**Vasilis Kotopoulos**

AI-native quant research · From strategy abundance to allocation conviction

NOVEMBER 2026

MANDATE ASSESSMENT · SAMPLE DELIVERABLE

# Market-neutral crypto pod: a 'crowding' long-short strategy

> Multi-manager crypto fund · Market-neutral pod

VERDICT

**Requires further research**

Almost no link to bitcoin and a strong held-back test period, but it misses the Sharpe bar and the evidence does not yet rule out luck.

PREPARED BY

**Vasilis Kotopoulos**

Case study against a publicly described mandate, not a past or current client. All results are simulated. Past and simulated performance is not indicative of future results. Not investment advice.

EXECUTIVE SUMMARY

# Executive summary

VERDICT

**Requires further research**

Almost no link to bitcoin and a strong held-back test period, but it misses the Sharpe bar and the evidence does not yet rule out luck.

THE MANDATE

A multi-manager crypto fund wants systematic, market-neutral pods that target about 20% a year net, with a Sharpe ratio above 1.5 (it accepts less if a pod is genuinely different from what it already holds) and little long-run exposure to the crypto market.

| 20.3% | 1.33 | -0.16 | -33% |
|---|---|---|---|
| a year in the held-back test period | Sharpe ratio, test period (bar: 1.5) | correlation with bitcoin, test period | worst fall, design period (2022 – early 2024) |

## Key findings

1. **Made money while bitcoin fell.** In the held-back test period (October 2025 – early November 2026) the pod returned 20.3% a year with a Sharpe ratio of 1.33 and a worst fall of −11%, while bitcoin fell about 30%. Correlation with bitcoin was −0.16.
2. **Positive but bumpy design period.** April 2020 – September 2025: 22.4% a year at 22.4% volatility (Sharpe 1.01), with a deep, long drawdown of −33% that ran from 2022 to January 2024.
3. **Works beyond one coin list.** On the 60 most-traded coins (design Sharpe 1.30) and on all eligible coins (1.39; 36% a year in the test period, worst fall −9%) it also held up. Only the narrow 20-coin version was weak (0.83, worst fall −52%).
4. **Cost-sensitive but survives.** Our base case already charges double our normal cost estimate. Design-period Sharpe is 1.19 at normal costs, 1.01 at double and 0.84 at triple. Turnover is about 39× a year.
5. **Not yet proven.** The 1.5 Sharpe bar is missed in both periods; at a 10% pod risk level the design period earns about 10.9% a year, below the 15–20% target; and luck-adjusted confidence is 21% (54% for the all-coin version) against the 95% we require.

RECOMMENDED NEXT STEP

**Three months of paper trading** at a pod size you choose, reported weekly against bitcoin and your targets, followed by a fixed-fee review. We suggest running the 40-coin and all-coin versions side by side so the choice between them is made on live evidence.

SECTION 1

# Mandate and constraints

The requirements below come from the fund's public description of what it looks for in a pod. We turned them into specific pass rules before testing.

| Requirement | What the fund asks for | How we tested it |
|---|---|---|
| Return | About 20% a year net | Annual return after costs, design and test periods; also at a 10% pod risk level |
| Return quality | Sharpe above 1.5; less if truly diversifying | Sharpe at least 1.5 in the test period; at least 1.0 in the design period |
| Market exposure | Little long-run crypto beta | Correlation with bitcoin below 0.2 in both periods |
| Diversification | Different from existing pods | Correlation with bitcoin; behaviour when bitcoin falls |
| Trading style | Preference for mid-to-high-frequency trading | Weekly rebalancing; flagged as a mismatch |
| Evidence | Robust, not a lucky backtest | Luck-adjusted confidence of at least 95% to call it client-ready |

> **Pass rules, fixed in advance:** test-period Sharpe of at least 1.5; correlation with bitcoin below 0.2 in both periods; design-period Sharpe of at least 1.0; at least 15% a year at a 10% risk level; and luck-adjusted confidence above 95% to be called client-ready. Otherwise the best result is reported as promising but not proven.

