

Polymarket Kink Strategy

Black-Scholes mispricing on crypto Up/Down markets: how it works, a historical replay on recorded data, and whether to re-enable it

Research paper 3 (papers 1-2 = Polymarket market overview, Momentum strategy)

Data: replay 2026-04-19 to 2026-06-17, 321 trades; live paper 2026-05-06 to 2026-05-21

Generated 2026-10-19

Internal research - not investment advice

0 Executive summary

Verdict: no edge after costs - keep Kink disabled. Replaying the strategy on 44 days of recorded data (2026-04-19 to 2026-06-17) produced 321 trades with a win rate of **44.5%** against a **45.8%** break-even, for **-0.022 \$ per 1-share stake** (-7.09 \$ in total). Before costs the signal is only just positive (**+0.029 \$/share**); the ~5.1c/share of slippage and fees is about **1.8x** the gross edge. The +0.114 \$/share the live paper run showed in May does not reproduce: the replay of the same days fires different, fewer trades and loses money.

Metric	Value	Comment
Trades / days	321 / 35 trading days of 44 replayed	run stopped at 44 of 123 available dates
Win rate vs break-even	44.5% vs 45.8%	break-even = average NO price paid (incl. slippage)
Net PnL	-7.09 \$ per 1-share stake	-0.022 \$/trade, Binance-graded
Gross PnL (no costs)	+9.37	cost drag 176%, cost = 1119 bps of capital staked
ROI on capital staked	-4.8%	per trade, fully collateralised
Kelly-sized \$ PnL	+33.75 on \$1,000	14 trades priced under \$0.10 made +241 \$
Sharpe (daily, annualised)	-2.80	calendar days incl. idle days, 1-share stakes
Max drawdown	11.44 \$ (1-share)	trough 2026-05-12
Positive days	13 of 35	best 2026-05-14 +3.79 / worst 2026-05-07 -3.05
Only positive slice	4h: +0.033 \$/trade	82 trades, WR 48.8% vs BE 44.4%

Recommendation: do not re-enable the 2-4 week paper run proposed on 2026-10-18; the replay already answers the question it was meant to answer. If anything is kept, it is a 4h-only variant, and that needs the remaining 79 days replayed and every trade re-graded on the real Polymarket outcome before it is worth more time (section 9).

1 Which markets it trades

Kink trades the same Polymarket **crypto Up/Down markets** as Momentum: rolling binary markets asking whether BTC, ETH, SOL, XRP or DOGE will be higher at the end of a fixed window than at its start. Each outcome token pays \$1 if right and \$0 if wrong. Kink scans every live window (5m, 15m, 1h, 4h, daily) but only **15m (239 trades) and 4h (82 trades)** passed its gates in the replay; the live paper run also took 15 trades on 5m markets on 7 May.

BTC dominates (256 of 321 trades), because the strategy takes only the single most mispriced market per scan and BTC's lower fixed volatility makes the model most confident on it. See paper 1 for the market structure, fees and the August 2026 switch to Chainlink TWAP settlement on 5m/15m/4h markets.

2 How the strategy works

Idea in one line: retail traders overpay for the side that has been moving; when the market's YES price sits well above a Black-Scholes fair value, buy the cheaper NO side and hold it to settlement.

Step 1 - fair value. For each live market, with S_0 the price at window open, S the current Binance price, τ the time left and σ a fixed annual volatility per coin (BTC 0.65, ETH 0.80, SOL 0.95, XRP 0.90, DOGE 1.05): **Theo_Yes** = $N(d_2)$, $d_2 = (\ln(S/S_0) - \sigma^2 \tau / 2) / (\sigma \sqrt{\tau})$, zero rate. Down markets use $1 - N(d_2)$.

Step 2 - detect the kink. Deviation = Live_Yes - Theo_Yes. Across all live markets, take the one with the largest deviation, and fire only if it exceeds **0.154**.

Step 3 - gates. At least 300 s to expiry; 24h volume of at least \$2,000; book snapshot under 300 s old; YES spread no wider than 0.08; and after costs, Theo_No $\times$ (1 - 2% fee) must still exceed the NO price paid.

Step 4 - sizing. Kelly on the NO side: $f^* = p - q/b$ with $p = \text{Theo_No}$ and $b = (1 - \text{price})/\text{price}$; stake $0.20 \times f^*$ of bankroll, capped at 1.5% (\$1,000 starting bankroll). Thin markets (24h volume under \$3,000) get a partial fill.

