

Config Watch #2

August 19, 2026 · MarketMania Research · Weekly series

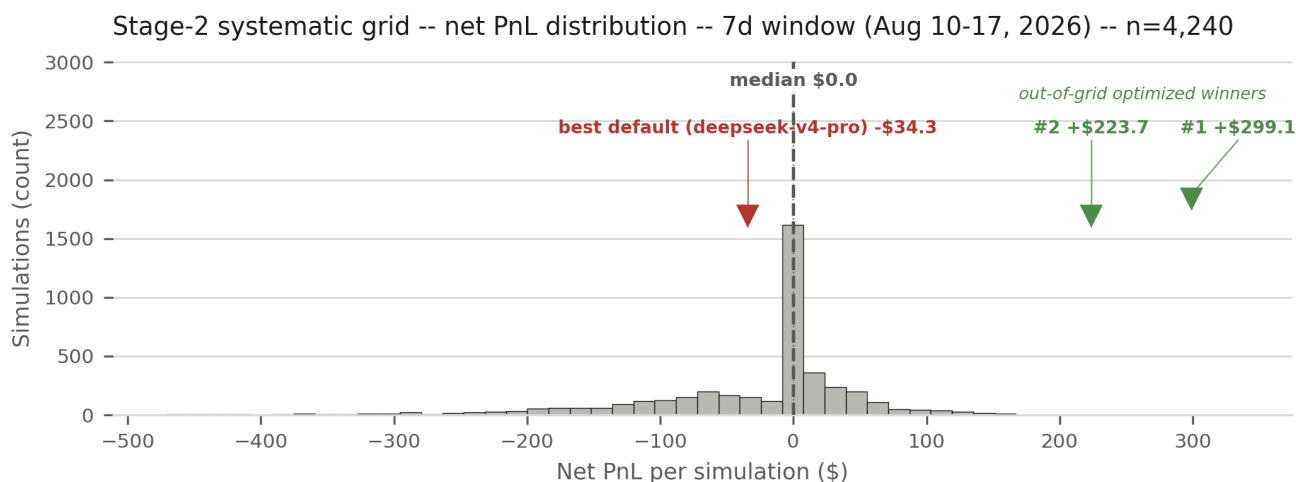
WEEKLY · CONFIG WATCH · Issue #2 · Issue date August 19, 2026 · Language English · Methodology v1.1 + Config Watch amendments (approved 12.08), hash e66c7e8c864a2233

Search windows: **7d** Aug 10-17, 2026 and **30d** Jul 18-Aug 17, 2026 (frozen date ranges). Windows now align to Mondays, as promised in issue #1 -- the 7d window IS the weekly reporting window the three sibling reports cover.

RESEARCH SNAPSHOT			
14,861 sims total (9,901 pipeline + 4,960 per-ticker)	7d branch grid 4,240 sims (Aug 10-17, 2026)	30d branch grid 4,240 sims (Jul 18-Aug 17, 2026)	6-stage search: scan -> grid -> refine -> per-ticker -> OOS -> dedicated per-coin
engine: \$1,000 deposit, 10x leverage	TOP-3 rule: ≥ 2 coins, distinct universes	methodology v1.1 + Config Watch amendments	Issue #2 -- first OOS verdict
market (report week): BTC net -3.08% (prior +2.09%)	ann. vol 10.0% (was 11.4%)	TOP5 volume \$7.75B , -8.9% w/w	pairwise corr 0.52 (was 0.40)
KEY FINDING · OBSERVATION (first OOS week)			
The first out-of-sample verdict: 10 of the 12 configs issue #1 published lost their edge on the fresh windows. Both survivors are 30d-tuned -- council 30d#2 (+\$421 on the new month) and the grok-4.5 solo 30ds#2 (+\$169).			

TL;DR

- OBSERVATION -- The first out-of-sample verdict is in, and it is harsh: 10 of the 12 configs issue #1 published lost their edge on the fresh windows. Both survivors are 30d-tuned: council 30d#2 (+\$421.0 on the new month) and solo 30ds#2, grok-4.5 (+\$168.9). Every 7d-tuned winner failed to carry.
- The issue-1 'robustness hint' family (fable-5+deepseek-v4-pro+opus-5 core) did not hold: its 7d#1 made +\$117.1 on the fresh week but -\$265.0 on the month, and the 30d#1/#3 pair went -\$107.3 / -\$167.8 on the fresh week. The hint is dead; the OOS decay series has its first data point.
- The fresh search still beats everything in-sample -- +\$299.1 (7d) and +\$744.7 (30d) vs. hand-tuned pre-search bests of +\$79.5 / +\$175.9 -- and the defaults still lose: grid medians \$0 (7d, 31.4% positive) and -\$159 (30d, 13.7% positive); all 7 solo defaults negative on both windows again.
- The dedicated per-ticker search debuts (added same-day on owner review: 4,960 sims, an independent 448-cell grid per coin-window): SOL was the month's goldmine -- a fable-5+deepseek-v4-pro+grok-4.5 council made +\$698.3 on 50 trades -- XRP's best solo made +\$543.7, and BTC was unextractable even at the top of its grid (best: -\$2.7). The calibrate axis helped 84% of the main grid's matched pairs (mean +\$244.1) yet hurt the per-ticker top-3 twins (7% wins on the month) -- the middle-vs-tail effect, live.



