trades while presenting as a clean "no-signal" result. S7 resolves every cutoff from the producer's realized distribution (a percentile of observed conviction), and a structurally unreachable cutoff is a calibration *failure*, surfaced loudly at structural review, not a quiet pass. Mean-reversion (z-score) and basis constructions are available to the producer under the same gate, but—as Section 10 reports—none has cleared the out-of-sample bar.

6. The Hybrid-AI Gate (Train = Serve)

S7 filters producer entries through the same multi-modal overlay the family uses—a directional learner, a financial-language sentiment channel, and a market-regime label—fused in the spirit of the reference hybrid architecture [9] but adapted

to per-configuration, train-equals-serve operation. The decisive engineering invariant is that the overlay *calibrated* is the overlay *deployed*: each model is trained in-cycle on the configuration's own data, and the content hash that gates live trades must equal the hash calibrated in-cycle. A configuration without a fresh, health-passing, hash-matched model is *blocked*, never traded gate-off. The model label is derived from the strategy's own self-identifier (`SCRIPT_FORK = "S7"`), so the pre-flight self-sufficiency and coherence gates discover this core as model-bearing and verify it against the on-disk artifacts.

6.1 Directional model — gradient-boosted trees (XGBoost)

The directional learner is an XGBoost classifier predicting whether the forward move is favorable—an additive ensemble of shallow trees fit by stagewise gradient boosting on a regularized logistic objective, outputting a calibrated probability. Following the reference recipe [9] (binary next-period-direction labels over a compact technical feature set, shallow trees, low learning rate) but disciplined to S7's setting: the model is trained in-cycle (no stale prior-cycle model) and admitted only above a per-direction discriminability floor and under an adversarial-leakage bound. Critically, its entry threshold is *anchored to the model's own observed probability distribution*, not a hardcoded constant. This is the direct fix for the unsatisfiable-gate failure class: a fixed probability cutoff a base-rate-calibrated model can never reach starves every configuration to zero trades—a defect, not a strategy result. The reference paper [9] exhibits exactly this anti-pattern (a hardcoded score≥2 entry and a fixed sentiment cutoff, with no calibration to the model's distribution); S7 deliberately does not inherit it.

6.2 Sentiment model — FinBERT

The sentiment channel is FinBERT [10], a finance-domain transformer mapping each news item to a polarity over {positive, neutral, negative}, aggregated per asset per day into a normalized score consumed strictly as a *risk filter and soft threshold skew, never a standalone entry signal*. For event contracts the channel is doubly motivated, since many contracts settle on the very macro prints (CPI, FOMC, employment) that the sentiment veto would suspend entries around. *Honesty note:* absent a time-aligned, point-in-time news feed for an evaluated window, the channel is run neutral under a signed no-feed exemption (honest-empty) rather than fabricating sentiment; the exemption is recorded in the cycle dossier, and train=serve is preserved by deploying the same neutral configuration that was calibrated.

6.3 The tandem and the regime channel

The third channel is a market-regime label conditioning both the gate and the exit envelope. The three combine as a tandem —the directional model proposes a probability, the sentiment channel can veto or attenuate it, and the regime label selects the operating envelope—the multi-modal fusion that [9] reports outperforms any isolated predictor across regime changes. All three are computed in-cycle from the same data the system trades, so the gate cannot be a flattering in-sample artifact of a model the live system never runs. We are explicit that the reference paper's empirical claim—a single uncosted two-year backtest with no fee model, no significance test, and no walk-forward—justifies the gate's *structure*, not any deployment edge; S7 imports the architecture, not the result.

7. Real Trade Management: Intra-Trade Exits and the Backstop

Each position is managed by real intra-trade exits, not merely by signal reversal. From entry, the signed return is tracked each bar: a move to `−stop fraction` triggers a stop-loss (which pays modeled adverse slippage, because stops are not guaranteed fills), and a move to `+take-profit fraction` triggers a limit take-profit (no slippage). These make the stop and take-profit distances *live strategy dimensions* rather than cosmetic parameters, and a real protective stop is what keeps aggregate drawdown under the venue line on adverse paths. The 23% non-latching backstop (Section 4) sits *below* the protective stop as a last-resort control for a gap-through that the stop could not fill; forced backstop flattens also pay slippage. The cost model amortizes venue funding (capped at the venue's ±2%/8h hard cap as a worst-case drag) into per-trade economics by hold duration, plus forced-exit slippage.

