Backtest Audit · Written Verdict

SAMPLE REPORT

Ref SV-000 (sample) · 22 July 2026 · prepared by Ayan Jain · finance-broski.github.io

> This is a sample. The subject below is a real strategy from my own research program, audited with the same pipeline and reported in the same format a paying client receives. Every number is from committed, rerunnable scripts. Only the subject changes in a client report.

SUBJECT AS RECEIVED

Strategy	Illiquid-tail reversal, NSE equities. Rank the full panel by trailing 60-day median turnover; in the least-liquid quintile, buy the worst 5-day performers and hold 5 days. Filters: price above ₹5, traded at least 200 of the trailing 252 sessions.
Claimed result	**+28.8%/yr** excess, **t = 11.65**, monotone across liquidity quintiles. 2011-2026, 3,819 names, 4.69M eligible name-days.
Materials	Rules, code and panel (equivalent of a full intake folder). Reconstructed and rerun independently; cross-checked against my survivorship-correct panel.

Reproduced, not tradeable.

2 PASS **2 FLAG** **2 FAIL**

The effect is statistically real. It reproduces a published academic result on independent data, survives every confound test with a significant residual, and was not manufactured by parameter search. It still cannot be collected: the claimed **+28.8%/yr** falls to **+17.9%/yr** once entries use a price a real order could get, and goes negative once it pays the trading costs of its own liquidity bucket: **−2.4%/yr** already at 40bps round trip, **−32.6%** at 100bps, against a bucket that realistically costs 100-300bps. The account that tries to harvest this edge pays more to trade it than the edge pays to hold it.

THE SIX CHECKS

01 Survivorship — FLAG

The panel itself is survivorship-correct (delisted names present). The test is not: positions whose price never prints again are silently dropped. Of **182,578** illiquid loser-positions, **3,107** have no price 5 days later; **517** (0.28%) are true deaths, the rest are illiquidity gaps that resume.

Measured band: deaths written to −100% cut the claim to **+14.5%/yr**; at −50%, **+21.7%/yr**.

FIX · assign an explicit delisting return to every position that stops pricing, and report the result as a band, not a point.

02 Look-ahead / point-in-time — FLAG

No calendar look-ahead found. The leak is finer: the entry uses the same stale trade print that generated the signal, a price no real order could have. Inserting one day between signal and entry cuts **+28.8%** (t 11.65) to **+15.1%** (t 6.27). Entering at the next day's open keeps **+17.9%** (t 7.22). A price-level control shows the residual is real rather than pure bid-ask bounce: if bounce were everything, cheap stocks would show ~83x the effect of dear ones; the observed ratio is **1.28x**.

FIX · compute signals on tradeable prices (next open), never on the signal print itself. Roughly 38-47% of the headline was the print artifact.

03 Cost realism — FAIL

The 5-day hold makes **50 round trips a year**, and the bottom-decile names it trades cost 100-300bps per round trip all-in before market impact. On the tradeable-print basis: **+7.7%/yr** at 20bps, **−2.4%** at 40bps, **−32.6%** at 100bps. Breakeven is ~36bps round trip; the bucket does not trade there. The 21-day variant cuts trips fourfold and still nets **−3.7%/yr** at 100bps.

FIX · none available at this bucket. Costs must be charged per liquidity bucket, never as one flat number; the flat 20bps assumption hid the whole result.

04 Overfitting / selection — PASS

Pass and kill criteria were preregistered before the run, with the significance bar raised to t ≥ 3.0 to account for ~150 prior constructions on the same panel. The parameter grid is small and fully disclosed (signal 5/21 days × hold 5/21 days), and the effect is monotone across a dimension the search never touched, which parameter mining does not produce.

NOTE · a pass here does not rescue the strategy. It means the number was measured, not manufactured. The strategy dies on checks 03 and 06 anyway.

05 Statistical validity — PASS

t = 11.65 on 4.69M name-days across 3,819 names; monotone across all five liquidity quintiles on two constructions; after the harshest confound correction the residual is still t ≥ 3.7. The edge is comfortably distinguishable from luck. This is what makes the verdict useful rather than dismissive: real and unreachable are different failures, and this is the second one.

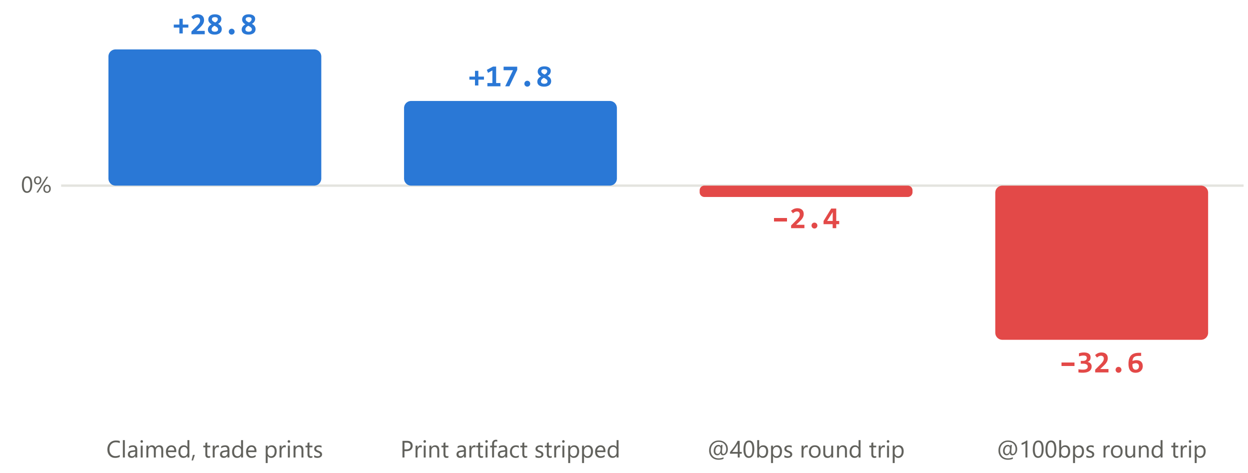
NOTE · it matches Chui, Ranganathan, Rohit & Veeraraghavan (2023), who find Indian reversal concentrated in exactly this illiquid tail.