Net PnL across all 4,240 stage-2 grid sims, 7d window (Aug 10-17, 2026). Grid median \$0.0, 31.4% positive. The spike at \$0 is mostly cells whose entry filters produced no trades (1,279 zero-trade sims on this window). Red marker = best solo default; green = the published councils (refinement stage, beyond this grid).

Why it matters

MarketMania publishes default LLM-council trading configs. This series asks a narrower, more useful question every week: how much apparent performance can large-scale in-sample search extract -- and how much of it survives out-of-sample? Issue #1 established the in-sample baseline. Issue #2 delivers the series' reason to exist: the first OOS verdict on the 12 configs issue #1 published, plus a fresh search on new, Monday-aligned windows.

How we searched

Five stages, run per window (7d and 30d branches, council and solo in one pipeline this issue):

- **1. Scan (s1)** -- 952 sims across council compositions and coarse settings.
- **2. Systematic grid (s2)** -- 8,480 sims (4,240 per window: 2,728 council + 1,512 solo cells), sweeping declared axes -- archetype, FH/TF, entry filters, universe, council size, membership -- and, new this issue, a **calibrate** axis (TP/SL calibration on/off) on advice cells. This grid is the denominator for every percentile claim.
- **3. Refinement (s2b+s3a+s3b)** -- 269 sims: ladder extension, seed flips and coordinate moves around the grid winners, producing the published cards.
- **4. Per-ticker probes (s5t)** -- 120 sims: the best finalists split onto single coins, both calibrate settings.
- **5. OOS replays (s4)** -- 80 sims: every issue-1 published config and four prior-wave seed families, replayed verbatim on both fresh windows, with and without calibration.
- **6. Dedicated per-ticker search (w112, added same-day on owner review)** -- 4,960 sims: an INDEPENDENT grid per coin (448 cells per coin-window: 112 solo + 336 councils covering every pair and triple of the 7 models), then one-knob refinement and calibrate twins on each cell's top-3. Zero errors.

Stage design per the Config Watch amendments (approved 12.08). Fixed economics (deposit, leverage, stake table, fees, liquidation modeling) are constant across every sim in every stage.

Models under test (exact engine IDs): claude-fable-5, claude-opus-5, gpt-5.6-sol, deepseek-v4-pro, gemini-3.1-pro, qwen-3.8-max, grok-4.5.

TOP councils -- 7d window (Aug 10-17, 2026)

#	Council	Coins	FH/TF	Net	Net %	maxDD	Trades	WR
#1	fable-5, opus-5, deepseek-v4-pro, gemini-3.1-pro	SOL+XRP	4h/1h	+\$299.1	+29.9%	0.0%	12	100%
#2	gpt-5.6-sol, deepseek-v4-pro, gemini-3.1-pro, qwen-3.8-max	BNB+XRP	1d/1d	+\$223.7	+22.4%	3.6%	9 *	57%
#	Ladder	BE	Shifts	Conf filter	Source	Entry	Positional	Stake
#1	50:40, 100:60	on	TP -10 / SL +75	>=60 (each)	farthest	2/0	reenter	\$450
#2	50:40, 100:60	on	TP -10 / SL +75	>=60 (each)	farthest	3/0	reenter	\$450

All are refinement-stage results -- better than 100.0% / 100.0% of the 4,240-config stage-2 grid on this window (the refinement stage searches beyond that grid). Grid median \$0.0.