8. Event / Correlation Cap

A control with no analogue in the single-instrument strategies: event contracts that share a settlement event are not independent, so a per-position clamp is insufficient. S7 caps aggregate stake-at-risk across positions sharing a settlement event or dependency-graph parent at a fixed fraction of equity, *on top of* the per-position clamp. A new position is blocked if it would push its shared-event cluster over the cap; an unknown event is treated, fail-safe, as its own singleton cluster (still capped). Cluster membership is populated from a dependency-graph / relationship-discovery layer—the construct that prevents several nominally distinct contracts on one underlying event from compounding into a single concentrated bet.

9. Calibration and Acceptance

S7 is calibrated under the program's precision pipeline, and the cornerstone is structural, not statistical. Before any search, a fail-closed Stage-0 establishes specification–calibration–execution equivalence at the formula level (engine liveness, searched-dimension consumption, attribution correctness, formula-footprint parity, runtime ship-smoke); any FAIL or UNKNOWN renders the pipeline RED and all downstream calibration advisory. Pre-flight self-sufficiency, hybrid-AI co-

herence, and ship-smoke gates must return CLEAR. Search optimizes every admissible dimension jointly under a sequential model-based optimizer with a fixed, disclosed evaluation budget, against a **drawdown-budget objective** that rewards *utilization* of the drawdown budget rather than variance minimization—preventing the degenerate collapse to near-zero exposure a naive ratio objective induces. The objective is calibrated on the hybrid-AI-gated equity curve, never the raw-signal curve. Acceptance is a simultaneous multi-criterion charter gate—win rate, gross profit factor, trade frequency, drawdown (soft, logged), and a hard commission-to-gross bound—evaluated out-of-sample on purged combinatorial cross-validation with embargo, where the *lower confidence bound* of the path distribution gates, not the point estimate, and the in-sample optimum is discounted by a deflated performance statistic to correct for selection across the disclosed trial count. A configuration that clears in-sample but not under a *fair sign-preserving permutation null* (timing and fees preserved, only direction randomized) is rejected; the broken full-array-shuffle null, which manufactures a fee-churn artifact and a spuriously significant p-value, is explicitly not used. Promotion additionally requires consecutive consistent cycles, specification-to-implementation parity within tolerance, and an independent re-derivation under separate credentials.

10. Empirical Findings: The Crypto-Perp Signal Search

Consistent with the program's standard of reporting evidence plainly, we record the current state of the S7 edge search, distinguishing two surfaces. **(a) The live Kalshi perpetual instrument** is, as of this writing, only ~27 days old (CFTC-approved and launched in June 2026), so the program's 90-day history floor—required for combinatorial-purged cross-validation to span regime diversity—is *physically unmeetable* on the instrument itself until it ages (~Q3 2026). Confirmation on the native instrument is therefore *pending data*, not refuted. **(b) The 364-day underlying** (BTC, ETH, SOL, XRP), the price surface S7 expresses through the Kalshi perpetual, does clear the floor; evaluated net of Kalshi economics (the 6× CFTC-regulated leverage cap, post-promotional taker fees, and 8-hour funding) under the fair sign-preserving null of Section 9, an uncalibrated momentum-confluence producer exhibits a *weak but non-null* directional signal—profit factors of 1.05–1.21 net of fees, strongest on ETH (PF 1.21; all four purged CV folds ≥ 1.0; fair-null $p = 0.035$, significant uncorrected). Decisively, however, **no asset clears the multiple-testing-corrected gate** (Bonferroni $\alpha = 0.0125$ across the four-asset family), and the funding-carry, delta-neutral-basis, and cross-sectional-pair classes remain null (a representative BTC pair-reversion with in-sample PF 1.36 collapsed out-of-sample to 0.957). Separately, prior testing of event-market niches (conditional weather, sports totals) found those markets at or near their irreducible information floor. We have since run the decisive test—the *calibrated, HAI-gated* (train-equals-serve) system, TPE-optimized on a 60% in-sample partition and evaluated on an untouched 40% out-of-sample holdout, scored under a deflated Sharpe ratio that discounts the selection across trials. It does *not* clear the gate. Two assets (BTC, SOL) overfit and collapse out-of-sample (in-sample profit factor ~1.6 → out-of-sample ~0.75, deflated Sharpe ≈ 0). Two assets (ETH, XRP) retain a persistent out-of-sample signal—profit factor ~1.9 with three of four purged folds well above 1.0—but fail the corrected bar on all three counts: the fair-null p (0.046 / 0.068) misses the four-asset Bonferroni threshold (0.0125), the purged-CV lower bound falls below 1.0, and the deflated Sharpe (0.64 / 0.78) misses the ~0.95 acceptance level, with thin samples (58–75 trades) the proximate cause. The honest verdict is therefore precise: ETH and XRP are genuine *candidates*—a real, persistent directional signal that survives the Kalshi fee and leverage structure out-of-sample—but not *confirmed* edges at the program's adversarial rigor, and BTC/SOL are overfit artifacts. This closes the first of the two confirmations negatively; the second—whether the signal reproduces on the native Kalshi instrument once it reaches the 90-day floor (~Q3 2026)—remains open. Until a configuration clears the full gate, S7 stays gate-first: it ships nothing on an *unconfirmed* signal, which is a fail-closed default, not a verdict that no edge can exist.