Step 5 - execution model. NO price = 1 - Live_Yes, plus 1.0c slippage, 1.2c adverse selection, 1.0c placement latency and 1.0c if the signal is more than 20 s after window open: **about 4.2c per share**. A win pays 1 - 2% fee. No early exit. Won/lost is decided from the Binance price at window end.

Operations. A 00:05 cron starts `strategy_kink.py` unless `state/kink.disabled` exists; it has been disabled since 2026-05-21. There is no order-placing code: everything here is paper.

3 Data and method

The replay runs `strategy_kink.py <date> unchanged` (config_kink.ini as deployed) on the recorder archives kept permanently on the T7 drive: the Polymarket book and trade CSVs and the Binance spot CSVs for each day. 123 dates have both archives (2026-04-19 to 2026-10-18). Each day is unpacked, replayed and deleted, four days at a time on shiv, in an isolated copy so that the live recorder and its logs are untouched.

Coverage. The run was stopped at the user's request after **44 days** (2026-04-19 to 2026-06-17). Three of those days (20-22 Apr) had no usable files, six more produced no signal, and **35 days traded**. 3 Jun and 16 Jun were still running at the stop and are excluded; together with the 77 dates after 17 Jun, 79 dates were not replayed, so this paper says nothing about the post-August TWAP-settlement regime.

April data is thin. The early archives hold about 1,300 Binance rows and 30 Polymarket tokens a day (the recorder was still being built), so April signals are few and noisy.

Units. One share per trade, so PnL is in \$ per \$1 of payout, matching the Momentum paper. Net = after the step-5 costs and fees, using the fill fraction of each trade. Gross = same trade at the raw NO price 1 - Live_Yes with no fee. Cost drag = (gross - net)/gross; cost bps = (gross - net)/capital staked. The Kelly-dollar column is reported only where noted, because it is dominated by a handful of cheap long-shot bets.