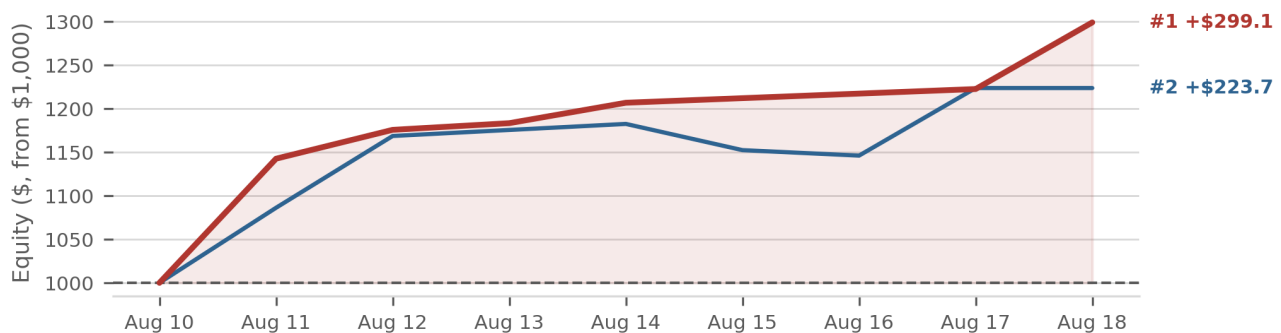
Net % = net PnL as a percentage of the fixed \$1,000 starting deposit; every simulation runs at **10x leverage** (liquidation modeled by the engine). * = fewer than 10 trades (thin).

Only two councils are published on this window: the branch finalists collapsed onto a single universe (SOL+XRP holds the top 9 of 10 slots), and the TOP-3 rule requires pairwise different universes. #2 (BNB+XRP) is the best finalist outside that universe -- 9 trades, flagged thin.

OPEN IN SANDBOX marketmania.ai/s/cw2-7d1 marketmania.ai/s/cw2-7d2

Each link opens this exact frozen window; results visible without sign-in (embedded share signature).

Published councils equity -- 7d (Aug 10-17, 2026) -- \$1,000 at 10x -- 21 settled trades



Daily settled equity for the councils above (\$1,000 start, 10x, frozen window Aug 10-17, 2026); ranks match the table (#1 red, #2 blue, #3 amber -- fixed rank colors across all issues). Sandbox-widget style; a curve past the window end is an open-at-cutoff trade settling at its horizon.

TOP-3 -- 30d window (Jul 18-Aug 17, 2026)

#	Council	Coins	FH/TF	Net	Net %	maxDD	Trades	WR
#1	opus-5, gpt-5.6-sol, deepseek-v4-pro, gemini-3.1-pro, qwen-3.8-max	BTC+ETH+SOL +BNB+XRP	1d/1d	+\$744.7	+74.5%	6.7%	25	73%
#2	opus-5, deepseek-v4-pro, gemini-3.1-pro	ETH+SOL+BNB +XRP	1d/1d	+\$655.1	+65.5%	17.3%	31	68%
#3	opus-5, deepseek-v4-pro, gemini-3.1-pro	BNB+XRP	1d/1d	+\$489.1	+48.9%	26.1%	23	67%
#	Ladder	BE	Shifts	Conf filter	Source	Entry	Positional	Stake
#1	50:40, 100:60	on	TP -10 / SL +75	>=60 (each)	farthest	3/1	reenter	\$450
#2	50:40, 100:60	on	TP -10 / SL +75	>=60 (each)	farthest	2/0	reenter	\$450
#3	50:40, 100:60	on	TP -10 / SL +75	>=60 (each)	farthest	2/0	reenter	\$450

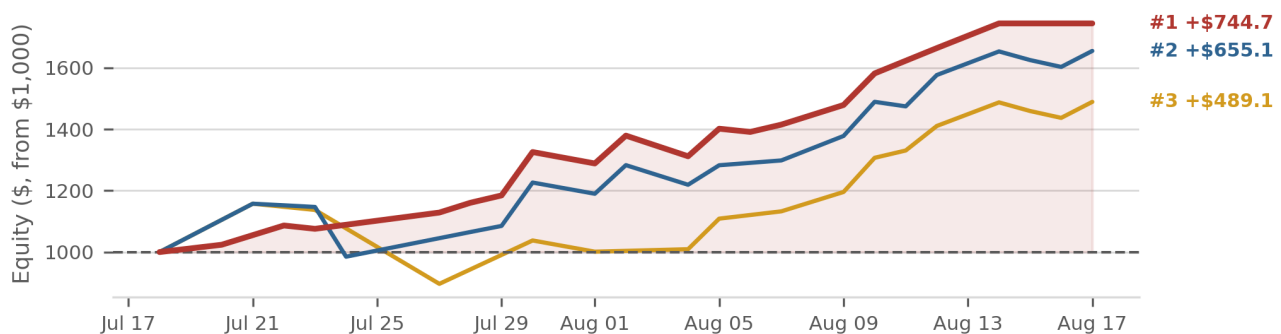
All are refinement-stage results -- better than 100.0% / 100.0% / 100.0% of the 4,240-config stage-2 grid on this window (the refinement stage searches beyond that grid). Grid median -\$159.0.

Net % = net PnL as a percentage of the fixed \$1,000 starting deposit; every simulation runs at **10x leverage** (liquidation modeled by the engine). * = fewer than 10 trades (thin).