11. Risk and Venue

S7 spans three venues with materially different microstructure and settlement. The perpetual leg (Kalshi) carries leverage and funding and is governed by the stop-bounded risk unit and the funding-amortized cost model. The options leg (a regulated broker) is defined-risk: maximum loss is the premium, so the per-trade clamp sizes against the full stake. The binary leg (a decentralized prediction venue) settles 0/1 at resolution with no protective-stop continuum, so sizing is full-stake-at-risk and the event/correlation cap is load-bearing. Across all three, the spine constants are structural and implemented identically in specification, calibrator, and executor; the 23% non-latching backstop marks to the bid/liquidation price, holding the realized ceiling under the venue auto-flatten line on every excursion. Commissions and venue fees enter per-trade economics conservatively, and the commission-to-gross hard bound is decisive on these venues, where the fee wall is where most candidate edges die.

12. Limitations

The architecture's integrity does not manufacture an edge. Section 10 is the central limitation: as of this writing S7 has *no confirmed* directional, carry, or relative-value edge that clears the program's multiple-testing-corrected gate, and the whitepaper claims none. What it does have is a persistent but

sub-threshold out-of-sample signal on two of four underlyings (ETH and XRP, out-of-sample PF ~1.9) that misses the deflated-Sharpe and multiple-testing-corrected bars on thin samples, alongside two that overfit (BTC, SOL), plus a native instrument too young (~27 days) to validate. The calibrated HAI-gated confirmation has been run and did not clear; a confirmed claim now awaits the instrument reaching the 90-day floor and the candidate signal reproducing out-of-sample on the native contract with an adequate trade count. The out-of-sample protocol bounds in-sample selection bias but does not immunize against structural breaks after deployment; out-of-sample here means out-of-sample *within the certified history*. The preventive sizing bounds drawdown against the worst gap in the instrument's history, and a future gap exceeding that envelope is a tail the construction reduces but does not eliminate. Train=serve and health gates constrain the hybrid-AI overlay but do not certify that any learned relationship is causal or durable; the sentiment channel, absent a point-in-time feed, runs neutral and contributes no signal in such windows. The one structural analogue to S2's edge in this venue class—a no-arbitrage tether between complementary or cross-venue contracts—was itself refuted for retail in prior testing and is treated as an efficiency play owned by faster, better-capitalized participants, not a present claim. Execution, leverage, liquidity, and venue/regulatory risks are not eliminated by architectural integrity. As with every Q7 deliverable, no realized-performance representation is made, and the acceptance criteria above are calibration gates, not outcomes.

References

[1] R. Cont, A. Kukanov, S. Stoikov. *The Price Impact of Order Book Events.* 2011.

[2] R. Cont, A. de Larrard. *Price Dynamics in a Markovian Limit Order Market.* 2011.

[3] M. Gould et al. *Limit Order Books* (survey). Quantitative Finance, 2013.

[4] M. López de Prado. *Advances in Financial Machine Learning* (purged combinatorial cross-validation, embargo). Wiley, 2018.

[5] D. Bailey, M. López de Prado. *The Deflated Sharpe Ratio: Correcting for Selection Bias, Backtest Overfitting, and Non-Normality.* Journal of Portfolio Management, 2014.

[6] R. A. Fisher. *The Design of Experiments* (permutation/randomization inference). 1935.

[7] J. Wolfers, E. Zitzewitz. *Prediction Markets.* Journal of Economic Perspectives, 2004.

[8] E. F. Fama. *Efficient Capital Markets.* Journal of Finance, 1970.

[9] V. N. K. Pillai, A. Ajith, S. K. J. *Generating Alpha: A Hybrid AI-Driven Trading System Integrating Technical Analysis, Machine Learning and Financial Sentiment for Regime-Adaptive Equity Strategies.* ComSIA 2026 (Springer LNNS).

[10] D. Araci. *FinBERT: Financial Sentiment Analysis with Pre-trained Language Models.* 2019.