SECTION 2

# Approach

## What we built

A weekly long-short strategy across the most-traded coins on a major crypto derivatives exchange. It is based on **crowding**: in perpetual futures, traders holding leveraged long positions pay a recurring fee when demand to be long is high. That fee tells us where bullish positioning is most crowded.

- **Signal.** Each week, coins are ranked by how much traders have been paying to hold long positions, averaged over several look-back windows fixed in advance.
- **Positions.** The pod buys the least crowded coins and sells the most crowded ones, sizing positions by risk so calmer coins get larger weights.
- **Market neutrality.** The long and short sides are balanced so that a move in bitcoin affects both sides about equally.
- **Universe.** The 40 most-traded eligible coins each week, chosen from a list that includes coins later delisted, so results are not flattered by looking only at survivors.

## How we tested it

- **Rules locked in advance.** Variants, pass rules and tests were written down and time-stamped before any result was seen. Every variant counts toward the luck adjustment.
- **Design and test separation.** All choices used April 2020 – September 2025 data. October 2025 – early November 2026 was held back and used once.
- **Conservative costs.** Every trade happens the day after its signal. The base case charges 0.20% per unit traded, double our normal estimate, plus the actual funding payments on every position.
- **Universe robustness.** The same rules on the 20 and 60 most-traded coins and on all eligible coins, plus a version run at a fixed 15% risk level.
- **Statistical checks.** Luck adjustment for all 741 crypto ideas tested to date, an overfitting check across variants, and a random sign-flip test.
- **Benchmark.** Bitcoin buy-and-hold over the same periods.

SECTION 3

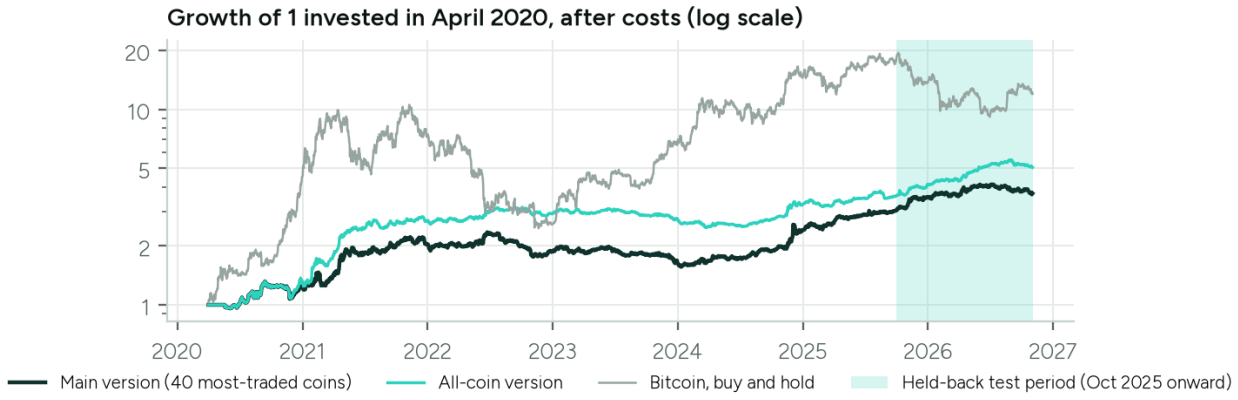
# Results

## Pass-rule scorecard

| Rule | Target | Result | Status |
|---|---|---|---|
| Test-period Sharpe | at least 1.5 | 1.33 | Not met |
| Correlation with bitcoin | below 0.2, both periods | −0.06 (design); −0.16 (test) | Met |
| Design-period Sharpe | at least 1.0 | 1.01 | Met |
| Return at a 10% risk level | at least 15% a year | 10.9% (design period) | Not met |
| Luck-adjusted confidence | above 95% | 21% (all-coin version 54%) | Not met |

## Results by period

| Measure | Design: Apr 2020 – Sep 2025 | Test: Oct 2025 – Nov 2026 |
|---|---|---|
| Annual return | 22.4% | 20.3% |
| Sharpe ratio | 1.01 | 1.33 |
| Worst fall | −33.4% | −11% |
| Correlation with bitcoin | −0.06 | −0.16 |
| Months with a positive return | 62% | — |
| Bitcoin buy-and-hold, same period | 69% a year, Sharpe 1.18 | about −30% |




Growth of 1 invested in April 2020, after costs. The shaded area is the held-back test period.




