Config Watch #12

October 28, 2026 · MarketMania Research · Weekly series

WEEKLY · CONFIG WATCH · Issue #12 · the ceiling slips, and the board with it · Issue date October 28, 2026 · Language English · Methodology v1.1 + Config Watch amendments (approved 12.08, hash e66c7e8c864a2233)
Search windows: **WK** Oct 19-26, 2026 and **M30** Sep 28-Oct 26, 2026 -- both frozen and Monday-aligned (WK = Mon Oct 19 00:00 -> Sun Oct 25 23:59 UTC; M30 = Mon Sep 28 00:00 -> Sun Oct 25 23:59 UTC, 4 full weeks), sharing the cut-off edge Mon Oct 26 00:00 UTC, so slots dated Oct 26 fall outside both. Slots fall inside the window; trades settle past its right edge at their horizon -- the same 2-3 day luft the engine has used since issue #1 and that the live presets run on.

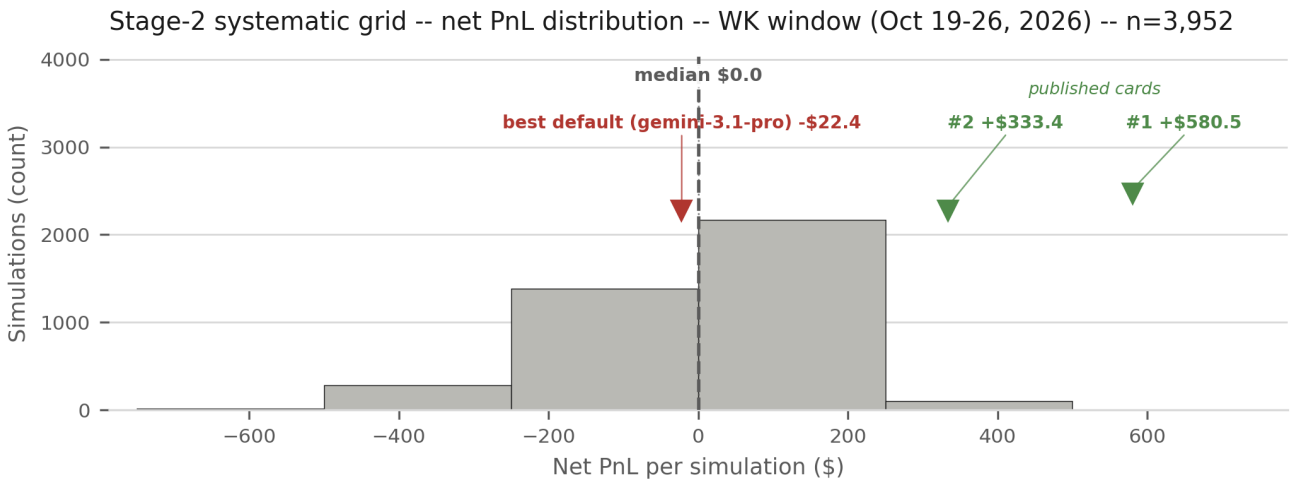
RESEARCH SNAPSHOT			
16,016 simulations, 0 errors, in two frozen collects	**WK** Oct 19-26, 2026	**M30** Sep 28-Oct 26, 2026	7-stage search: scan -> grid -> refine -> stake sweep -> per-ticker -> OOS replays
engine 1.1; **$1,000** deposit, **10x** leverage	stake is **universe-aware** from sim #1 (table below)	maturity gate: >=**15** closed (WK) / >=**20** (M30)	criteria frozen in writing before launch (w360 plan)
walk-forward: **74** past cards + **4** live presets replayed verbatim on both windows	**13 of 74** past cards Survive (net-positive on both windows)	stability: **1** config across WK and M30 (night run)	models: **7** (3 lineage axes in M30; WK on the new builds)
best WK council **+$580.5** (+58.1%)	best M30 council **+$880.9** (+88.1%)	zero-trade cells: **1,200** (weekly) / **2,414** (night run)	margin-skip alerts on leaders: **0** (WK) / **9** (M30) rows

KEY FINDING · OBSERVATION (the ceiling slips, and the board with it)

The ceiling of the fresh search moved from +$822.1 to +$580.5 on the week and 29 of the 74 cards this series has published are net-positive on that same week, and under the standing two-window rule 13 of the 74 Survive: 24 clear the month, 29 clear the week.

TL;DR

- OBSERVATION -- **The standing board grew to 74 cards.** Issue #11 kept 21 of its 69 past cards under the same rule; this issue replays **74 cards** -- issues #1, #2, #3, #4, #5, #6, #7, #8, #9, #10, #11, Monthly Config Watch #1 and #2 and Quarterly Config Watch #1 -- on two fresh windows, and **13 of the 74 Survive** (net-positive on both). On the week alone 29 of 74 are positive, on the four-week window 24 of 74. Issue #11's own week champion, replayed verbatim, makes +$20.7 on the week and +$546.0 on the month.
- **The search ceiling, window by window.** The best mature council on WK makes +$580.5 (+58.1% of a $1,000 deposit, 73.8% win rate, 5.6% maxDD, 42 closed) against issue #11's +$822.1; on M30 it makes +$880.9 (+88.1%, 73.5% win rate, 36.8% maxDD, 34 closed) against issue #11's +$1,484.4. The grids behind them: median $0.0 with 29.4% positive across 3,952 WK cells (2 of them above +$500), median $0.0 with 40.4% positive across 3,952 M30 cells (144 above +$500).
- **Calibration and the universe, window by window.** Across the whole stage-2 grid calibration improves 29.5% of the 1,336 matched WK pairs (mean -$24.0) and 38.5% of the 1,342 matched M30 pairs (mean -$12.3); pooled, 34.0% of 2,678. On the cards, **1 of the 2 published WK councils are calibrated cells** and **no M30 card is a calibrated cell**. The mature net top-10 leans on one universe per window: SOL+XRP on WK (9 of 10 slots), SOL+XRP on M30 (6 of 10).
- **The stability board, a ninth reading of the same pair.** The night collect computes WK and M30 in one pass, so this issue's cross-window board is measured on the two windows the issue prints -- and **1 configuration** is in the mature net top-40 of both. Issue #11 measured the same pair one week earlier and held 0, issue #10 the week before that 4, issue #9 7, issue #8 4, issue #7 16, issue #6 9, issue #5 0 and issue #3 21. The hand-tuned anchor keeps its standing measurement against the whole mature field of each window -- 211 of 1,882 mature WK configs and 1,637 of 2,809 mature M30 configs beat it on win rate, maxDD and smoothness at the same time. Net PnL stays the headline board (continuity with issues #1-#11).



Net PnL across all 3,952 stage-2 grid sims, WK window (Oct 19-26, 2026), universe-aware stake throughout. Grid median **$0.0**, 29.4% positive (issue #11's week: median +$61.2, 59.1% positive; its month: $0.0, 43.0% positive). The spike at $0 is mostly cells whose entry filters produced no trades (1,100 zero-trade sims on this window). Range -$645.2 to +$580.5; 2 cells finished above +$500. Red marker = best solo default at table stake; green = the published cards. The same board definition one week earlier put its top-5 at +$822.1 down to +$767.3; this week it runs +$580.5 down to +$451.4 -- a property of the week, not evidence about either issue's cards. The M30 histogram appears after the M30 cards. Source: grid_hist_wk.json.

Why it matters

MarketMania publishes default LLM-council trading configs. This series asks one narrower 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 set the in-sample baseline, issue #2 delivered the first walk-forward verdict (brutal), issue #3 the second (the opposite), issue #4 the third on two windows at once, issue #5 the fourth, issue #6 the fifth, issue #7 the sixth, issue #8 the seventh, issue #9 the eighth, issue #10 the ninth and issue #11 the tenth, each on the widest board the series had had. Issue #12 delivers the eleventh, on a board that carries every generation the series has produced: 74 cards from eleven weekly issues, the first two monthly ones and the first quarterly one, each replayed verbatim on the same two fresh windows. That is the point of the exercise -- a rule that is applied to a growing, never-pruned list is the only walk-forward number that cannot be chosen after the fact.

How we searched

Two frozen collects, run to the same plan with the criteria frozen in writing *before* launch (w360 plan) and printed into the meta of every JSON deliverable: the weekly collect over the WK window, and the night collect, which sweeps the M30 window and the same WK week together in one pass. Each stage below prints **weekly + night** and both totals are printed whole; no calendar-month window is computed this issue. **Every WK number in this issue comes from the weekly collect and every M30 number from the night collect** -- the night run's own second pass over the WK week is used for one thing only, the cross-window stability board, and no number from it is printed beside a weekly-collect number.

- **1. Scan (s1)** -- 476 + 952 sims across council compositions and coarse settings.
- **2. Systematic grid (s2)** -- 3,952 + 7,904 sims sweeping the declared axes (archetype, FH/TF, entry filters, universe, council size, membership, TP/SL source, ladder, break-even, TP/SL shift, calibrate and stake) -- the denominator for every distribution claim below.
- **3. Refinement: ladder + stake sweep (s2b+s3b)** -- 124 + 247 sims around the grid winners.
- **4. Preset neighbourhood grid (s2p)** -- 53 + 106 sims: one-knob neighbours of the live presets that have a local grid this issue.
- **5. Per-ticker probes (s5t)** -- 60 + 110 sims: best finalists split onto single coins.
- **6. Dedicated per-ticker (s6t+s6tc)** -- 564 + 1,110 sims: a compacted independent per-coin grid plus calibrate twins (the standing branch from issue #2).
- **7. OOS replays (s4cw11+s4cw11o+s4cw10+s4cw10o+s4cwq1+s4cwq1o+s4cw9+s4cw9o+s4cwm2+s4cwm2o+s4cw8+s4cw8o+s4cw7+s4cw7o+s4cw6+s4cw6o+s4cw5+s4cw5o+s4cw4+s4cw4o+s4cwm1+s4cwm1o+s4cw3+s4cw3o+s4cw2+s4cw2o+s4cw1+s4pre)** -- 140 + 218 sims: every published card and all four live

presets, verbatim, on every fresh window, plus origin re-runs of the issue-11 and issue-10 cards for the drift baseline.

Total 16,016 simulations (5,369 weekly + 10,647 night), 0 errors, 3,614 zero-trade cells (1,200 + 2,414); stage sums reconcile exactly in each collect (counts.by_stage). The axis grid was NOT widened relative to issue #11 (anti-overfit rule): the replay branch grew by one cohort, the search space did not. Equity series, grid histograms, baselines and the calibrate pack are produced inside the same two collects -- there is no separate same-day extras pass this issue.

Stake: universe-aware from simulation #1 ($1,000 deposit)

Instruments in universe	Stake cap	Sweep values also simulated
1 coin	$450	$300 / $375
2 coins	$450	$300 / $375
3 coins	$300	$225 / $375
4 coins	$250	$200 / $300
5 coins	$200	$150 / $250

A $1,000 deposit cannot fund five concurrent $450 positions. Stake is a function of how many instruments the config trades, recomputed on every universe change and swept as its own axis; any simulated stake above the cap for its universe size is marked **concentrated** and excluded from the headline boards (it stays in the open data). Every card also prints **skipped_nofunds** -- entries the engine could not fund. Margin-skip alert rows on the leaders this issue publishes: 0 on WK (weekly collect) and 9 on M30 (night collect); the night collect raised 9 alerts in total, all of them on M30. Source: criteria in both collects.