4 Backtest results

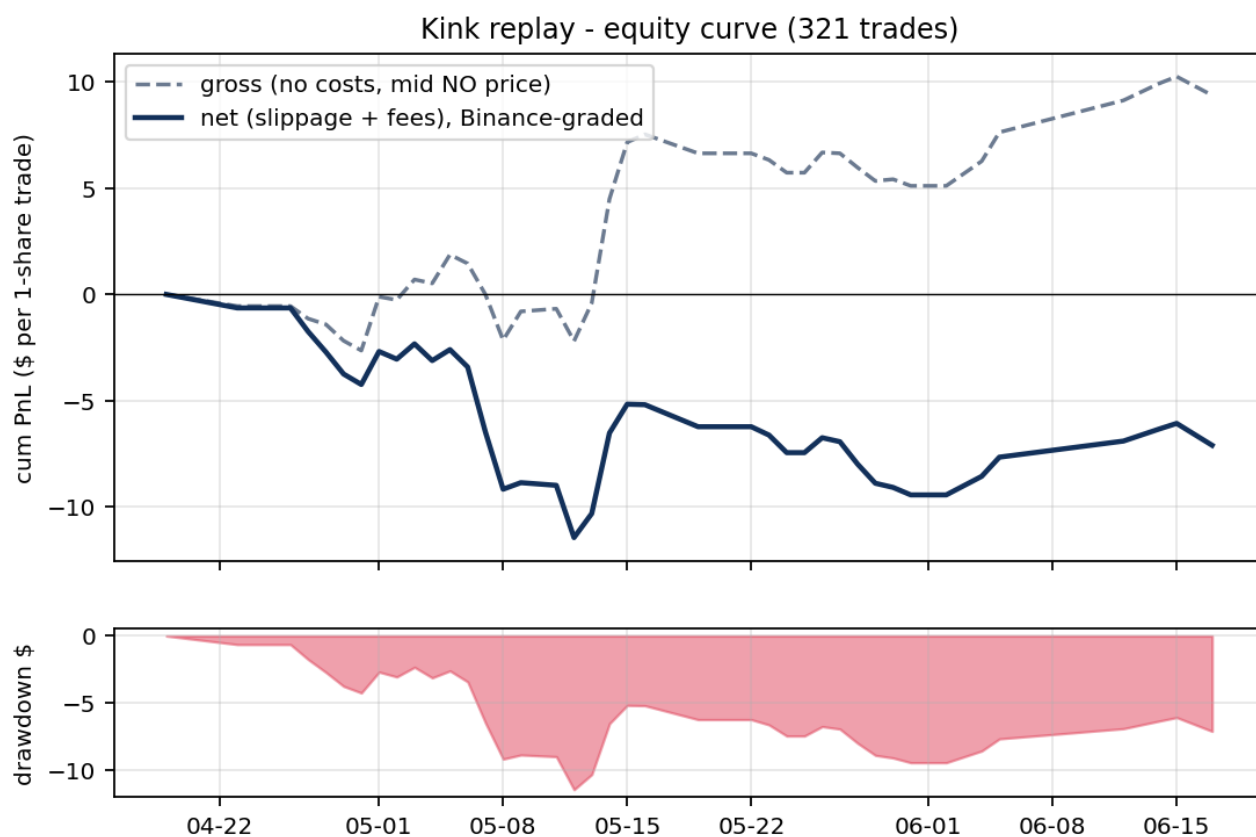


Fig 4.1 - Equity curve. Grey dashed = gross (no costs); navy = net. 1-share stakes, Binance-graded.

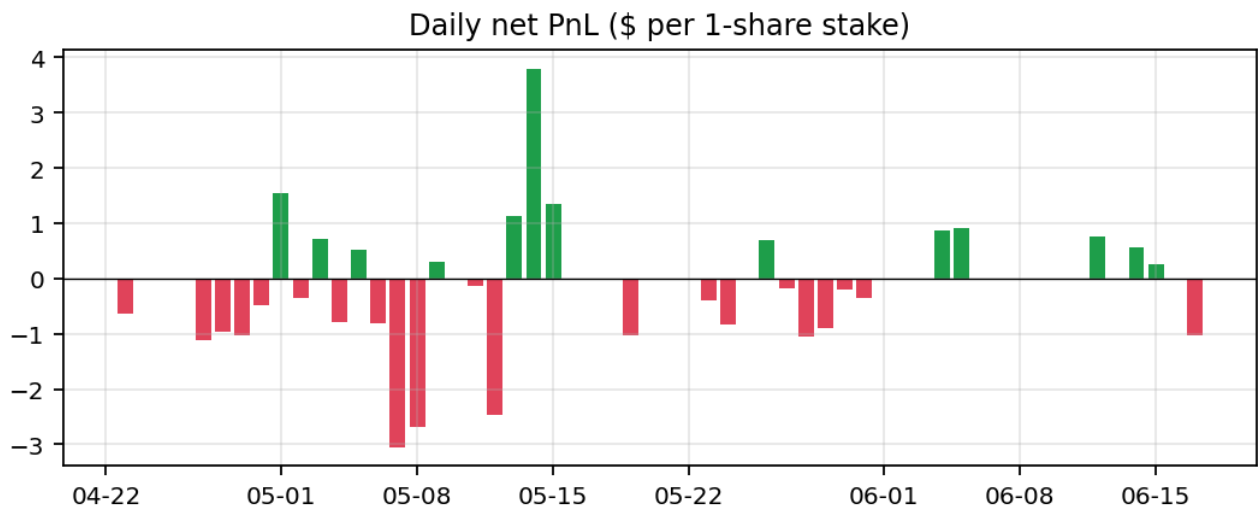


Fig 4.2 - Daily net PnL. 13 of 35 trading days were positive.

Month	Days	Trades	WR	Break-even	Net PnL	\$/trade	ROI	Kelly \$
2026-04	5	35	31.4%	43.1%	-4.24	-0.121	-28.1%	-204
2026-05	24	255	43.9%	45.1%	-5.19	-0.020	-4.5%	+154
2026-06	6	31	64.5%	55.0%	+2.33	+0.075	+13.7%	+84

June covers 1-17 Jun only, excluding 3 and 16 Jun (31 trades); the positive month is the smallest sample.

5 Where it wins and where it loses

By horizon. 15m loses; 4h is the only slice with a win rate clearly above break-even.

Group	Trades	WR	Break-even	\$/trade net	Net PnL	Gross PnL	Cost drag	Cost bps
15m	239	43.1%	46.3%	-0.041	-9.79	+2.33	520%	1095
4h	82	48.8%	44.4%	+0.033	+2.70	+7.04	62%	1191
ALL	321	44.5%	45.8%	-0.022	-7.09	+9.37	176%	1119

By coin. BTC and ETH both lose; SOL, XRP and DOGE have 12 trades between them, too few to read.