OPEN IN SANDBOX marketmania.ai/s/cw2-30d1 marketmania.ai/s/cw2-30d2 marketmania.ai/s/cw2-30d3

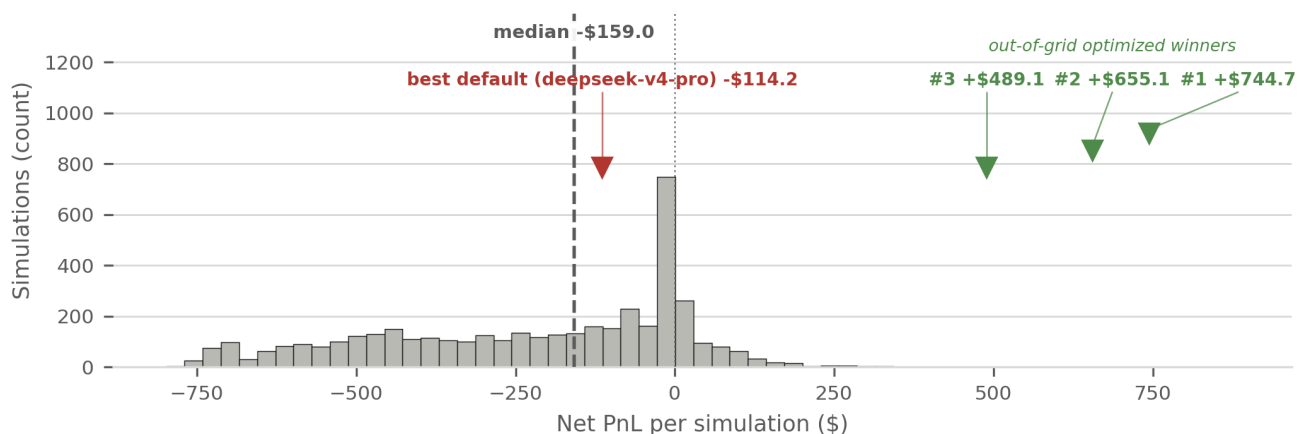
Each link opens this exact frozen window; results visible without sign-in (embedded share signature).

Published councils equity -- 30d (Jul 18-Aug 17, 2026) -- \$1,000 at 10x -- 79 settled trades



Daily settled equity for the councils above (\$1,000 start, 10x, frozen window Jul 18-Aug 17, 2026); ranks match the table (#1 red, #2 blue, #3 amber -- fixed rank colors across all issues). Sandbox-widget style; a curve past the window end is an open-at-cutoff trade settling at its horizon.

Stage-2 systematic grid -- net PnL distribution -- 30d window (Jul 18-Aug 17, 2026) -- n=4,240



Net PnL across all 4,240 stage-2 grid sims, 30d window (Jul 18-Aug 17, 2026). Grid median -\$159.0, only 13.7% positive -- the harder window again. The \$0 spike is 604 zero-trade cells. Red = best solo default; green = TOP-3.

Walk-forward: the first OOS verdict on issue #1

Issue #1 promised that the out-of-sample fate of its published configs becomes a standing section. Here it is: all 12 issue-1 configs (six council TOP-3, six solo top), replayed VERBATIM -- same knobs, no re-tuning, calibration off -- on both fresh windows.

Config	Family	Coins	Issue #1 (home)	New 7d	New 30d	Verdict
7d#1	fable-5, deepseek-v4-pro, opus-5, gpt-5.6-sol	SOL+XRP	+\$439.2	+\$117.1 (19)	-\$265.0 (92)	Faded
7d#2	fable-5, qwen-3.8-max, gemini-3.1-pro, opus-5	XRP+BNB	+\$429.5	+\$167.9 (16)	-\$581.9 (17)	Faded
7d#3	gpt-5.6-sol, opus-5	BNB+XRP+SOL	+\$398.6	+\$65.8 (26)	-\$428.9 (70)	Faded
30d#1	fable-5, deepseek-v4-pro, opus-5, qwen-3.8-max	ETH+SOL	+\$868.4	-\$107.3 (10)	+\$192.2 (44)	Faded
30d#2	gemini-3.1-pro, gpt-5.6-sol, deepseek-v4-pro	SOL+XRP+ETH	+\$590.6	+\$59.8 (4) *	+\$421.0 (25)	Survived
30d#3	fable-5, deepseek-v4-pro, opus-5, qwen-3.8-max	BTC+ETH+SOL	+\$580.5	-\$167.8 (14)	+\$24.4 (61)	Faded
7ds#1	gemini-3.1-pro	XRP+BNB	+\$417.8	-\$134.9 (33)	-\$473.0 (81)	Faded
7ds#2	gpt-5.6-sol	XRP+BNB	+\$366.3	+\$4.6 (28)	-\$370.8 (67)	Faded
7ds#3	qwen-3.8-max	XRP+BNB	+\$267.0	-\$78.6 (25)	-\$605.7 (19)	Faded
30ds#1	opus-5	SOL+XRP	+\$442.8	-\$8.4 (14)	-\$292.1 (44)	Faded
30ds#2	grok-4.5	SOL+XRP	+\$262.9	+\$15.0 (12)	+\$168.9 (46)	Survived
30ds#3	deepseek-v4-pro	XRP+BNB	+\$119.0	-\$77.4 (24)	-\$372.8 (37)	Faded