Models under test (engine IDs; claude-opus-5.5 is stored as claude-opus-5-5): claude-fable-5, claude-opus-5.5, gpt-6.1-sol, deepseek-v4-pro, gemini-3.1-pro, qwen-3.8-max, grok-4.7. Three axes run as lineages -- **claude-opus-5.5 (incl. opus-5-era), gpt-6.1-sol (incl. 5.6-era), grok-4.7 (incl. 4.6-era)** -- across the model cutover at Oct 17, 2026 09:01 UTC, which lies inside this issue's M30 window (day 20 of 28) and before its WK week: Oct 19-25 is the first full week on the new builds. Published configs that named claude-opus-5, gpt-5.6-sol, grok-4.6 or grok-4.5 are replayed out-of-sample with the same lineage remap ([remap] rows); the five issue-11 cards already name the heads. The vendors' names are Claude Opus 5.5, GPT-6.1 Sol and Grok 4.7.

TOP councils -- WK window (Oct 19-26, 2026)

#	Council	Coins	FH/TF	Net	Net %	maxDD	Trades	WR
#1	fable, gemini	SOL+XRP	4h/4h,1h	**+$580.5**	+58.1%	5.6%	42	73.8%
#2	gpt, gemini, qwen	BTC+ETH	4h/1h	**+$333.4**	+33.3%	0.7%	41	90.2%

#	Ladder	BE	Shifts	Conf filter	Source	Entry	Positional	Stake
#1	50:40,100:60	on	TP -10 / SL +75	>=60 (each)	farthest	3/0	reenter	$450
#2	50:40,100:60	on	TP -10 / SL +75	>=60 (each)	farthest	1/0	reenter	$450

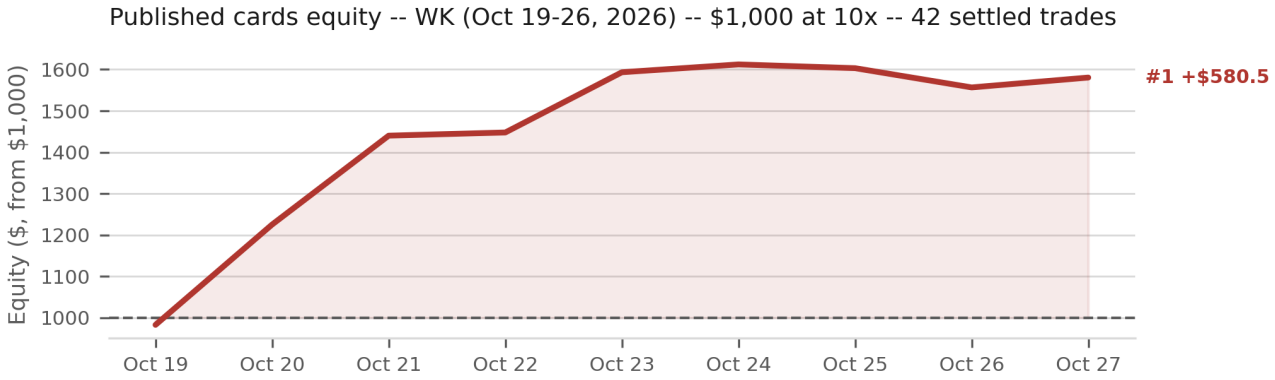
Against the 3,952-cell stage-2 grid on this window: card #1 lands in the +$500 to +$750 bucket (2 of 3,952 cells) and card #2 in the +$250 to +$500 bucket (100 cells); 2 cells in the whole grid finished above +$500 (grid median $0.0). Card #1 is the same simulation as the raw net-board leader (i=723), so its exported percentile is the card's -- above **100.0%** of the grid; the other exported card percentiles cover the raw net-board leaders and the replay rows, not card #2, for which no percentile is claimed. Net % = net PnL as a percentage of the fixed $1,000 starting deposit; every simulation runs at **10x leverage** (liquidation modeled by the engine). Entry = max_sideways / max_diff_side.

Only two councils are published on this window. 9 of the ten slots of the mature net board are the same 2-coin universe (SOL+XRP), so the pairwise-different-universes rule admits exactly one; #2 is the best mature multi-coin council outside that universe on the quality boards -- it comes off the **Calmar** board (+$333.4 on BTC+ETH, R2 0.972, ulcer 0.21, 0.7% maxDD, 0 unfunded entries) -- disclosed rather than padded. Read the cards as two universes deep, not three.

On the calibrate axis 1 of the 2 cards are calibrated cells, where issue #11 published 0 of its 2 WK councils as calibrated cells. Card #1 runs the 2-coin table stake of $450. Neither card carries an unfunded entry.

OPEN IN SANDBOX marketmania.ai/s/cw12-wk1 marketmania.ai/s/cw12-wk2

Each link opens this exact frozen window; results visible without sign-in (embedded share signature). The codes are minted by the publish script before this file goes live; every card's full frozen query signature is in the open-data JSON next to its simulation index.



Daily settled equity for card #1 above ($1,000 start, 10x leverage, frozen window; #1 red -- fixed rank colors across all issues). A curve running past the window end is an open-at-cutoff trade settling at its horizon (the 2-3 day luft); the curve's last point reconciles to its config's net to the cent. There is no contrast line on this window: the net-board maximum is the same simulation as card #1 (i=723, 42 closed, above the >=15 maturity gate), so a second curve would be the same curve. Card #2 has no daily series in this issue's equity pack, and the solo and smoothness series stay in the open data (equity.cards) -- the PDF carries one equity chart per window. Source: equity_wk.json.

TOP councils -- M30 window (Sep 28-Oct 26, 2026)

#	Council	Coins	FH/TF	Net	Net %	maxDD	Trades	WR
#1	opus, gpt	SOL+XRP	1d/1d	**+$880.9**	+88.1%	36.8%	34	73.5%
#2	gpt, deepseek	BNB+XRP	1d/1d,4h	**+$821.0**	+82.1%	21.4%	25	88.0%

#	Ladder	BE	Shifts	Conf filter	Source	Entry	Positional	Stake
#1	50:40,100:60	on	TP -10 / SL +75	>=60 (each)	farthest	1/0	reenter	$375
#2	50:40,100:60	on	TP -10 / SL +75	>=60 (each)	farthest	3/0	reenter	$450

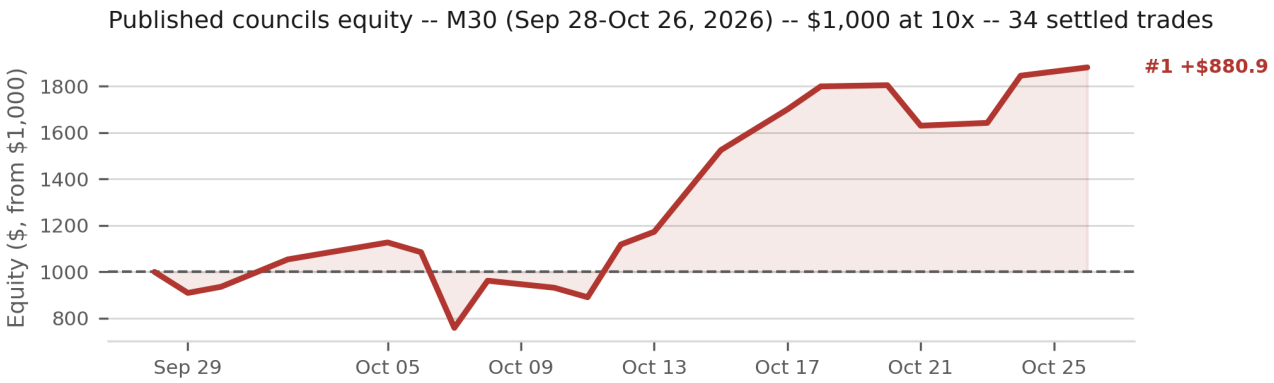
Against the 3,952-cell stage-2 grid on this window: card #1 is the same simulation as the raw net-board leader (i=9347), so its exported percentile is the card's -- above **99.9%** of the grid; card #2 lands in the +$750 to +$1,000 bucket (20 cells) with 99.5% of the grid finished below it. Grid median $0.0, 144 of the 3,952 cells above +$500. Net % = net PnL as a percentage of the fixed $1,000 starting deposit; every simulation runs at **10x leverage** (liquidation modeled by the engine). Entry = max_sideways / max_diff_side.

Two councils are published on this window too. The mature net board carries more than one universe: SOL+XRP holds 6 of its ten slots (card #1 among them) and BNB+XRP 4. Card #2 is the best mature multi-coin council outside card #1's universe (the pairwise-different-universes rule), and it comes off the mature net board itself (+$821.0 on BNB+XRP, R2 0.866, ulcer 4.90, 21.4% maxDD, 0 unfunded entries).

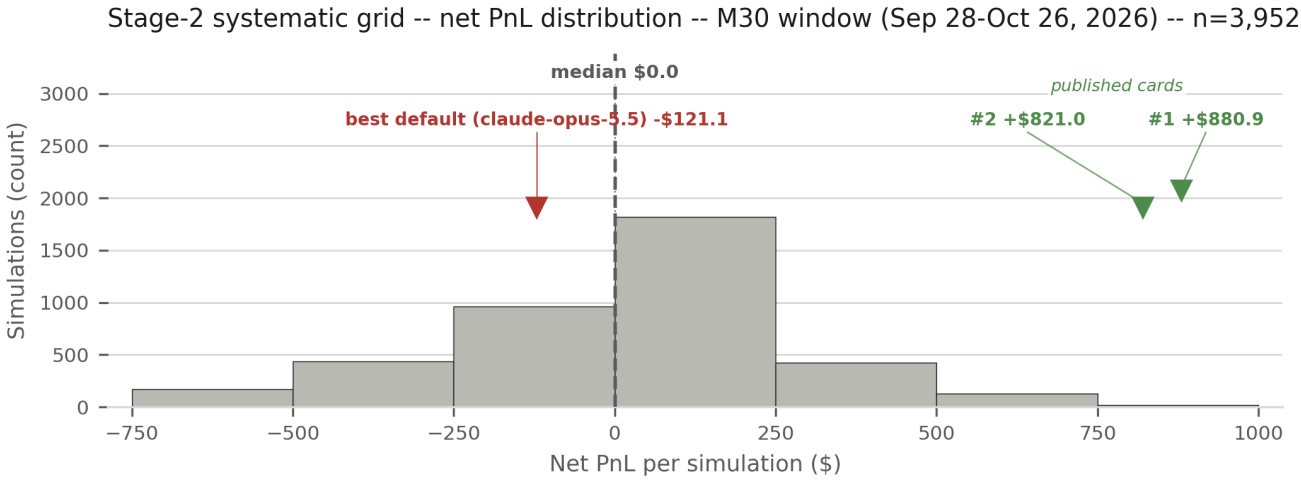
On the calibrate axis no M30 card is a calibrated cell, where 1 of the 2 published WK councils are calibrated cells -- the same axis, read on two windows (Calibrate axis below). Card #1 pays for its +$880.9 with a **36.8% drawdown** on 34 closed trades and 0 unfunded entries, card #2 with 21.4% on 25 and 0. The M30 mature pool is 2,809 configs against 1,882 on WK, because four weeks clear the >=20-trade gate that one week does not.