Group	Trades	WR	Break-even	\$/trade net	Net PnL	Gross PnL	Cost drag	Cost bps
BTC	256	44.1%	45.2%	-0.019	-4.92	+8.04	161%	1120
DOGE	1	100.0%	53.2%	+0.224	+0.22	+0.51	56%	5372
ETH	53	43.4%	46.0%	-0.036	-1.89	+0.85	322%	1124
SOL	6	33.3%	51.6%	-0.174	-1.04	-0.86	n/a (gross <= 0)	600
XRP	5	80.0%	67.8%	+0.106	+0.53	+0.82	35%	855
ALL	321	44.5%	45.8%	-0.022	-7.09	+9.37	176%	1119

By which token was overpriced. 'Up' = the Up token looked too expensive, so Kink bought its NO (a bet on Down). Fading an overpriced Up loses; fading an overpriced Down is flat. That is the opposite of a stable retail-FOMO effect.

Group	Trades	WR	Break-even	\$/trade net	Net PnL	Gross PnL	Cost drag	Cost bps
Down	162	45.7%	44.2%	+0.006	+0.97	+9.17	89%	1145
Up	159	43.4%	47.5%	-0.051	-8.06	+0.19	4235%	1094
ALL	321	44.5%	45.8%	-0.022	-7.09	+9.37	176%	1119

By NO price paid. Bands alternate between positive and negative with no pattern, as expected from noise. The long-shots under \$0.20 (48 trades, 8 wins) are positive per share, and Kelly sizing magnifies them into most of the dollar PnL because each win pays 5-10x.

Group	Trades	WR	Break-even	\$/trade net	Net PnL	Gross PnL	Cost drag	Cost bps
0.0-0.1	14	14.3%	7.8%	+0.062	+0.87	+1.50	42%	5762
0.1-0.2	34	17.6%	15.5%	+0.021	+0.71	+2.16	67%	2750
0.2-0.3	56	23.2%	24.9%	-0.021	-1.19	+1.42	183%	1870
0.3-0.4	45	24.4%	35.0%	-0.111	-4.99	-2.88	n/a (gross <= 0)	1335
0.4-0.5	38	42.1%	45.1%	-0.037	-1.39	+0.42	431%	1058
0.5-0.6	29	65.5%	54.8%	+0.087	+2.51	+4.34	42%	1150
0.6-0.7	35	57.1%	65.3%	-0.093	-3.25	-1.38	n/a (gross <= 0)	819
0.7-0.8	37	67.6%	73.9%	-0.078	-2.89	-0.78	n/a (gross <= 0)	771
0.8-0.9	33	93.9%	84.3%	+0.076	+2.51	+4.56	45%	738
ALL	321	44.5%	45.8%	-0.022	-7.09	+9.37	176%	1119

6 Is the model right, or the market?

Kink's thesis is that the Black-Scholes value is closer to the truth than the Polymarket price. Fig 6.1 tests it: for each bucket of predicted NO probability, the realised NO win rate. If the model were better, the navy line would sit on the diagonal and the red line (the market's own price) off it.