n in parentheses; * = fewer than 10 trades (insufficient cell). Window overlaps: the new 7d window shares 2 of 7 days with the old 7d window and sits inside the old 30d window; the new 30d window shares 26 of its 31 days with the old 30d window. This table is therefore a first decay read, not pure out-of-sample -- purity improves every week as the windows roll forward.

Honest read: the decay is broad and fast -- 10 of 12 configs are negative on at least one fresh window, and every 7d-tuned config (all six) failed to carry. The two survivors are both 30d-tuned: **30d#2**

(gemini-3.1-pro+gpt-5.6-sol+deepseek-v4-pro, 1d/1d, SOL+XRP+ETH) at +\$421.0 on the fresh month (its fresh-week read is n=4 -- insufficient), and **30ds#2** (grok-4.5 solo, 4h/1h, SOL+XRP) at +\$168.9 / +\$15.0. Issue #1's 'robustness hint' family (fable-5+deepseek-v4-pro+opus-5 core) did not hold anywhere cleanly. One corroboration worth naming: 30d#2 is the same 3-model 1d family as pre-search hand-tuned config A (the 1d_3_llm_v02 preset -- a tuned variant of this family), which also stayed positive on both fresh windows (+\$34.0 * / +\$175.9, Baselines below) -- three independent positive reads on one family, each individually thin.

Calibrate axis: does TP/SL calibration help a config?

New axis this issue. For matched config pairs -- identical knobs, calibration OFF vs ON (v1 model multipliers learned from forecast history) -- the search measured the net-PnL delta on each window.

Window	Pairs	Calibrate wins	Mean delta	Best delta	Worst delta
7d	4	75%	+\$154.6	+\$323.4	-\$327.6
30d	396	84%	+\$244.1	+\$636.1	-\$639.9

On the month, calibration improved 84% of 396 pairs with a +\$244.1 mean delta -- a strong broad-population effect. The honest tension: **every published champion above runs calibrate OFF** -- at the very top of the distribution, raw configs still edged out their calibrated twins this issue. Read as: calibration lifts the middle of the population, not (yet) the extreme tail. The 7d window produced only 4 matched pairs (the axis ran on advice cells) -- no conclusion from n=4.

Secondary analysis -- solo-model top

Sidebar to the main council narrative. Solo configs ran inside the same pipeline this issue (1,512 grid cells per window plus refinement). Same publication gates as issue #1: ≥ 2 coins; ≥ 15 trades on 7d, ≥ 20 on 30d; one config per model.

7d: no qualifier. No 7d solo finalist cleared the gates this week -- the best solo by net (gpt-5.6-sol, BNB+XRP, 1d/1d, +\$245.0) closed 11 trades, under the ≥ 15 gate, and the next finalists are single-coin. Reported unranked for completeness, not as a card.

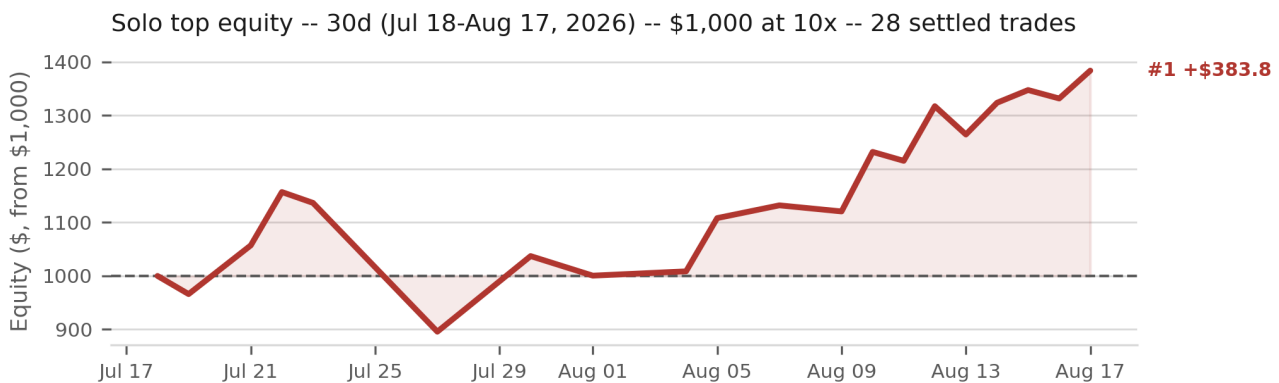
Solo top -- 30d window (Jul 18-Aug 17, 2026)