OPEN IN SANDBOX **marketmania.ai/s/cw12-m301** **marketmania.ai/s/cw12-m302**

Each link opens this exact frozen window; results visible without sign-in (embedded share signature). The codes are minted by the publish script before this file goes live; every card's full frozen query signature is in the open-data JSON next to its simulation index.



Daily settled equity for card #1 above ($1,000 start, 10x leverage, frozen window; #1 red -- fixed rank colors across all issues). A curve running past the window end is an open-at-cutoff trade settling at its horizon (the 2-3 day luft); its last point reconciles to the card's net to the cent. There is no contrast line on this window: the net-board maximum is the same simulation as card #1 (i=9347, 34 closed, above the >=20 maturity gate). The solo, Calmar, WR and smoothness series stay in the open data (equity.cards) -- the PDF carries one equity chart per window. Source: equity_night.json.

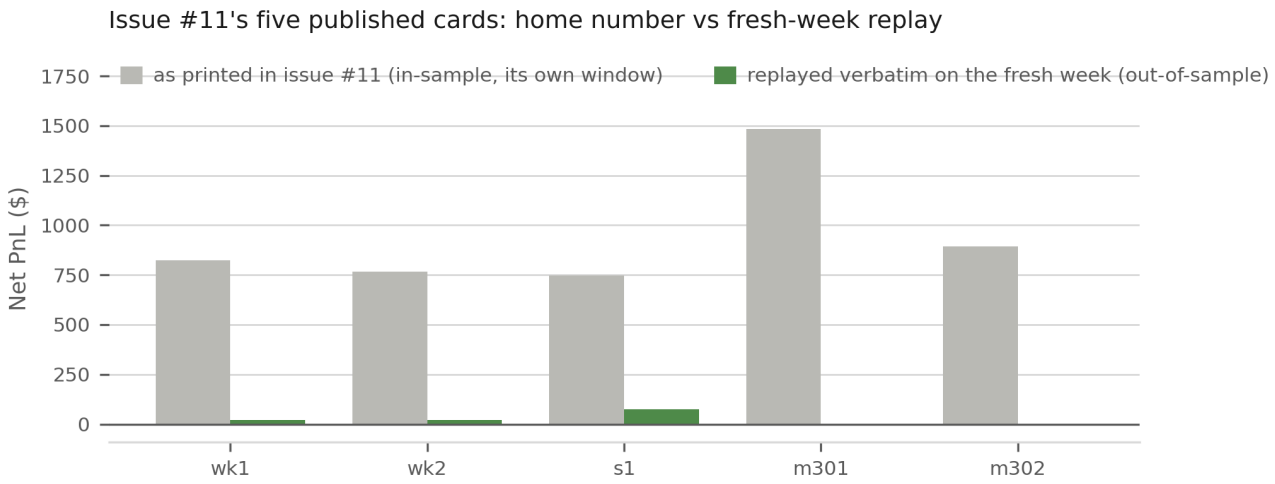


Net PnL across all 3,952 stage-2 grid sims, M30 window (Sep 28-Oct 26, 2026), universe-aware stake throughout. Grid median **$0.0**, 40.4% positive against 29.4% on the week (issue #11's month: median $0.0, 43.0% positive). The spike at $0 is mostly cells whose entry filters produced no trades (788 zero-trade sims on this window). Range -$669.0 to +$940.1; 144 cells finished above +$500, against 2 on the week. Red marker = best solo default at table stake; green = the published cards. Source: grid_hist_night.json.

Walk-forward: the standing OOS verdict

Every config this series has ever published -- all 74 cards from issues #1, #2, #3, #4, #5, #6, #7, #8, #9, #10, #11, Monthly Config Watch #1 and #2 and Quarterly Config Watch #1 -- replayed VERBATIM (same knobs, same stake as printed, no re-tuning) on both fresh windows, plus the four hand-tuned presets running live on the platform, listed first. The 9 cards of the two monthly issues and the quarterly one are judged by the weekly rule, like every other row. The replays and their origin re-runs cost 140 simulations in the weekly collect and 218 in the night collect, which replays every row on both of its own windows.

Verdict rule, unchanged since issue #2. Survived = net-positive on BOTH fresh windows; **Faded** = net-negative on at least one. The live presets, the monthly cards and the quarterly cards are judged by the same rule. New WK is pure out-of-sample for every row -- the week begins exactly at issue #11's right edge (Oct 19). New M30 (Sep 28-Oct 26, 2026) shares 21 of its 28 days with issue #11's own M30 window (shared stretch Sep 28-Oct 18), 14 of its 28 days with issue #10's own M30 window (shared stretch Sep 28-Oct 11), 7 of its 28 days with issue #9's own M30 window (shared stretch Sep 28-Oct 4), 3 of its 28 days with Monthly Config Watch #2's calendar-September window (shared stretch Sep 28-30) and 3 of its 28 days with Quarterly Config Watch #1's calendar-quarter window (shared stretch Sep 28-30), so for those five cohorts that column is decay context rather than pure out-of-sample: the same asymmetry issues #3, #4, #5, #6, #7, #8, #9, #10 and #11 disclosed for their own second columns.



Issue #11's 5 published cards: the number its own issue printed (grey, in-sample on its own window) against the same configuration replayed verbatim on this week (green, out-of-sample). None of the 5 is net-negative on the fresh week, and cw11-m302 closed no trade there (it stands at $0.0); the 2 M30 cards were measured at home on a four-week window and the 3 WK cards were on a one-week window, so the grey bars are not all the same kind of number. Source: oos_cw11_wk.json + issue #11 open data.

Config	Family	Coins	Home	New WK (n)	New M30 (n)	Verdict
preset 4h4	fable, qwen, deepseek, opus [remap]	ETH+SOL	-- live	-$223.7 (5) *	-$299.9 (27) +	Faded
preset bc2	gemini, gpt, deepseek [remap]	SOL+XRP	-- live	-$484.1 (4) * +	-$582.8 (14) +	Faded
preset sol	deepseek, fable, opus [remap]	SOL [conc]	-- live	-$197.1 (5) *	-$977.2 (40)	Faded
preset xrp	gemini, gpt, deepseek [remap]	XRP [conc]	-- live	-$427.5 (1) * +	-$279.0 (1) * +	Faded
cw11-m301	gpt, deepseek	SOL+XRP	+$1,484.4	+$0.6 (7) * +	+$735.9 (37) +	**Survived**
cw11-m302	fable, opus, gpt, gemini, qwen	BNB+XRP	+$894.5	$0.0 (0) *	+$717.3 (15)	Faded
cw11-s1	gpt	BTC+ETH+ SOL+BNB+XRP	+$747.1	+$73.7 (27) +	+$503.2 (84) +	**Survived**
cw11-wk1	opus, gpt, grok	SOL+XRP	+$822.1	+$20.7 (4) *	+$546.0 (50) +	**Survived**
cw11-wk2	gpt, grok	BTC+ETH+ SOL+BNB+XRP	+$767.3	+$22.3 (19)	+$239.7 (66) +	**Survived**
cw10-m301	fable, gpt, qwen [remap]	SOL+XRP	+$1,023.9	+$66.9 (16)	+$519.0 (55)	**Survived**
cw10-m302	fable, gpt, qwen [remap]	BNB+XRP	+$972.4	-$194.9 (18)	+$143.8 (50)	Faded
cw10-s1	gpt [remap]	SOL+XRP	+$415.2	-$56.9 (3) *	+$367.9 (60) +	Faded
cw10-wk1	fable, opus, gpt, gemini, qwen [remap]	SOL+XRP	+$528.8	-$139.1 (5) *	+$862.3 (77)	Faded
cw10-wk2	qwen, grok [remap]	BTC+ETH	+$176.5	+$121.7 (38)	+$228.8 (112)	**Survived**
cwq1-q1	fable, opus, gpt, gemini [remap]	SOL+XRP	+$2,271.8	-$22.2 (19) +	-$93.9 (107) +	Faded
cwq1-q2	fable, opus, deepseek, gemini [remap]	BNB+XRP	+$1,107.1	-$257.7 (5) *	+$145.3 (21)	Faded
cwq1-s1	gemini	SOL+XRP	+$1,491.0	+$314.0 (57)	+$97.3 (174) +	**Survived**
cw9-m301	fable, opus, gpt, qwen, grok [remap]	BNB+XRP	+$797.6	-$224.5 (8) *	-$297.7 (40) +	Faded
cw9-m302	fable, opus, qwen, grok [remap]	SOL+XRP	+$784.9	-$236.8 (7) *	-$532.2 (46) +	Faded