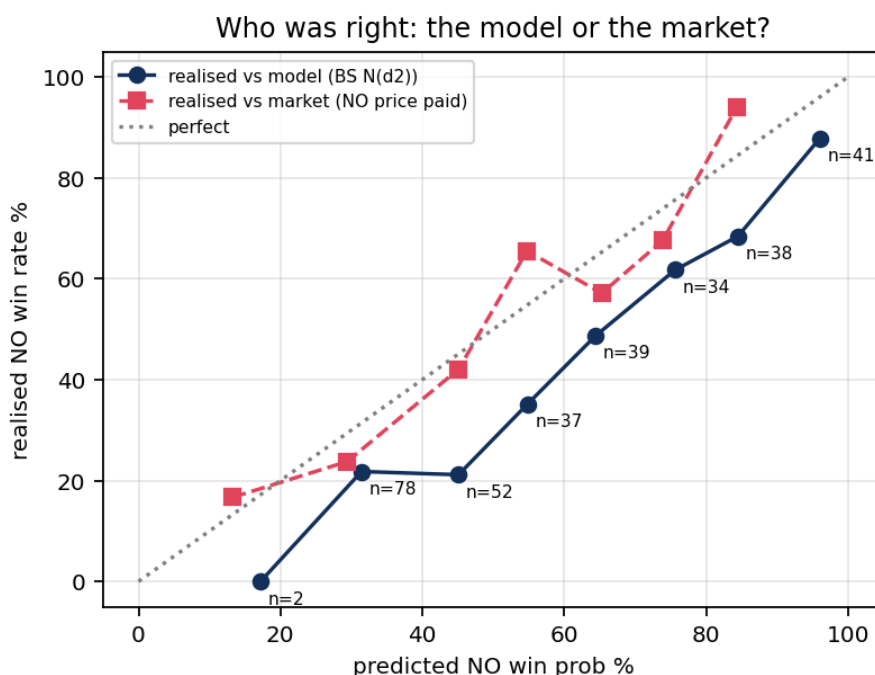


Fig 6.1 - Calibration of the model (navy) and of the market price (red). Below the diagonal = overconfident.

Reading. The model is overconfident on the NO side: its high-probability buckets win less often than it says. The market price tracks realised outcomes at least as well. Three reasons the 'kink' is mostly not mispricing:

- **Fixed volatility.** sigma is a constant per coin. When realised volatility is higher than that, a large move since the open is less decisive than $N(d2)$ claims, and the market (which sees current volatility) is right to price the trailing side higher.
- **Thin tails.** A Gaussian underprices reversals in crypto, the same overconfidence found in Momentum.
- **Selection on the largest deviation.** Picking the single biggest gap across all markets each scan selects the markets where the model is most wrong, not where the market is.

7 Live paper record vs replay

The live paper run (May 6-21) recorded 316 trades at **+0.114 \$/share** (WR 50.0% vs BE 37.6%), and was the reason Kink was called 'most promising' on 2026-10-18. The replay of the same days does not reproduce it:

Date	Live trades	Live \$/share	Replay trades	Replay \$/share	Same market+side
2026-05-06	3	-0.028	8	-0.102	0
2026-05-07	164	+0.129	31	-0.098	15
2026-05-08	111	+0.117	12	-0.225	5
2026-05-13	6	-0.065	12	+0.094	0
2026-05-14	22	+0.028	21	+0.180	2
2026-05-21	10	+0.174	not replayed	-	0

- **7-8 May are the whole live record:** 275 of 316 live trades and most of its profit. The replay takes 43 trades on those two days, overlaps with only 20 of the live ones, and loses on both days.
- **The cause of the gap was not investigated.** Candidates: the live loop reads the recorder's hot CSVs every 5 s while the replay reads the archived files after the fact. Either way, the live record is a 6-day sample that the recorded data does not confirm.
- 21 May has no T7 archive, so it could not be replayed.

8 Costs and caveats

Costs are the deciding factor. Gross edge +2.9c/share against costs of 5.1c/share (1119 bps of capital staked). Even with zero slippage the 2% fee on wins alone (about 1c/share) would leave the edge marginal. The gross figure is itself optimistic: it uses 1 - Live_Yes, not the NO ask, so the spread is not charged.

- **Wrong settlement source.** Trades are graded on the Binance end price; Polymarket settles on Chainlink (TWAP from Aug 2026). In Momentum this inflated PnL by about 5-8%; real numbers are likely a little worse.
- **Partial history.** 44 of 123 dates; nothing after 17 Jun, and 3 and 16 Jun missing.
- **Fill realism.** The cost model is a fixed slippage add-on, not a walk of the recorded order book (fill_sim.py was built for Momentum, not Kink).
- **Kelly sizing hides the result.** The +33.75 \$ Kelly PnL is 14 long-shot trades; per share the strategy loses.

9 Next steps

#	Action	Effort	Decides
1	Keep Kink disabled; drop the proposed 2-4 week paper re-enable	none	-
2	Optional: replay the remaining 79 dates (shiv ~/kink_bt, run_day.sh resumes)	~3 h machine time	whether 4h holds up
3	If 4h holds: re-grade on the real outcome (verify_resolutions.py, needs column adaptation)	half a day	settlement-source effect
4	If still positive: TWAP-aware fair value + realised vol in place of fixed sigma	days	whether a kink exists at all

Files: shiv ~/kink_bt/kink_trades_all.csv (all trades), out/kink_<date>.err (per-day replay logs).