#	Model	Coins	FH/TF	Net	Net %	maxDD	Trades	WR
#1	gemini-3.1-pro	BNB+XRP	1d/1d	+\$383.8	+38.4%	26.1%	28	65%
#	Ladder	BE	Shifts	Conf filter	Source	Entry	Positional	Stake
#1	50:40, 100:60	on	TP -10 / SL +75	≥ 60 (each)	median	0/0	reenter	\$450

Only one 30d solo cleared the gates (gemini-3.1-pro; the next finalists repeat the same model or run single-coin -- grok-4.5 x SOL alone at +\$337.4 appears in Per-ticker bests below). Net % on the fixed \$1,000 deposit at **10x leverage**.

OPEN IN SANDBOX marketmania.ai/s/cw2-30ds1

Each link opens this exact frozen window; results visible without sign-in (embedded share signature).



Daily settled equity for the solo model above (\$1,000 start, 10x, frozen window Jul 18-Aug 17, 2026); ranks match the table (#1 red, #2 blue, #3 amber -- fixed rank colors across all issues). Sandbox-widget style; a curve past the window end is an open-at-cutoff trade settling at its horizon.

Council vs. solo: on 30d the best solo (+\$383.8) sits below the council TOP-3 (+\$744.7 / +\$655.1 / +\$489.1); on 7d no solo qualified while two councils did. The ensemble edge held at the top of both windows this issue. Same in-sample caveats apply.

Per-ticker bests: a dedicated single-coin search (new standing branch)

Added same-day after owner review: the pipeline's finalist-split probe (s5t, 120 sims) was not a real per-coin search, so each coin got its own -- an independent 448-cell grid per coin-window (112 solo cells: 7 models x 4 FH/TF x conf x sideways; 336 council cells: every pair and triple of the 7 models x 3 FH/TF x sideways), plus one-knob refinement and calibrate twins on each cell's top-3. 4,960 sims, zero errors, same fixed economics. Single-coin results remain excluded from the TOP-3 by rule (idiosyncratic-coin risk) -- this section answers a narrower question: which coin was extractable, and by whom.

7d (Aug 10-17, 2026)

Coin	Best council	FH/TF	Net (trades)	Best solo	Net (trades)
BTC	opus, gpt, qwen	4h/1h	+\$137.0 (8) *	fable	+\$69.1 (15)
ETH	fable, grok, qwen [cal]	4h/4h	+\$18.6 (5) *	qwen [cal]	+\$23.1 (1) *
SOL	fable, deepseek, grok	4h/1h	+\$260.4 (9) *	fable	+\$230.0 (8) *
BNB	deepseek, gemini, grok	1d/1d	+\$92.1 (2) *	qwen	+\$43.9 (3) *
XRP	deepseek, gpt	1d/1d	+\$268.2 (6) *	grok	+\$287.1 (7) *

30d (Jul 18-Aug 17, 2026)

Coin	Best council	FH/TF	Net (trades)	Best solo	Net (trades)
BTC	fable, deepseek, grok	4h/4h	-\$2.7 (50)	fable	-\$94.0 (58)
ETH	gpt, grok, qwen [cal]	1d/1d	+\$235.9 (7) *	fable	+\$68.6 (7) *
SOL	fable, deepseek, grok	4h/1h	+\$698.3 (50)	grok	+\$337.4 (39)
BNB	deepseek, gemini, grok	1d/1d	+\$342.8 (6) *	gemini	+\$211.4 (14)
XRP	deepseek, gemini	1d/1d	+\$357.6 (15)	opus	+\$543.7 (16)

n in parentheses; * = fewer than 10 trades (thin); [cal] = the winning cell is a calibrate=ON twin. The reads: **SOL was the month's goldmine** -- the fable-5+deepseek-v4-pro+grok-4.5 council made +\$698.3 on 50 trades (a grid cell, not even a refined one), double what the finalist-split probe found on the same coin (+\$337.4). **XRP's best solo** (claude-opus-5, 1d, single rung) made +\$543.7. **BTC was unextractable**: the best of its 448-cell month grid finished at -\$2.7, the best solo at -\$94.0 -- no configuration made BTC pay this month. ETH's month winner is a calibrate twin (+\$235.9) -- the only place in the issue where calibration produced a champion. The week side is thin nearly everywhere (2-9 trades) -- read as direction only.