Config	Family	Coins	Home	New WK (n)	New M30 (n)	Verdict
cw9-s1	gemini	SOL+XRP	+$266.9	+$314.0 (57)	+$97.3 (174) +	**Survived**
cw9-wk1	deepseek, gemini, grok [remap]	SOL+XRP	+$250.4	-$201.5 (43) +	-$145.3 (128) +	Faded
cw9-wk2	opus, gpt, deepseek, gemini [remap]	BTC+ETH	+$216.9	-$130.1 (10) +	-$67.0 (81)	Faded
cwm2-m1	fable, opus, gemini, qwen [remap]	BNB+XRP	+$1,041.5	-$217.8 (9) *	-$395.3 (49) +	Faded
cwm2-m2	fable, opus, gemini, qwen [remap]	SOL+XRP	+$1,021.7	-$131.1 (8) *	-$547.2 (52) +	Faded
cwm2-s1	gemini	SOL+XRP	+$923.2	+$54.9 (57)	-$229.4 (157) +	Faded
cw8-m301	fable, opus, gemini, qwen [remap]	SOL+XRP	+$1,018.6	-$131.1 (8) *	-$547.2 (52) +	Faded
cw8-m302	fable, opus, gemini, qwen [remap]	BNB+XRP	+$943.0	-$217.8 (9) *	-$395.3 (49) +	Faded
cw8-s1	gemini	BNB+XRP	+$445.4	-$156.2 (51) +	-$562.6 (54) +	Faded
cw8-wk1	gpt, gemini, qwen [remap]	SOL+XRP	+$583.4	-$6.1 (24)	+$106.7 (81)	Faded
cw8-wk2	gpt, gemini, qwen, grok [remap]	BNB+XRP	+$570.6	-$253.8 (26) +	-$511.1 (69) +	Faded
cw7-m301	fable, gpt, gemini, qwen [remap]	SOL+XRP	+$1,351.8	+$123.6 (43) +	-$312.8 (99) +	Faded
cw7-m302	opus, gpt, gemini, qwen [remap]	BTC+ETH	+$528.4	+$9.3 (1) *	-$271.3 (13) +	Faded
cw7-s1	gemini	SOL+XRP	+$1,022.0	+$261.7 (57)	+$36.5 (188) +	**Survived**
cw7-wk1	fable, opus, gpt, gemini [remap]	SOL+XRP	+$1,467.8	-$22.2 (19) +	-$93.9 (107) +	Faded
cw7-wk2	fable, opus, gpt [remap]	BTC+ETH	+$953.2	-$60.7 (2) *	-$356.8 (38) +	Faded
cw6-m301	opus, deepseek, gemini [remap]	SOL+XRP	+$1,793.8	+$46.2 (17)	-$553.0 (93) +	Faded
cw6-m302	gpt, deepseek, gemini, grok [remap]	BTC+ETH	+$547.8	-$78.8 (5) * +	-$407.4 (27) +	Faded
cw6-s1	grok [remap]	SOL+XRP	+$404.8	-$202.6 (29) +	-$1.4 (87) +	Faded
cw6-wk1	opus, gpt, grok [remap]	SOL+XRP	+$609.5	-$145.7 (9) * +	-$559.8 (67) +	Faded
cw6-wk2	fable, opus, deepseek [remap]	BTC+ETH	+$135.4	+$20.2 (6) *	-$129.8 (60) +	Faded
cw5-m301	fable, opus, gpt, gemini [remap]	SOL+XRP	+$1,717.1	+$181.7 (3) *	+$204.8 (18) +	**Survived**
cw5-s1	gemini	BNB+XRP	+$178.7	+$361.5 (59)	-$264.9 (141) +	Faded
cw5-wk1	gpt, gemini [remap]	SOL+XRP	+$510.4	+$53.3 (56)	-$102.9 (181) +	Faded
cw5-wk2	fable, opus, gpt, gemini, grok [remap]	BTC+ETH	+$183.5	+$104.3 (19)	-$79.1 (78) +	Faded
cw4-m301	gpt, deepseek, gemini [remap]	SOL+XRP	+$1,676.0	-$633.5 (5) * +	-$639.9 (33) +	Faded
cw4-m302	fable, gpt, deepseek, gemini [remap]	BNB+XRP	+$1,155.2	-$230.4 (2) * +	-$558.2 (13) +	Faded
cw4-s1	deepseek	BTC+ETH+ SOL+BNB+XRP	+$325.0	+$84.4 (18)	+$516.2 (69)	**Survived**
cw4-wk1	opus, deepseek, gemini, qwen [remap]	BTC+ETH+ SOL+BNB+XRP	+$399.4	-$130.7 (7) *	-$200.0 (52) +	Faded

Config	Family	Coins	Home	New WK (n)	New M30 (n)	Verdict
cw4-wk2	fable, deepseek, qwen	BNB+XRP	+$128.2	-$193.9 (41)	-$443.0 (75) +	Faded
cwm1-m1	fable, gemini	SOL+XRP	+$1,618.4	-$396.6 (3) * +	+$45.6 (26) +	Faded
cwm1-m2	fable, gpt, deepseek, gemini, qwen [remap]	ETH+SOL+ BNB+XRP	+$756.6	-$197.5 (6) *	-$212.7 (37) +	Faded
cwm1-s1	gemini	SOL+XRP	+$1,721.9	-$64.8 (7) * +	-$300.7 (26) +	Faded
cw3-m301	opus, deepseek, gemini [remap]	SOL+XRP	+$2,137.6	+$306.7 (8) *	+$208.3 (69) +	**Survived**
cw3-m302	fable, opus, deepseek, gemini, qwen [remap]	BTC+ETH	+$412.3	$0.0 (0) *	+$14.7 (10)	Faded
cw3-s1	gemini	SOL+XRP	+$1,604.3	+$326.2 (49)	-$639.1 (26) +	Faded
cw3-s2	gemini	SOL+XRP	+$1,449.0	-$64.8 (7) * +	-$300.7 (26) +	Faded
cw3-wk1	opus, deepseek, gemini [remap]	SOL+XRP	+$2,175.1	+$306.7 (8) *	+$208.3 (69) +	**Survived**
cw3-wk2	opus, deepseek, gemini, grok [remap]	BTC+ETH	+$837.6	+$129.5 (8) *	-$550.6 (28) +	Faded
cw2-30d1	opus, gpt, deepseek, gemini, qwen [remap]	BTC+ETH+ SOL+BNB+XRP [conc]	+$744.7	-$73.3 (4) * +	-$35.7 (22) +	Faded
cw2-30d2	opus, deepseek, gemini [remap]	ETH+SOL+ BNB+XRP [conc]	+$655.1	-$508.6 (4) * +	-$27.8 (35) +	Faded
cw2-30d3	opus, deepseek, gemini [remap]	BNB+XRP	+$489.1	-$148.0 (4) *	+$13.7 (26) +	Faded
cw2-30ds1	gemini	BNB+XRP	+$383.8	-$159.7 (3) *	+$273.2 (25) +	Faded
cw2-7d1	fable, opus, deepseek, gemini [remap]	SOL+XRP	+$299.1	-$56.2 (8) *	-$586.8 (50) +	Faded
cw2-7d2	gpt, deepseek, gemini, qwen [remap]	BNB+XRP	+$223.7	-$217.9 (3) *	-$224.3 (18) +	Faded
cw2-pt1	fable, deepseek, grok [remap]	SOL	+$698.3	-$150.0 (19)	-$105.4 (56)	Faded
cw2-pt2	opus [remap]	XRP	+$543.7	$0.0 (0) *	+$76.3 (7) *	Faded
cw1-30d1	fable, deepseek, opus, qwen [remap]	ETH+SOL	+$868.4	-$184.3 (5) *	-$305.5 (25) +	Faded
cw1-30d2	gemini, gpt, deepseek [remap]	SOL+XRP+ETH	+$590.6	-$99.8 (9) * +	-$40.7 (39) +	Faded
cw1-30d3	fable, deepseek, opus, qwen [remap]	BTC+ETH+SOL	+$580.5	-$87.2 (6) *	-$368.3 (41) +	Faded
cw1-30ds1	opus [remap]	SOL+XRP	+$442.8	+$107.6 (31)	-$164.1 (67) +	Faded
cw1-30ds2	grok [remap]	SOL+XRP	+$262.9	+$44.9 (27)	-$273.2 (68) +	Faded
cw1-30ds3	deepseek	XRP+BNB	+$119.0	+$273.1 (38)	-$285.8 (85) +	Faded
cw1-7d1	fable, deepseek, opus, gpt [remap]	SOL+XRP	+$439.2	-$121.8 (12)	-$366.1 (73) +	Faded
cw1-7d2	fable, qwen, gemini, opus [remap]	XRP+BNB	+$429.5	-$90.6 (30)	-$557.6 (27) +	Faded
cw1-7d3	gpt, opus [remap]	BNB+XRP+SOL	+$398.6	-$48.3 (24)	-$720.7 (74) +	Faded
cw1-7ds1	gemini	XRP+BNB	+$417.8	+$229.5 (46)	-$558.0 (74) +	Faded
cw1-7ds2	gpt [remap]	XRP+BNB	+$366.3	+$6.4 (41)	-$284.3 (96) +	Faded

Config	Family	Coins	Home	New WK (n)	New M30 (n)	Verdict
cw1-7ds3	qwen	XRP+BNB	+$267.0	+$46.8 (29)	-$284.0 (79) +	Faded

n in parentheses; * = fewer than 10 closed trades (insufficient cell); + = non-zero skipped_nofunds (entries the engine could not fund -- the printed economics are then not the economics that ran). Home = the number the card's origin issue printed; the live presets have no home issue. [conc] = stake above this issue's cap for its universe size; [remap] = a superseded model id replayed on its lineage head (grok-4.5 / grok-4.6 -> grok-4.7, claude-opus-5 -> claude-opus-5.5, gpt-5.6-sol -> gpt-6.1-sol; none of the issue-11 rows carries it). Verdict rule: Survived = net-positive on both fresh windows, Faded = net-negative on at least one. Survivors are tinted green. The 74 cards are 68 distinct configurations window aside (cwq1-q1 = cw7-wk1; cwq1-s1 = cw9-s1; cwm2-m1 = cw8-m302; cwm2-m2 = cw8-m301; cwm1-s1 = cw3-s2; cw3-m301 = cw3-wk1): the two rows of each pair carry the same WK and M30 numbers.

Counted: 13 of 74 past cards **Survive** both fresh windows (issue #11 on its two windows: 21 of 69) -- by cohort 4 of 5 issue-11 cards, 2 of 5 issue-10 cards, 1 of 3 Quarterly-1 cards, 1 of 5 issue-9 cards, 0 of 3 Monthly-2 cards, 0 of 5 issue-8 cards, 1 of 5 issue-7 cards, 0 of 5 issue-6 cards, 1 of 4 issue-5 cards, 1 of 5 issue-4 cards, 0 of 3 Monthly-1 cards, 2 of 6 issue-3 cards, 0 of 8 issue-2 cards, 0 of 12 issue-1 cards. Taken a window at a time: 29 of 74 are positive on the week and 24 of 74 on the four weeks, and the two columns do not nest: a row can clear one and fail the other, so both windows remove cards. The survival curve by generation, each on its own first, second, third, fourth, fifth, sixth, seventh, eighth, ninth, tenth and eleventh fresh week, reads: issue-1 50.0% -> 91.7% -> 8.3% -> 16.7% -> 25.0% -> 91.7% -> 8.3% -> 8.3% -> 0.0% -> 58.3% -> **50.0%** now (6 of 12); issue-2 100.0% -> 50.0% -> 12.5% -> 25.0% -> 62.5% -> 75.0% -> 25.0% -> 25.0% -> 62.5% -> **0.0%** now (0 of 8); issue-3 33.3% -> 33.3% -> 33.3% -> 83.3% -> 33.3% -> 33.3% -> 16.7% -> 66.7% -> **66.7%** now (4 of 6); issue-4 20.0% -> 0.0% -> 80.0% -> 80.0% -> 60.0% -> 0.0% -> 80.0% -> **20.0%** now (1 of 5); Monthly-1 66.7% -> 0.0% -> 33.3% -> 100.0% -> 66.7% -> 0.0% -> 66.7% -> **0.0%** now (0 of 3); issue-5 25.0% -> 75.0% -> 50.0% -> 75.0% -> 0.0% -> 25.0% -> **100.0%** now (4 of 4); issue-6 80.0% -> 40.0% -> 60.0% -> 0.0% -> 80.0% -> **40.0%** now (2 of 5); issue-7 80.0% -> 80.0% -> 0.0% -> 40.0% -> **60.0%** now (3 of 5); issue-8 40.0% -> 0.0% -> 60.0% -> **0.0%** now (0 of 5); Monthly-2 100.0% -> 0.0% -> 0.0% -> **33.3%** now (1 of 3); issue-9 0.0% -> 60.0% -> **20.0%** now (1 of 5); Quarterly-1 33.3% -> 100.0% -> **33.3%** now (1 of 3); issue-10 80.0% -> **40.0%** now (2 of 5); issue-11 **80.0%** now (4 of 5). Live presets: 0 of 4 Survive, 0 of 4 positive on the week, 0 of 4 on the month.