Fall from previous peak in the design period. The main version's deepest fall (−33%) bottomed in January 2024 after a long slide through 2022 and 2023.

## Robustness across coin universes (design period; test period in last column)

| Version | Annual return | Sharpe | Worst fall | Test period |
|---|---|---|---|---|
| 20 most-traded coins | 25.3% | 0.83 | −52.1% | positive but weaker |
| 40 most-traded coins (main) | 22.4% | 1.01 | −33.4% | 20.3% a year, worst fall −11% |
| 60 most-traded coins | 24.6% | 1.30 | −27.7% | positive |
| All eligible coins | 26.2% | 1.39 | −20.8% | 36% a year, worst fall −9% |
| Main version at a fixed 15% risk level | 14.3% | 0.87 | −31.2% | 25% a year, worst fall −14% |




Design-period Sharpe ratio of the main version as trading costs rise. The base case (dark bar) already assumes double our normal cost estimate.

## Calendar-year returns (design period)

| Year | Main version (40 coins) | All-coin version | Bitcoin |
|---|---|---|---|
| 2020* | +26.0% | +32.1% | +351.2% |
| 2021 | +62.0% | +103.0% | +59.8% |
| 2022 | -7.9% | +9.0% | -64.2% |
| 2023 | -12.3% | -6.6% | +155.6% |
| 2024 | +46.0% | +19.8% | +121.3% |
| 2025* | +26.1% | +10.2% | +21.9% |

*2020 from April; 2025 to September. Test-period figures are reported exactly as they came out; only the figures shown are reported for that period.

SECTION 4

# Robustness and risk assessment

## How much confidence the evidence supports

- **Unlikely to be pure noise.** A random sign-flip test on the design period gives a probability of about 0.6% that a return this good arises by chance.
- **But not proven once we allow for how many ideas were tried.** 741 crypto ideas have been tested in our research. After that adjustment, the chance that the main version's record reflects real skill is about 21%; for the all-coin version it is about 54%. We require 95%.
- **Moderate overfitting risk.** The overfitting check across the variants came out at about 25%. Selection added value, but not overwhelmingly.
- **The test period was better than the design period** (Sharpe 1.33 versus 1.01). That is encouraging, but a single year with a falling bitcoin market is a narrow sample.

## Key risks

- **Deep, long drawdowns.** The main version lost 7.9% in 2022 and 12.3% in 2023, and took from 2022 to early 2024 to bottom at −33%. A pod with tight drawdown limits would likely have been stopped out.
- **Costs and turnover.** At about 39× turnover a year, results depend on execution quality. Triple costs cut the design-period Sharpe to 0.84.
- **Return at the pod's risk level.** Scaled to a 10% risk level, the design period earns about 10.9% a year, well below the 15–20% target.
- **Crowding of the idea itself.** The signal is based on a publicly observable fee. If more capital trades it, the edge may shrink.
- **Style mismatch.** Weekly rebalancing is slower than the mid-to-high-frequency trading the fund prefers.
- **Exchange and venue risk.** Results assume trading on a single derivatives exchange; counterparty, listing and liquidity risks are not modelled.

> **Bottom line.** The pod is genuinely different from bitcoin exposure and performed well when bitcoin fell. It is not yet a 1.5-Sharpe strategy, and the statistics do not rule out luck. We would offer it as a diversifier on paper, not as a finished pod.