The calibrate twins on per-ticker top-3s LOST on average (month: 7% wins, mean -\$187.5 across 30 pairs; week: 20% wins, -\$69.9) -- the opposite of the 84%-wins result on the main grid's full population. Direct, live evidence for the middle-vs-tail read in the Calibrate section -- with one honest caveat: top-3 twins carry winner's-curse selection, so a negative mean is partly expected from regression alone; read the direction, not the magnitude.

OPEN IN SANDBOX marketmania.ai/s/cw2-pt1 marketmania.ai/s/cw2-pt2

pt1 = the SOL month council (+\$698.3); pt2 = the XRP month solo (+\$543.7). Each link opens this exact frozen window; results visible without sign-in (embedded share signature).

Baselines

Solo-model and council-of-7 engine defaults, plus the three hand-tuned configs saved before the issue-1 search began (independent baselines, not derived from any search). All re-run fresh on this issue's windows.

7d (Aug 10-17, 2026)

Baseline	Net PnL	Trades	WR	maxDD	PF
Best solo default (deepseek-v4-pro)	-\$34.3	286	42%	3.4%	0.37
Worst solo default (grok-4.5)	-\$69.9	389	47%	7.0%	0.45
Council-of-7 default	-\$18.0	108	48%	1.8%	0.43
Hand-tuned config A (pre-search)	+\$34.0	3	100% (N=3)	0.0%	n/a
Hand-tuned config B (pre-search)	+\$79.5	31	77%	7.9%	1.44
Hand-tuned config C (pre-search)	-\$95.4	11	50%	11.8%	0.58

30d (Jul 18-Aug 17, 2026)

Baseline	Net PnL	Trades	WR	maxDD	PF
Best solo default (deepseek-v4-pro)	-\$114.2	1014	50%	12.0%	0.51
Worst solo default (grok-4.5)	-\$179.1	1441	55%	17.9%	0.60
Council-of-7 default	-\$43.2	437	56%	5.2%	0.66
Hand-tuned config A (pre-search)	+\$175.9	13	82%	10.7%	1.85
Hand-tuned config B (pre-search)	-\$114.8	89	67%	42.0%	0.89
Hand-tuned config C (pre-search)	+\$170.8	44	65%	23.0%	1.21

WR and PF greyed with (N=X) where fewer than 10 trades closed in the window -- insufficient sample; shown for completeness, not comparison.

Search validation: the fresh TOP beat the best hand-tuned pre-search config on both windows -- +\$299.1 vs +\$79.5 (7d); +\$744.7 vs +\$175.9 (30d). A small w/w shift: the council-of-7 default was the least-bad default on BOTH windows this issue (-\$18.0 / -\$43.2); in issue #1 it sat below the best solo default on both.

Patterns: what the finalists look like

Knob modes across each branch's unique finalists (top-10 council + top-10 solo per window, deduplicated). Issue #1 read this table as 'winning knobs differ by window'; this issue the OPPOSITE happened.

Knob	7d finalists mode	30d finalists mode
Forecast horizon	1d (11/20)	1d (16/20)
Ladder shape (steps)	50:40,100:60 (20/20)	50:40,100:60 (20/20)
Break-even stop (BE)	1 (20/20)	1 (20/20)
Min confidence	60 (20/20)	60 (20/20)
SL shift	75 (20/20)	75 (20/20)
TP/SL source	farthest (15/20)	farthest (15/20)

Both windows' finalists converged on the SAME core: 2-rung 50:40,100:60 ladder, break-even on, conf>=60, SL shift +75, farthest source (20/20 on four of six knobs, both branches) -- with 1d the modal horizon on both. Two honest caveats: the refinement stage seeds finalists from the same grid winners, so some convergence is search-design echo, not market signal; and a converged knob core that produced 10-of-12 OOS decay last issue is not a recommendation.

Liquidation note

Liquidation is modeled by the simulation engine at the fixed 10x leverage used throughout this search. Across the published configs, observed stop-loss placements stay well inside the distance that would approach the liquidation threshold. Max drawdown (maxDD) is reported on every config in this issue so realized risk is visible directly, rather than relying on this note alone.

Limitations

- **Multiple testing:** ~14,861 simulations were searched; with a search this large, some winners are expected by chance alone. Antidotes: the full stage-2 histograms (not just winners), declared axes, independent pre-search baselines, IS labeling, and -- new this issue -- the OOS walk-forward table itself. No formal multiple-testing correction yet (see What we're testing next).
- **Per-ticker calibrate read:** the -\$187.5 mean on the month rests on 30 top-3 twins only -- winner's-curse selection makes a negative mean partly expected from regression alone. Direction, not magnitude.
- **OOS purity:** the fresh windows overlap the issue-1 windows (2 of 7 days on 7d; 26 of 31 on 30d). The walk-forward table is a first decay read; purity improves weekly as windows roll.
- **Small trade counts:** several key cells run on 3-16 trades (30d#2's fresh-week read: 4; hand-tuned config A on 7d: 3) -- flagged and never silently pooled.
- **Single regime:** 6 of 7 days in the report week were flat; the month leans one way. No regime-change evidence yet.
- **Thin card set on 7d:** two council cards (universe collapse) and no qualifying solo -- a narrower publication than the rules were written for, disclosed rather than padded.
- Research output, not financial advice.