The four live presets lead the table by standing rule. On the week: preset 4h4 -$223.7 on 5 closed, preset bc2 -$484.1 on 4 closed, preset sol -$197.1 on 5 closed, preset xrp -$427.5 on 1 closed -- 0 of the four net-positive. On the four-week window: preset 4h4 -$299.9 on 27 closed, preset bc2 -$582.8 on 14 closed, preset sol -$977.2 on 40 closed, preset xrp -$279.0 on 1 closed -- 0 of four. 0 Survive both. **sol** and **xrp** are marked [conc]: both run a $900 stake on a single coin, twice the $450 cap the stake table sets for a 1-coin universe, so their printed economics assume funding this issue's own rule would not grant. Passports for all four are in the open data (presets_recon_wk.json and presets_recon_night.json).

Origin re-runs (drift baseline). Re-simulating the issue-11 cards on their OWN origin windows today lands within -$139.7..$0.0 of the printed issue-11 numbers and the issue-10 cards within $0.0..$0.0 of theirs; 9 of the 10 re-runs reproduce to the cent. The widest deviation is cw11-m301 (+$1,484.4 printed, +$1,344.6 today) -- cards with positions still open when their origin window closed re-run to their settled result; issue #11 printed open positions for cw11-m301. Snapshot principle: no past issue is restated; the full drift table is in the open data (oos_walk_forward.drift_baseline).

Honest read. The fresh search and its own back catalogue moved in the same direction this week. On the fresh week 29 of 74 past cards are net-positive -- issue #11 counted 42 of 69 on its own week -- and the ceiling of the fresh search moved to +$580.5 from +$822.1. On the fresh four weeks 24 of 74 are net-positive. Same replay machinery, same frozen knobs; what changed is the window and the size of the board. Three things keep this from being a statement about search. The board is not a fixed sample: it grows every issue and nothing is ever pruned, so a share computed on 74 cards is not comparable with one computed on 69 without saying which generations joined -- 5 of the 74 rows are new this issue. The month column is not clean evidence for five of the fourteen cohorts: it shares 21 of its 28 days with issue #11's own M30 window (shared stretch Sep 28-Oct 18), 14 of its 28 days with issue #10's own M30 window (shared stretch Sep 28-Oct 11), 7 of its 28 days with issue #9's own M30 window (shared stretch Sep 28-Oct 4), 3 of its 28 days with Monthly Config Watch #2's calendar-September window (shared stretch Sep 28-30) and 3 of its 28 days with Quarterly Config Watch #1's calendar-quarter window (shared stretch Sep 28-30), so those cards are partly being scored on their own data. And the size of the in-sample number is not what decides the give-back on this board: the row that lost most on the week is cw4-m301 (-$633.5 on the week against +$1,676.0 printed at home), while the largest home number on the board, cwq1-q1 at +$2,271.8, made -$22.2. Eleven fresh weeks of verdicts now: brutal, the opposite, split by window, the fourth, the fifth, the sixth, the seventh, the eighth, the ninth, the tenth, and this one. That is a series of weathers, not a strategy.

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

Matched pairs -- identical knobs, calibration OFF vs ON -- measured on net PnL. The table is the **whole stage-2 grid** of each window, and the **Both** row pools the two windows' pairs and nothing else. The board-leader slice is a second, separate measurement (12 WK pairs and 14 M30 pairs whose OFF or ON leg reached the top of a board); it is reported in the paragraph below and in the open data, and the two are never pooled together.

Window	Pairs	Calibrate wins	Mean delta	Best delta	Worst delta
WK	1,336	29.5%	-$24.0	+$501.9	-$524.5
M30	1,342	38.5%	-$12.3	+$819.1	-$676.7
Both	**2,678**	**34.0%**	**-$18.1**	**+$819.1**	**-$676.7**

The two windows answer the axis, and the whole-grid reading and the board-leader slice are separate measurements. On the week calibration improves 29.5% of the 1,336 matched pairs (mean -$24.0, median -$2.3) and 28.1% of the 520 mature pairs (mean -$81.3); on the four weeks it improves 38.5% of 1,342 (mean -$12.3, median $0.0) and 50.3% of the 700 mature pairs (mean +$13.4). Pooled over both windows the axis improves 34.0% of 2,678 pairs with a mean of -$18.1 -- a number that describes neither window on its own. The board-leader slices are separate: 0.0% of the 12 WK slice pairs improve (mean -$383.3, median -$437.3, OFF legs already at a median of +$414.0), and 14.3% of the 14 M30 slice pairs (mean -$207.0, OFF legs at +$833.0). Issue #11 measured a 24-pair WK slice (1 improving) and a 13-pair M30 slice (0). Where the axis reached the cards is visible above: 1 of the 2 published WK councils are calibrate=ON cells and 30 of the 51 WK finalists run it, while 0 of the 2 M30 cards are and 18 of the 49 M30 finalists do.

The open-data pack also carries a 400-pair export (calibrate_effect) per collect. Its delta distribution is not the population's (median -$145.8 against the WK grid's -$2.3), so no claim in this section is computed from it; it is published for inspection only.

Secondary analysis -- solo-model top

Sidebar to the council narrative. Solo configs run inside the same pipeline and face the same gates: >=2 coins, the maturity gate for the window (>=15 closed on WK, >=20 on M30), one config per model.

Solo top -- WK (Oct 19-26, 2026)

#	Model	Coins	FH/TF	Net	Net %	maxDD	Trades	WR
#1	gemini	SOL+XRP	4h/4h,1h	**+$373.1**	+37.3%	22.5%	64	64.1%
#2	fable	BTC+ETH	4h/4h,1h	**+$355.5**	+35.6%	3.9%	43	79.1%

#	Ladder	BE	Shifts	Conf filter	Source	Entry	Positional	Stake
#1	50:40,100:60	on	TP -10 / SL +75	>=60 (each)	farthest	1/0	reenter	$450
#2	50:40,100:60	on	TP -10 / SL +75	>=60 (each)	median	1/0	reenter	$450

2 solo cards on this window. Both qualifying solos run the 2-coin table stake of $450. No single-coin row outranks them on the mature net board this issue; single-coin results are excluded from the cards by the >=2-coin rule either way and appear under Per-ticker bests. Source: summary_wk.json top_mature_by_window_branch.

Solo top -- M30 (Sep 28-Oct 26, 2026)

#	Model	Coins	FH/TF	Net	Net %	maxDD	Trades	WR
#1	gpt	SOL+XRP	1d/1d	**+$973.1**	+97.3%	37.2%	40	77.5%
#2	deepseek	SOL+XRP	1d/1d,4h	**+$817.0**	+81.7%	29.9%	30 *	76.7%

#	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	$375
#2	50:40,100:60	on	TP -10 / SL +75	>=60 (each)	median	1/0	reenter	$450

Two solo cards on M30. The mature solo board carries two models, gpt and deepseek, both on SOL+XRP, so the one-config-per-model rule admits two. Unfunded entries on the cards: #1 0, #2 1 (flagged * on Trades); the best row on the board with 0 skips is card #1 itself. Source: top_mature_by_window_branch in the night collect.

OPEN IN SANDBOX marketmania.ai/s/cw12-s1

Each link opens this exact frozen window; results visible without sign-in (embedded share signature). The codes are minted by the publish script before this file goes live; every card's full frozen query signature is in the open-data JSON next to its simulation index.

Council vs solo. On the week the best solo sits below the best council (+$373.1 vs +$580.5 on the mature board, 64 closed against 42, 22.5% drawdown against 5.6%). On the four weeks the best solo sits above the best council (+$973.1 vs +$880.9, 40 closed against 34, 37.2% against 36.8%). Same in-sample caveats apply to both sides.

Quality boards

Net PnL stays the primary board (continuity with issues #1-#11); four selection views sit on top of it at zero additional simulations. **Calmar-like** = net / max(maxDD, 1.0), positive net only; **WR board** = win rate among mature configs (>=15 closed on WK, >=20 on M30); **Smoothness** = R2 of a linear fit through the daily closed-equity series, mature rows with net > 0 and at least 5 days, ties broken by ulcer index; **Quality composite** = mean rank over net, Calmar-like, win rate and R2, lower is better. Top-3 of each board on each window below; all 15 rows of all eight boards are in the open-data JSON.

Board	Config	Coins	FH/TF	Net	WR	maxDD	Cls	R2
WK Calmar	gpt, gemini	SOL	4h/4h,1h	+$412.4	73.9%	0.0%	23	0.937
WK Calmar	gpt, gemini, qwen [cal]	BTC+ETH	4h/1h	+$333.4	90.2%	0.7%	41	0.972
WK Calmar	fable, gemini [cal]	BTC+ETH	4h/4h	+$321.2	71.9%	0.0%	64	0.962
WK WR	gemini, qwen, grok [cal]	ETH	4h/1h	+$170.9	100.0%	0.0%	16	0.965
WK WR	gemini, qwen, grok [cal]	BTC+ETH	4h/1h	+$319.2	97.1%	0.0%	35	0.970
WK WR	gemini, qwen, grok [cal]	BTC+ETH	4h/1h	+$266.0	97.1%	0.0%	35	0.970
WK Smoothness	fable, gpt, gemini, qwen [cal]	BTC+ETH	4h/1h	+$236.2	92.3%	0.0%	26	0.975
WK Smoothness	fable, gpt, gemini, qwen, grok [cal]	BTC+ETH	4h/1h	+$237.2	92.3%	0.0%	26	0.975
WK Smoothness	fable, gemini, grok [cal]	BTC+ETH	4h/4h	+$255.9	76.1%	0.0%	46	0.975
WK Quality	gpt, gemini, qwen [cal]	BTC+ETH	4h/1h	+$333.4	90.2%	0.7%	41	0.972
WK Quality	gemini, qwen, grok [cal]	BTC+ETH	4h/1h	+$319.2	97.1%	0.0%	35	0.970
WK Quality	gpt, gemini, qwen [cal]	BTC+ETH	4h/1h	+$277.9	90.2%	0.6%	41	0.972
M30 Calmar	gpt, deepseek [cal]	BNB+XRP	1d/1d,4h	+$685.8	96.0%	3.7%	25	0.961
M30 Calmar	gpt, deepseek [cal]	BNB+XRP	1d/1d,4h	+$685.8	96.0%	3.7%	25	0.961
M30 Calmar	gpt, deepseek [cal]	BNB+XRP	1d/1d,4h	+$571.5	96.0%	3.1%	25	0.961
M30 WR	gpt, deepseek [cal]	BNB+XRP	1d/1d,4h	+$685.8	96.0%	3.7%	25	0.961
M30 WR	gpt, deepseek [cal]	BNB+XRP	1d/1d,4h	+$685.8	96.0%	3.7%	25	0.961
M30 WR	gpt, deepseek [cal]	BNB+XRP	1d/1d,4h	+$571.5	96.0%	3.1%	25	0.961
M30 Smoothness	gpt, deepseek [cal]	BNB+XRP	1d/1d,4h	+$457.2	96.0%	2.5%	25	0.961
M30 Smoothness	gpt, deepseek [cal]	BNB+XRP	1d/1d,4h	+$457.2	96.0%	2.5%	25	0.961
M30 Smoothness	gpt, deepseek [cal]	BNB+XRP	1d/1d,4h	+$571.5	96.0%	3.1%	25	0.961
M30 Quality	gpt, deepseek [cal]	BNB+XRP	1d/1d,4h	+$685.8	96.0%	3.7%	25	0.961
M30 Quality	gpt, deepseek [cal]	BNB+XRP	1d/1d,4h	+$685.8	96.0%	3.7%	25	0.961
M30 Quality	gpt, deepseek [cal]	BNB+XRP	1d/1d,4h	+$571.5	96.0%	3.1%	25	0.961

[cal] = calibrate=ON cell; the board column names the window. The WK Calmar board is degenerate by construction -- its top row sits at 0.0% maxDD, under the 1.0 floor of net / max(maxDD, 1.0), so the score collapses onto net PnL; read it as 'net among configs that barely drew down'. The M30 board's top row sits at 3.7%, above the floor, so that board ranks by the ratio. The Smoothness boards are the ones that can leave a window's dominant universe: the WK board's top row is a BTC+ETH config at R2 0.975 on a net of +$236.2, the M30 board's a BNB+XRP config at R2 0.961 on +$457.2. That trade -- straightness bought with size -- is exactly what the board exists to expose. The M30 boards are drawn from a mature pool of 2,809 configs against 1,882 on WK, because four weeks clear a >=20-trade gate that one week does not.