06 Fill feasibility — FAIL

The median bottom-decile name turns over **₹21,838 per day** (median price ₹45). A ₹1 lakh position is 4.6 full days of that volume; at a 5% participation cap it takes ~92 sessions to build, against a 5-day hold. **13.1%** of entries follow a ≥4.5% prior-day move (9.5% in liquid names), so circuit locks sit on the entry path. Deployable capital at honest participation: effectively nil.

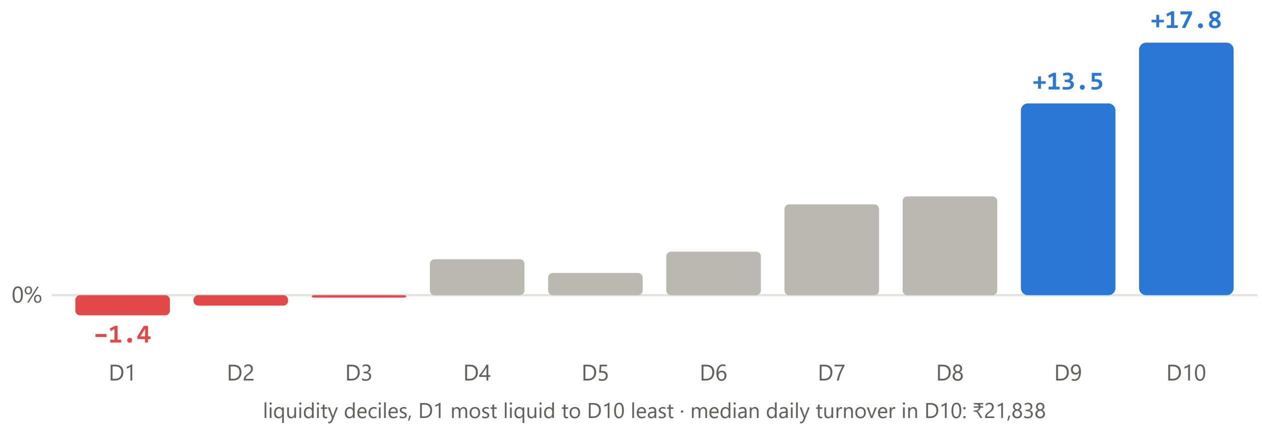
FIX · re-simulate with a participation cap and limit-order fills. One decile up (D9, ₹1.3 lakh/day) the same math still fails: gross +13.5% against its own realistic cost.

What the claim is worth at each correction



%/yr excess return. Cost bars charge the strategy's 50 round trips a year on the artifact-stripped decile-10 basis; the bucket's realistic all-in cost is 100-300bps per round trip.

The effect is real, and it lives exactly where trading is dearest



Gross %/yr with one day between signal and entry (print artifact stripped), 5-day signal / 5-day hold, 2011-2026.

AGAINST THE INTAKE

In a paid Check this table quotes your own intake answers, and it is the refund baseline: if the verdict does not go beyond what you wrote, the fee comes back. The intake column here is illustrative.

INTAKE SAID (ILLUSTRATIVE)	MEASURED
"20bps flat should cover costs"	Bucket breakeven ~36bps round trip; realistic 100-300bps; net −2.4% to −32.6%/yr.
"No look-ahead that I know of"	38-47% of the headline is the entry print itself; +28.8% becomes +17.9% at next-open.
"Universe includes delisted names"	True for the panel, not the test: 3,107 unpriced exits silently dropped; band +14.5% to +21.7%.
"Tried 4 variants"	Consistent with the disclosed grid; monotonicity across an untouched dimension supports it.

PRIORITISED FIXES

1. Charge each trade its own bucket's cost. The flat 20bps assumption concealed a −30pp/yr swing; this one line reverses the verdict of the whole backtest.
2. Move every entry to the next day's open. Removes the untradeable-print artifact and restates the honest gross at +17.9%/yr.
3. Assign an explicit delisting return to positions that never price again; report the result as the +14.5% to +21.7% band.
4. Re-simulate with a 5% participation cap and limit-order fills before believing any size number. At the median D10 name, ₹1 lakh is 92 sessions of building.
5. If the idea must survive, test it one decile up at D9 with the same discipline. It fails there too on current numbers; that result is worth having in writing.

WHAT WOULD CHANGE THIS VERDICT

An execution channel at or below ~36bps all-in round trip in the bottom decile (none exists for a taker; a patient limit-order version is a different strategy with a fill-selection bias, and needs its own audit), or the effect clearing its measured breakeven in a participation-capped simulation at D8-D9. This verdict is about this construction on this data, not about reversal in India generally.

SCOPE AND TERMS

This document is the deliverable of a Bias Check: ₹7,500 in India, $250 international, written verdict within 48 hours of payment and a complete folder. Guaranteed: if the verdict does not go beyond your intake answers, the fee comes back. This is an audit of software and data quality. It is not investment advice and not a prediction that any strategy will make money. I audit the process; the market grades the strategy.

ABOUT THIS SAMPLE

The subject is from my own research program: a reproduction of a published anomaly on an independent survivorship-correct panel, chosen because it fails in an instructive way. The verdict format, checks, evidence standard and guarantee mechanics are identical in a client report. Receipts and method notes: finance-broski.github.io · pip install backtest-bias · financebroski.substack.com