SECTION 5

# Verdict and rationale

**Requires further research**

**Almost no link to bitcoin and a strong held-back test period, but it misses the Sharpe bar and the evidence does not yet rule out luck.**

### Why further research is required

- Three of the five pre-set pass rules were missed: test-period Sharpe (1.33 versus 1.5), return at a 10% risk level (10.9% versus 15%) and luck-adjusted confidence (21% versus 95%).
- The design period included a −33% drawdown lasting into 2024, which is hard to reconcile with a pod mandate.
- On the positive side, the market-neutral property held in both periods, the test period was strong, and the result was not confined to one coin list.

### What would move it to a pass

- A live paper record that sustains the test-period profile, especially a Sharpe ratio near 1.5 with drawdowns well inside the fund's limits.
- Evidence that the broader all-coin version, which had the better design-period profile, holds up live at realistic costs.

SECTION 6

# Recommended next steps

1. **Three-month paper trial.** Run the 40-coin and all-coin versions on paper at a pod size you choose, reported weekly against bitcoin and your targets.

2. **Execution check.** Measure live trading costs against the 0.20% per-unit base case, since results are sensitive to costs.

3. **Fixed-fee review.** At the end of the trial, a written review against the same pre-set pass rules, with a clear recommendation to proceed or stop.

APPENDIX

# Definitions in plain English

| Term | What it means |
|---|---|
| **Annual return** | The average yearly growth rate over the period, compounded, after modelled trading costs. |
| **Volatility** | How much returns swing around, expressed per year. Higher means a bumpier ride. |
| **Sharpe ratio (return per unit of risk)** | Average return divided by volatility, per year. Above 1 is generally considered good for a single strategy. Unless stated otherwise, we show it on total return. |
| **Worst fall (maximum drawdown)** | The largest drop from a previous high point to a subsequent low, before a new high is reached. |
| **Correlation** | A measure from −1 to +1 of how closely two return streams move together. Near zero means they move independently. |
| **Market-neutral** | Long and short positions are balanced so that the overall book has little exposure to the direction of the crypto market. |
| **Funding fee (crowding signal)** | A recurring payment between long and short holders of perpetual futures. When many traders want leveraged long exposure, longs pay shorts; a high fee signals crowded bullish positioning. |
| **Turnover** | How much of the book is traded per year. 39× means the full book is replaced about 39 times a year. |
| **Delisted coins** | Coins that were later removed from the exchange. Including them avoids flattering results by testing only on survivors. |
| **Design period** | The historical window in which all choices were made. Nothing after 30 September 2025 was used to design or select anything. |
| **Held-back test period** | Data from 1 October 2025 to early November 2026, kept aside and used only once to check the final rules. A simulation on unseen data, not live trading. |
| **Luck-adjusted confidence** | Testing many ideas makes it likely that some look good by chance. This statistic (a "deflated Sharpe ratio") estimates the probability that the record reflects real skill, after allowing for the total number of ideas we have tested. We require 95% to call a result statistically proven. |
| **Overfitting check** | Estimates how often the version that looked best in one half of history ends up below average in the other half. Lower is better; below 50% means selection added value. |
| **Sign-flip test** | We randomly flip the sign of daily returns 1,000 times. If the real average return is rarely matched, the result is unlikely to be noise. |

## Important information

This document is a sample deliverable. It was prepared against a mandate described in public sources; the organisation that published that mandate is not a past or current client and did not commission, review or endorse this work.

All performance figures are simulated (backtested) and hypothetical. They were produced by applying fixed rules to historical market data, after modelled trading costs, and do not reflect actual trading or real money. Simulated results have inherent limitations: they are prepared with the benefit of hindsight, cannot fully account for liquidity, market impact, financing costs, operational constraints or the effect of real trading decisions, and may differ materially from results actually achieved. The held-back test period is also simulated, not live trading.

**Past and simulated performance is not indicative of future results.** Any investment can lose value, and leverage magnifies losses as well as gains.

This document is provided for information and discussion only. It is not investment advice, a recommendation, or an offer or solicitation to buy or sell any security, fund or strategy, and it does not take into account the objectives, financial situation or needs of any investor. Recipients should make their own independent assessment and seek professional advice before making any investment decision.