Stability, and what it is across this issue. The board is defined as **mature net top-40 on BOTH windows** of the run that computes it. The run that produced it here is the night collect, whose two windows are **WK and M30** -- the same two windows this issue prints, computed in a second execution of the WK plan -- so this issue's board reads **stable across WK and M30**. **1 configuration qualifies.** It is the solo gemini config on BNB+XRP, clean (zero unfunded entries on both windows), ranked 34 on M30 and 15 on WK. The shared board is narrower than either window's own: 6 of the 20 rows exported from its WK board are BTC+ETH; 16 of the 20 rows exported from its M30 board are SOL+XRP, and no row exported from one window's mature board appears on the other's. The count is not comparable with issue #4's 25 (its board was measured across M30 and the calendar month); the boards measured across WK and M30 were issue #11's, at 0, issue #10's, at 4, issue #9's, at 7, issue #8's, at 4, issue #7's, at 16, issue #6's, at 9, issue #5's, at 0, and issue #3's, at 21. Full list in the open data (stability).

Hand-tuned baseline vs the grid

Standing section, ninth issue. One hand-tuned configuration is treated as the reference the search has to beat -- not on net PnL, where any large search wins by construction, but on the three properties a configuration is actually tuned for: **win rate, shallow drawdown and a straight equity line**. The reference (referred to below as the **hand-tuned anchor**) is the live preset 1d_3_llm_BC2_v02, a 3-model 1d council on SOL+XRP at the $450 two-coin stake; it is distinct from the pre-search hand-tuned set A/B/C carried in Baselines below. Its replay row is in the walk-forward table above; here it is the yardstick.

Beat rule (frozen with the criteria): a config beats the anchor only if it does so on all three at once -- win rate higher, maxDD lower and smoothness R2 higher -- among mature, non-concentrated search rows. Net PnL is not part of the rule; it is reported next to the winners so the cost of the improvement is visible.

The anchor on each window

Window	Config	Coins	FH/TF	Net	WR	maxDD	R2	Ulcer	Cls
WK	**Hand-tuned anchor (1d_3_llm_BC2_v02)**	SOL+XRP	1d/1d	-$484.1	25.0%	52.8%	0.825	29.79	4
M30	**Hand-tuned anchor (1d_3_llm_BC2_v02)**	SOL+XRP	1d/1d	-$582.8	50.0%	58.3%	0.265	24.55	14

211 of the 1,882 mature WK configs (11.2%) clear all three bars at once, and 1,637 of the 2,809 mature M30 configs (58.3%). The anchor's own two windows are different rows: on the week -$484.1 net at 25.0% win rate with a 52.8% drawdown, an R2 of 0.825 and 4 closed trades; on the four weeks -$582.8 at 50.0%, a 58.3% drawdown, an R2 of 0.265 and 14 closed. The week row is thin against the >=15 gate the search rows must pass (the anchor is exempt from that gate by construction: it is the reference, not a candidate). Read against net PnL, all 10 of the 10 best WK winners by net also beat it on net PnL; all 10 of the 10 best M30 winners by net also beat it on net PnL. The share itself moves with the window (11.2% of the WK field, 58.3% of the M30 field), so read it as one field at a time, not as a property of the anchor.

Top-10 by net among the 211 WK configs that beat the anchor on win rate, maxDD and smoothness at once

#	Config	Coins	FH/TF	Stake	Net	WR	maxDD	R2
#1	gpt, gemini	SOL	4h/4h,1h	$450	+$412.4	73.9%	0.0%	0.937
#2	gemini, qwen, grok	BTC+ETH	4h/4h,1h	$450	+$401.8	70.4%	7.2%	0.827
#3	gemini, qwen, grok	BTC+ETH	4h/4h,1h	$450	+$376.9	63.3%	7.2%	0.835
#4	gemini, qwen, grok	BTC+ETH	4h/1h	$450	+$375.7	78.8%	3.1%	0.892
#5	gpt, gemini, qwen	BTC+ETH	4h/1h	$450	+$354.6	74.4%	3.5%	0.834
#6	gemini, qwen, grok	BTC+ETH	4h/4h,1h	$375	+$334.8	70.4%	6.0%	0.827
#7	gemini, qwen, grok	BTC+ETH	4h/1h	$450	+$334.6	76.7%	4.3%	0.848
#8	gpt, gemini, qwen [cal]	BTC+ETH	4h/1h	$450	+$333.4	90.2%	0.7%	0.972
#9	gpt, gemini, qwen [cal]	BTC+ETH	4h/1h	$450	+$326.4	81.8%	3.9%	0.927
#10	gpt, gemini, qwen [cal]	BTC+ETH	4h/1h	$450	+$323.0	81.5%	3.9%	0.926

Top-5 by net among the 1,637 M30 configs that beat the anchor on win rate, maxDD and smoothness at once

#	Config	Coins	FH/TF	Stake	Net	WR	maxDD	R2
#1	gpt	SOL+XRP	1d/1d	$375	+$973.1	77.5%	37.2%	0.764
#2	gpt	SOL+XRP	1d/1d	$375	+$973.1	77.5%	37.2%	0.764
#3	gpt	SOL+XRP	1d/1d	$450	+$911.7	77.8%	44.7%	0.647
#4	gpt	SOL+XRP	1d/1d	$450	+$911.7	77.8%	44.7%	0.647
#5	opus, gpt	SOL+XRP	1d/1d	$375	+$880.9	73.5%	36.8%	0.762

Anchor passport (verbatim, as stored): members=gemini-3.1-pro,gpt-6.1-sol,deepseek-v4-pro tokens=SOL,XRP fh=1d tf=1d conf_mode=each min_conf=60 max_sideways=2 max_diff_side=0 tp_sl_source=nearest trade_mode=position same_side=update opposite=reverse no_signal=hold time_stop=3x steps=40:80,100:20 steps_on=1 be=1 tp_shift=30 sl_shift=50 deposit=1000 lev=10 refill=0 stake=450 calibrate=- variant=v0

Per-ticker bests

Which coin was extractable this window, and by what. Single-coin results are excluded from the council cards by rule (idiosyncratic-coin risk) and reported here, never ranked against the cards. The dedicated per-coin branch contributed 564 + 1,110 sims -- a compacted grid, one best per coin per window; the branch tag in brackets says which kind won.

WK (Oct 19-26, 2026)

Coin	Best config (branch)	FH/TF	Net (trades)
BTC	grok [solo]	1d/4h	+$135.8 (4) *
ETH	gemini, qwen, grok [council]	4h/1h	+$305.8 (21)
SOL	gpt, gemini [council]	4h/4h,1h	+$412.4 (23)
BNB	gemini [solo]	4h/1h	+$156.0 (14)
XRP	fable, gemini [council]	4h/4h,1h	+$376.5 (19)

M30 (Sep 28-Oct 26, 2026)

Coin	Best config (branch)	FH/TF	Net (trades)
BTC	gpt [solo]	4h/4h	+$225.6 (37)
ETH	gpt, deepseek [council]	1d/1d,4h	+$568.4 (15)
SOL	gpt [solo]	1d/1d	+$758.1 (23)
BNB	gemini [solo]	4h/4h,1h	+$394.4 (31)
XRP	opus, gpt [council]	1d/1d	+$606.1 (16)

n in parentheses; * = fewer than 10 closed trades (thin -- direction only). 1 of the ten winners runs calibrate ON; 3 of the five WK winners and 2 of the five M30 winners are councils.

The reads. On the week SOL leads at +$412.4 and 1 of the five rows is thin -- 4 to 23 closed trades against a 15-trade gate. On the four weeks no row is thin: 15 to 37 closed, SOL leading at +$758.1, and the same coin leads both windows. The 1d horizon wins 1 of the five WK coins and 3 of the five M30 coins. **Winner's curse applies to this whole table**: each row is the maximum of a per-coin grid, biased high by selection alone, and none has an out-of-sample read until next issue.

Baselines

Three independent reference points, none of them search output, re-simulated fresh on both of this issue's frozen windows inside the same two collects (engine 1.1): the engine solo default for each of the 7 models, the default council-of-7, and the pre-search hand-tuned set A/B/C carried since issue #1. The tables show the universe-aware **table-stake** run -- this issue's canonical mode ($200 for the 5-coin defaults; A and B already at their $300 3-coin cap; C moves $400 -> $450 on 2 coins) -- with the best and worst of the seven solo defaults on each window; all eleven rows per window and both stake modes are in the open data.

WK (Oct 19-26, 2026)

Baseline	Net PnL	Trades	WR	maxDD
Best solo default (gemini-3.1-pro)	-$22.4 *	131	55%	4.0%
Worst solo default (qwen-3.8-max)	-$55.4 *	283	58%	5.6%
Council-of-7 default	-$8.4 *	170	69%	1.2%
Hand-tuned config A (pre-search)	-$76.2 *	8	63%	30.6%
Hand-tuned config B (pre-search)	-$16.2	24	63%	12.9%
Hand-tuned config C (pre-search)	-$184.3	5	40%	20.4%

M30 (Sep 28-Oct 26, 2026)

Baseline	Net PnL	Trades	WR	maxDD
Best solo default (claude-opus-5.5)	-$121.1 *	652	62%	12.1%
Worst solo default (qwen-3.8-max)	-$216.0 *	979	56%	21.6%
Council-of-7 default	-$52.1 *	495	65%	5.6%
Hand-tuned config A (pre-search)	+$239.2 *	20	80%	44.0%
Hand-tuned config B (pre-search)	-$493.5 *	97	60%	54.1%
Hand-tuned config C (pre-search)	-$305.5 *	25	52%	38.0%

* = non-zero skipped_nofunds at table stake. On the 5-coin defaults those counts are the stake table doing its job: at $200 a $1,000 deposit funds at most five concurrent positions while the defaults fire hundreds of entries per window (196 to 412 skipped each on the week, 674 to 1,346 on the four weeks). The other five solo defaults run between the two printed rows on each window. **The conclusions do not change between modes**: on WK they do so for every row, and on M30 they do so for every row. 2 of the three hand-tuned configs are already at their own table cap, so their two modes are identical on both windows. Source: baselines_wk.json (22 sims) and baselines_night.json (44 sims, of which 22 are the M30 half), inside this issue's two collects.