Counters & lineage

Stage	Sims	Stage	Sims
Scan (s1)	952	Per-ticker probes (s5t)	120
Systematic grid (s2)	8,480	OOS replays (s4 + s4cw1)	80
Refinement (s2b+s3a+s3b)	269	Pipeline subtotal	9,901
Dedicated per-ticker (w112, same-day)	4,960	-- its grid / refine / cal	4,480 / 420 / 60
Total	14,861	Errors / zero-trade	0 / 1,851

Engine 1.1 across every sim in every stage.

Scale note (vs issue #1): issue #1 ran 15,416 sims across two dedicated multi-day searches (council 12,199 + solo 3,217). Issue #2 runs ONE overnight pipeline -- 9,901 sims -- with three new axes folded in (calibrate, per-ticker, OOS replays) and a DENSER grid per window (4,240 vs 3,000 cells). The cut fell on the refinement waves (269 vs 8,204 sims), the stage issue #1's own OOS verdict just showed to be the most overfit-prone. The same-day dedicated per-ticker addition (4,960 sims) brings issue #2 to 14,861 -- near issue #1's scale with three more axes covered.

Generated at: 2026-08-19T04:45:23Z · Methodology v1.1 + Config Watch amendments (approved 12.08), hash e66c7e8c864a2233.

Snapshot principle: this report is a frozen snapshot of the search taken at the generated-at timestamp above; numbers are not updated retroactively. Each issue re-runs the full pipeline fresh over that week's windows.

RESEARCH TO DATE

THIS REPORT -- search effort: 14,861 simulations (9,901 pipeline + 4,960 dedicated per-ticker) · 30 verification sims (baselines + equity re-runs)

MARKETMANIA RESEARCH TO DATE (as of Aug 17, 2026 cutoff): 23,458 resolved forecasts since Jul 11 · 9 models tracked (7 current frontier + 2 archived legacy) · 5 assets · 5 forecast horizons · hourly cadence · 9 published research reports incl. this wave

Weekly counting rule (resolved = matured by the Monday cutoff), matching the three sibling weekly reports. The wider horizon-expired counter quoted in issue #1 (59,253) uses a different rule and was not re-probed this issue; the two are not comparable. Source: platform_counters_2026-08-17.

Issue #2. The walk-forward promise from issue #1 is now a standing section: every published config's OOS fate gets tracked weekly, decay and all. Windows are Monday-aligned from this issue on. Engine 1.1 powers the sandbox since Aug 18; this search ran on engine 1.1 throughout -- no cross-engine PnL comparisons are claimed.

What we're testing next

- Next issue: the OOS fate of THIS issue's cards -- and a second decay point for 30d#2 and 30ds#2. Two survivors of 12 is a base rate to beat, not a strategy.
- The per-ticker branch is now standing (owner decision): weekly per-coin winners get tracked OOS like the main cards -- starting with the +\$698.3 SOL council and the +\$543.7 XRP solo.
- Calibrate at the tail: why does calibration lift 84% of the population but not the champions? Candidate test: calibrated twins of the TOP cards over several weeks.
- Monthly: formal multiple-testing correction once several OOS windows accumulate.

Related research

Consensus Watch #2 -- <https://marketmania.ai/research/reports/consensus-watch-2026-08-10.pdf>

Weekly Calibration #2 -- <https://marketmania.ai/research/reports/weekly-calibration-2026-08-10.pdf>

Weekly Model Watch #2 -- <https://marketmania.ai/research/reports/model-watch-2026-08-10.pdf>

Config Watch #1 -- <https://marketmania.ai/research/reports/config-watch-2026-08-05.pdf>

The three weekly reports publish together as one issue each week; Config Watch follows on its own cycle. Direct links are the posting rule from this wave on.

CITE THIS REPORT

```
@techreport{mm_config_watch_2026w34,
  title      = {Config Watch #2},
  author     = {{MarketMania Research}},
  institution = {MarketMania},
  year       = {2026}, month = aug, day = {19},
  type       = {Weekly Research Report},
  series     = {Config Watch}, number = {2},
  note       = {Methodology v1.1 + Config Watch amendments, hash e66c7e8c864a2233},
  url        = {https://marketmania.ai/research/reports/config-watch-2026-08-12.pdf}
}
```