Every engine default lost money on both windows; the hand-tuned set lost money on WK and split on M30. On WK the seven solo defaults run -$22.4 to -$55.4 at table stake, the council-of-7 -$8.4, and the three hand-tuned configs A -$76.2, B -$16.2, C -$184.3 -- all negative. On M30 the defaults run -$121.1 to -$216.0 solo (council-of-7 -$52.1) and the three hand-tuned configs A +$239.2, B -$493.5, C -$305.5 -- 1 of the three positive. The upper half of the standing ordering (search > hand-tuned > defaults) holds on both windows, in-sample as ever -- the fresh search makes +$580.5 on the week against a best baseline of -$8.4, and +$880.9 on the four weeks against +$239.2. The live presets are a separate row of evidence: 0 of the four are net-positive on the week and 0 of four on the month (walk-forward table above), and 2 of them run stakes above this issue's cap.

Live presets & preset neighbourhoods

Four hand-tuned configurations trade live on the platform (not search output, distinct from the pre-search A/B/C set): preset bc2 = '1d_3_llm_BC2_v02'; preset 4h4 = '4h_4_llm_v02'; preset sol = 'Sol'; preset xrp = 'XRP'. Their replay rows lead the walk-forward table; passports and window detail are in the open data (presets_recon_wk.json and presets_recon_night.json). The neighbourhood grid asks a narrower question: does a one-knob neighbour beat the preset as configured? Two of the four have a local grid this issue (53 cells per window); the two single-coin presets do not (0 cells -- the local grid is not built for them, so no neighbour claim is made about either).

Live preset	Window	As configured	Best neighbour in local grid	Cells	Verdict
preset bc2	WK	-$484.1	+$249.5 (1d/4h, SOL+XRP, 8 cls)	34	neighbour +$733.6
preset bc2	M30	-$582.8	+$273.2 (1d/4h, SOL+XRP, 23 cls)	34	neighbour +$856.0
preset 4h4	WK	-$223.7	-$85.0 (4h/1h, BTC+ETH, 5 cls)	19	neighbour +$138.7
preset 4h4	M30	-$299.9	+$10.3 (4h/1h, ETH+SOL, 35 cls)	19	neighbour +$310.1

In **4 of the 4** preset-window cells with a grid, a one-knob neighbour beat the configuration as it is actually running. The largest single edit is preset bc2's 1d/1d -> 1d/4h horizon move on M30, which turns -$582.8 into +$273.2. The presets are not at a local optimum -- a testable, low-risk edit, unlike adopting a search champion wholesale. Cell counts are small (34 and 19 per window) and every neighbour is an in-sample maximum of its own little grid.

Patterns: what the finalists look like

Knob modes across each window's unique finalists (mature net top-10 per branch plus all four quality boards, deduplicated by full signature; 51 unique configs on WK, 40 of them councils; 49 on M30, 29 councils).

Knob	WK finalists mode	M30 finalists mode
Forecast horizon	4h (51/51)	1d (36/49)
Ladder shape (steps)	50:40,100:60 (51/51)	50:40,100:60 (49/49)
Break-even stop (BE)	on (51/51)	on (49/49)
Min confidence	60 (49/51)	60 (47/49)
SL shift	75 (51/51)	75 (49/49)
TP shift	-10 (51/51)	-10 (49/49)
TP/SL source	farthest (45/51)	farthest (41/49)
Calibrate ON	30/51	18/49
Universe	BTC+ETH (25/51)	SOL+XRP (18/49)

The mechanical core is the same on both windows and unchanged for the twelfth issue running: 50:40,100:60 ladder, break-even on, conf 60, SL +75 / TP -10, farthest source. The horizon is **4h** on WK (51/51) and **1d** on M30 (36/49). **Calibrate=ON** runs in 30 of 51 WK finalists against 18 of 49 on M30, where issue #11 measured 19/58 on its week and 29/52 on its month. The universe: BTC+ETH takes 25 of the 51 WK finalists while SOL+XRP takes 18 of the 49 M30 finalists. Membership concentrates too -- gemini leads the WK council finalists (40 of 40) and gpt the M30 ones (25 of 29). Standing caveats: refinement seeds from the same grid winners, so part of the convergence is search-design echo; and a knob core that produced 21-of-69 survival one issue ago and 13-of-74 this issue is describing the weather, not a recommendation.

Liquidation note

Liquidation is modeled by the engine at the fixed 10x leverage used throughout. Across the published cards, stop-loss placement stays inside the distance that would approach the liquidation threshold; maxDD is printed on every config so realized risk is visible directly. This issue the published cards draw down 5.6% and 0.7% on the week and 36.8% and 21.4% on the month, and the replay table is deeper: 7 replayed rows drew down more than 40% on the week and 46 on the month, 1 and 26 of them past 55%. A high headline net and a survivable path are different claims, and so are a shallow in-sample drawdown and a shallow one next week.

Watch amendments (this issue)

Amendment 3 (standing) -- smoothness and a quality composite. Every simulation stores its daily closed-equity series and the three numbers read off it (n_days, smooth_r2, ulcer); the smoothness board ranks mature rows with net > 0 and at least 5 days by R2 with ulcer as tie-break, and the quality composite ranks by the mean of the net, Calmar-like, win-rate and R2 ranks. Both still cost zero additional simulations and the headline board is still net PnL.

Amendment 4 (standing) -- the monthly split. The weekly Config Watch keeps two windows, the calendar week (WK) and the four calendar weeks that end at the same edge (M30, Sep 28-Oct 26, 2026); the calendar-month board belongs to the Monthly Config Watch line and no calendar-month number is printed in a weekly issue. This issue's night collect computes no calendar-month window at all: calendar September belongs to Monthly Config Watch #2, published Oct 3, and Q3 2026 to Quarterly Config Watch #1, published Oct 9, whose 3 cards are on the walk-forward table here.

Amendment 5 (standing) -- the standing OOS board grows with every issue. Every card this series publishes joins the walk-forward table and is never removed: this issue replays **74 cards** -- issue #11's 5 join issue #10's 5, Quarterly Config Watch #1's 3, issue #9's 5, Monthly Config Watch #2's 3, issue #8's 5, issue #7's 5, issue #6's 5, issue #5's 4, issue #4's 5, Monthly Config Watch #1's 3 and issues #1, #2 and #3's 26 -- plus the four live presets, on both fresh windows. **Monthly and quarterly cards are judged by the weekly rule** (net-positive on both of this issue's fresh windows), the same rule every other row is judged by; their own calendar-month or calendar-quarter origin window is the Home column and nothing else. One consequence for the counts: a survival share is computed on a board that changes size every issue, so it is reported with the cohort breakdown beside it and never as a trend on its own. **Stability** is defined inside the run that computes it: this issue's night collect computes WK and M30, so the board reads across WK and M30 -- the issue-#3 definition, the one issues #5, #6, #7, #8, #9, #10 and #11 used as well -- where issue #4's read across M30 and the calendar month.

All three amendments are standing, provisional and scoped to this series. **Formalization is slated for methodology v1.2**; until then this issue runs on v1.1 plus the Config Watch amendments approved 12.08 (hash e66c7e8c864a2233).

Limitations

- **Two windows, one regime, and they overlap.** M30 (Sep 28-Oct 26, 2026) contains WK (Oct 19-26, 2026) -- 7 of the week's 7 days are inside it -- and it shares 21 of its 28 days with issue #11's own M30 window (shared stretch Sep 28-Oct 18), 14 of its 28 days with issue #10's own M30 window (shared stretch Sep 28-Oct 11), 7 of its 28 days with issue #9's own M30 window (shared stretch Sep 28-Oct 4), 3 of its 28 days with Monthly Config Watch #2's calendar-September window (shared stretch Sep 28-30) and 3 of its 28 days with Quarterly Config Watch #1's calendar-quarter window (shared stretch Sep 28-30). The two columns of this issue are therefore not two independent tests: agreement between them is partly arithmetic, and for the issue-11, issue-10, issue-9, Monthly-2 and Quarterly-1 cohorts the M30 replay column is decay context rather than pure out-of-sample. Do not read them as a cross-validation.
- **The two windows come from two collects, and the two runs of the same week do not agree.** WK is the weekly run (generated 2026-10-28T06:10:18Z), M30 the night run (generated 2026-10-28T06:15:57Z), which recomputed the same WK week in the same pass on its own grid: the night run's WK ceiling is +$226.4 (gpt, deepseek, qwen on BTC+ETH) against the weekly collect's +$580.5. Every WK number this issue prints is the weekly collect's; the night run's WK half is used only for the cross-window stability board. Same engine, same criteria, same collect code, but not the same execution -- a number is comparable across the two windows only as far as that is.
- **Stability is defined inside the run that computes it.** The board comes from the night run, whose two windows are WK and M30, so it reads across those two -- and 1 configuration is in the mature net top-40 of both. Issue #11 read the same pair one week earlier and held 0, issue #10 the week before that held 4, issue #9 7, issue #8 4, issue #7 16, issue #6 9, issue #5 0 and issue #3's board (the other one across the same pair) held 21; issue #4's 25 is not comparable at all, having been measured across M30 and the calendar month. A board this size is one observation per issue and no claim is made from a single reading.
- **Model lineage is a splice.** Three model axes run as lineages (claude-opus-5.5 (incl. opus-5-era), gpt-6.1-sol (incl. 5.6-era), grok-4.7 (incl. 4.6-era)) across the cutover at Oct 17, 2026 09:01 UTC, which lies inside this issue's M30 window (Sep 28-Oct 26, 2026, day 20 of 28) and before its WK week (Oct 19-26, 2026): every WK number in this issue is a single-build week on the new ids, every M30 number joins the two builds of each line under one id; every published card that named a superseded id is replayed on the lineage head ([remap] rows: 57 of the 78 rows -- 53 of the 74 cards (by cohort issue-10 5 of 5, Quarterly-1 2 of 3, issue-9 4 of 5, Monthly-2 2 of 3, issue-8 4 of 5, issue-7 4 of 5, issue-6 5 of 5, issue-5 3 of 4, issue-4 3 of 5, Monthly-1 1 of 3, issue-3 4 of 6,

issue-2 7 of 8 and issue-1 9 of 12) and all four live presets), and a remapped replay is not the same simulation the origin issue ran. The grok 4.5 -> 4.6 cutover of 2026-08-24 lies entirely before both windows.

- **The cutover hours.** The two roll-out slots of Oct 17 (07:01 and 08:01 UTC; invalid rows for the three superseded ids, council of four in those hours) fall inside the M30 window and outside the WK week.
- **Calibrated replays are pinned to their origin point.** 12 replay rows carry calibrate=1 with the calibration point fixed in their stored query (calibrate_asof: 2026-08-25, 2026-09-01, 2026-09-07, 2026-09-14, 2026-09-15, 2026-09-21, 2026-10-01, 2026-10-06, 2026-10-20), so their TP/SL multipliers are the ones in force when the origin issue was computed; up to issue #11 a calibrated replay used the live point of the day it ran. The rows: cw11-m302, cw3-m302, cw4-wk2, cw5-wk1, cw5-wk2, cw6-m302, cw6-wk2, cw7-m302, cw9-wk2, cwm1-m2, cwm2-s1, cwq1-q2. Their week-over-issue trajectory therefore changes method at this issue, once.
- **Concentrated rows are published, not hidden.** 4 rows run stakes above this issue's cap for their universe size and are marked [conc] (preset_sol, preset_xrp, cw2-30d1, cw2-30d2); their printed economics assume funding the current rule would not grant. They stay in the table because removing them would flatter the preset row.
- **The stake rule changes what runs, not only its size.** Capping the stake changes WHICH entries get funded: the seven table-stake solo defaults skip 196 to 412 entries each, and 22 replay rows carry unfunded entries (up to 4 on cw4-m301). Where skipped_nofunds is non-zero the printed economics are not the economics that ran; the rows are flagged, never silently pooled.
- **Snapshot principle.** This report is a frozen snapshot taken at the generated-at timestamps; numbers are not updated retroactively and past issues are not restated. The drift baseline exists because a re-run does not reproduce every past number; the widest this issue (cw11-m301: +$1,484.4 printed, +$1,344.6 today) is a card with positions still open when its origin window closed, and such a card re-runs to its settled result.
- **Research-to-date counter is pinned.** The counter below is pinned from wave 3 onward and is read at the Oct 26, 2026 16:00 UTC cut-off; issue #11 printed 71,945 resolved and 58 published reports against 77,801 and 62 here. The model line keeps the definition issue #4 introduced (7 tracked, current line-up; earlier versions folded into their successors' lineage), so the nine issues' model counts are comparable and issue #3's is not.
- **No calendar month in this issue.** The night collect computes M30 and WK and nothing else, so no claim about a calendar month is made anywhere here. Monthly Config Watch #2 published Oct 3 and Quarterly Config Watch #1 Oct 9; Monthly Config Watch #3 follows in early November. Quarterly Config Watch #1's 3 cards, Monthly Config Watch #2's 3 and Monthly Config Watch #1's 3 appear in this issue only as replay rows, judged by the weekly rule.
- **Multiple testing.** 16,016 simulations across the two collects (5,369 weekly + 10,647 night); at this scale some winners are expected from chance alone. Antidotes: axes frozen before launch, the walk-forward table, independent baselines, in-sample labeling. No formal correction yet.
- **Winner's curse and thin cells.** Every card, board row and per-ticker best is the maximum of a search, biased high by selection alone. On the week the per-ticker winners closed 4-23 trades and 43 of the 78 replay rows are under 10 closed; on the four weeks 2 replay rows are thin and 0 per-ticker winners are. Flagged with *, reported for completeness.
- Research output, not financial advice.

Counters & lineage

Weekly collect (WK window)

Stage	Sims	Stage	Sims
Scan (s1)	476	Dedicated per-ticker (s6t+s6tc)	564
Systematic grid (s2)	3,952	OOS replays (s4cw11+s4cw11o+s4cw10+s4cw10o+s4cwq1+s4cwq1o+s4cw9+s4cw9o+s4cwm2+s4cwm2o+s4cw8+s4cw8o+s4cw7+s4cw7o+s4cw6+s4cw6o+s4cw5+s4cw5o+s4cw4+s4cw4o+s4cwm1+s4cwm1o+s4cw3+s4cw3o+s4cw2+s4cw2o+s4cw1+s4pre)	140
Refinement: ladder + stake sweep (s2b+s3b)	124	**Total**	**5,369**
Preset neighbourhood grid (s2p)	53	Errors / zero-trade	0 / 1,200
Per-ticker probes (s5t)	60	Engine	1.1

Night collect (M30 and a second execution of the WK week; this issue publishes M30 from it, and its WK half feeds the cross-window stability board only)

Stage	Sims	Stage	Sims
Scan (s1)	952	Dedicated per-ticker (s6t+s6tc)	1,110
Systematic grid (s2)	7,904	OOS replays (s4cw11+s4cw11o+s4cw10+s4cw10o+s4cwq1+s4cwq1o+s4cw9+s4cw9o+s4cwm2+s4cwm2o+s4cw8+s4cw8o+s4cw7+s4cw7o+s4cw6+s4cw6o+s4cw5+s4cw5o+s4cw4+s4cw4o+s4cwm1+s4cwm1o+s4cw3+s4cw3o+s4cw2+s4cw2o+s4cw1+s4pre)	218
Refinement: ladder + stake sweep (s2b+s3b)	247	**Total**	**10,647**
Preset neighbourhood grid (s2p)	106	Errors / zero-trade	0 / 2,414
Per-ticker probes (s5t)	110	Engine	1.1

Scale note (vs issue #11). Issue #11 ran 15,740 sims over two windows in two collects; issue #12 runs 16,016 over two windows in two -- 5,369 in the weekly collect (WK) and 10,647 in the night collect (M30 and the same WK week together). The per-window grid is close to unchanged: 3,952 WK cells and 3,952 M30 cells here against 3,760 WK and 3,920 M30 cells there. The per-coin branch is 564 and 1,110 against 545 and 1,105. The replay branch grew in rows (78 configs against 73) and in sims (140 and 218 against 130 and 203), because one more cohort joined and every row is replayed on every window of its collect. The two selection boards still cost zero simulations.

No extras pass in either collect. Equity series, the grid histograms (0 extra sims), the baselines and the calibrate packs are produced by the same two collects that wrote the boards. The baseline runs are verification, outside both search totals: 22 sims on WK (11 configs x 2 stake modes) and 44 in the night collect, of which the 22 M30 rows belong to this issue. 0 equity replays were needed. **Model lines in these windows.** Seven lines; three are lineages spliced at the cutover of Oct 17, 2026 09:01 UTC, inside M30 and before WK: claude-opus-5.5 -- claude-opus-5 rows until Oct 17 08:01 UTC, then claude-opus-5.5 from 09:01 (M30: 4,095 + 1,785 rows; WK: 1,470 rows, all under claude-opus-5.5; succession claude-opus-5, claude-opus-4.8); gpt-6.1-sol -- gpt-5.6-sol rows until Oct 17 08:01 UTC, then gpt-6.1-sol from 09:01 (M30: 4,095 + 1,785 rows; WK: 1,470 rows, all under gpt-6.1-sol; succession gpt-5.6-sol); grok-4.7 -- grok-4.6 rows until Oct 17 08:01 UTC, then grok-4.7 from 09:01 (M30: 4,095 + 1,785 rows; WK: 1,470 rows, all under grok-4.7; succession grok-4.6, grok-4.5). The other four lines (claude-fable-5, deepseek-v4-pro, gemini-3.1-pro, qwen-3.8-max) carry one id each.

Generated at: 2026-10-28T06:10:18Z (weekly collect) and 2026-10-28T06:15:57Z (night collect) · 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 timestamps above; numbers are not updated retroactively. Each issue re-runs the pipeline fresh over that issue's windows.

Research to date

RESEARCH SNAPSHOT

THIS REPORT -- search effort: 16,016 simulations (0 errors) across two frozen collects, of which 358 walk-forward replays and origin re-runs · 66 verification sims (baseline runs inside the same collects; 0 equity replays needed)

MARKETMANIA RESEARCH TO DATE (as of cutoff): 77,801 directional forecasts resolved since Jul 11 (as of Oct 26, 2026 16:00 UTC cutoff) · 7 models tracked (current line-up; earlier versions folded into their successors' lineage) · 5 assets · 5 horizons · hourly · 62 published reports

Counter pinned from wave 3 onward and read at the Oct 26, 2026 16:00 UTC cut-off: issue #11 printed 71,945 resolved and 58 published reports. The model line keeps the definition issue #4 introduced (7 tracked, current line-up; earlier versions folded into their successors' lineage), so the nine issues' model counts are comparable. Config Watch #12 is the 62nd published report.

Issue #12. The standing walk-forward board reaches 74 cards -- issue #11's 5 join it, and monthly and quarterly cards are judged by the weekly rule (Amendment 5). Two fresh windows again, the calendar week and the four calendar weeks ending at the same edge, both computed in the night collect as well as the week: the table keeps 13 of 74 past cards where the week alone keeps 29 and the month alone 24. The cross-window stability board is measured on this issue's own two windows (WK and M30) and comes back with **1** configuration. Formalization of the amendments is still slated for methodology v1.2.

What we're testing next

- **The first fresh week of this issue's own cards.** Config Watch #13 replays the 5 cards published here that carry a short link (both WK councils, WK solo #1, both M30 councils; WK solo #2 and both M30 solo cards have no link and are not replayed, as issue #11's two unlinked solo cards were not) on its own fresh week, alongside the 74 already on the board. The number to check against: 80.0% of issue #11's 5 linked cards were net-positive on their first fresh week, which is this issue's week.
- **Whether the WK-and-M30 stability board reads the same twice.** 1 configuration is in the mature net top-40 of both windows of this issue's night run, against 0 on issue #11's board, 4 on issue #10's, 7 on issue #9's, 4 on issue #8's, 16 on issue #7's, 9 on issue #6's, 0 on issue #5's and 21 on issue #3's, all across the same pair. Config Watch #13's night run computes the same pair one week on, so the question is whether the count moves when the windows do.
- **What the calibrate share reads on a tenth window.** Calibration improves 29.5% of the 1,336 matched WK pairs and 38.5% of the 1,342 M30 pairs this issue, against 27.9% and 42.5% in issue #11. The test is the same two measurements on the next pair of windows -- the whole grid and the board-leader slice, reported separately and never pooled.

Related research

Consensus Watch #12 -- https://marketmania.ai/research/reports/consensus-watch-2026-10-19.pdf
Weekly Calibration #12 -- https://marketmania.ai/research/reports/weekly-calibration-2026-10-19.pdf
Weekly Model Watch #12 -- https://marketmania.ai/research/reports/model-watch-2026-10-19.pdf
Config Watch #11 -- https://marketmania.ai/research/reports/config-watch-2026-10-14.pdf
Config Watch #10 -- https://marketmania.ai/research/reports/config-watch-2026-10-07.pdf
Quarterly Config Watch #1 -- https://marketmania.ai/research/reports/config-watch-quarterly-2026-q3.pdf

The three weekly reports publish together as one issue each week; Config Watch follows its own cycle. Market-regime figures, where this issue refers to them, come from that wave and are not re-tabulated here. Quarterly Config Watch #1 is listed because its 3 cards are on the walk-forward table above.

CITE THIS REPORT

```
@techreport{mm_configwatch_2026w44,
  title        = {Config Watch #12},
  author       = {{MarketMania Research}},
  institution  = {MarketMania},
  year         = {2026},
  month        = oct,
  day          = {28},
  type         = {Weekly Research Report},
  series       = {Config Watch},
  number       = {12},
  note         = {Methodology v1.1 + Config Watch amendments, hash e66c7e8c864a2233},
  url          = {https://marketmania.ai/research/reports/config-watch-2026-10-21.pdf}
}
```