Buying new Solana memecoins after launch lost money in every way we tested

Results of the Solana Memecoin Lab, 17-26 Oct 2026.

A forward paper-trading study of every tradable new Solana token seen over nine days (1,411 tokens from 91,657 new pools), with four pre-registered secondary studies, a held-out exploratory search, 416 alternative trading strategies, and a study of the top 10% of winners across all discovered tokens. All times are UK time. All returns are net of modelled fees, slippage and network costs unless marked raw.

Report date: Monday 26 October 2026 (UK time)

Contents

Abstract

1. Background and hypotheses
2. Methods
3. Data and coverage
4. Results: buying every tradable new token
5. Secondary studies
6. Exploratory edge search
7. The funnel: from 91,657 pools to 1,411 tradable tokens
8. Every trading method tested on the 1,411
9. The top 10% of winners across all discovered tokens
10. Discussion
11. What would change our conclusion
12. Limitations
13. Conclusion

Abstract

1,411	-61.8%	-74.2%	66.6%	1.3%
tradable new tokens from 91,657 new pools	**mean net return at 6 h** 95% CI -68.0% to -53.4%, n = 1,187	**at 6 h, likely honeypots excluded** 95% CI -80.8% to -65.1%, n = 1,061 (section 4.4)	**dead within 6 h** 790 of 1,187	**doubled or better at 6 h** 16 of 1,187, after costs

Objective. To measure what an ordinary buyer earns by buying newly launched Solana memecoins about 30 minutes after their first liquidity pool appears, using realistic retail trading costs.

Design. A forward paper-trading experiment ran from 17 to 26 October 2026. Every new Solana DEX pool seen on GeckoTerminal and Jupiter was enrolled (91,657 in the 09:37 database copy used for the results; 91,438 in the 09:02 weekly report). Every token tradable at the 30-minute decision point (liquidity of at least $5,000 and at least 20 transactions an hour; 1,411 tokens) was bought on paper at $25, with modelled fees, slippage, price impact and latency, and followed for 1, 6 and 24 hours. Four pre-registered secondary studies (a pump.fun launch filter, an on-chain rug check, an established-token plus X-sentiment strategy, and an X promotion check) and an exploratory search on a 70/30 development/held-out split ran over the same period. Two further analyses use the same data: 416 alternative trading strategies on the tradable universe, with Bonferroni correction, White's reality check and held-out testing; and a study of the winners among all 67,734 discovered tokens with a tradable price history, including chronologically validated classifiers that try to pick winners before purchase and a post-hoc look at the 143 tokens that reached 10x, using 1-minute candles, launch-second on-chain trades and matched look-alike controls (section 9.9).

Results. Buying the 1,411 tradable launches at about 30 minutes lost a mean of 31.4% after 1 hour (n = 1,293; 95% CI -37.1% to -24.9%), 61.8% after 6 hours (n = 1,187; CI -68.0% to -53.4%) and 55.7% after 24 hours (n = 760; CI -68.4% to -38.9%), in both chronological halves. The median token lost the whole stake at 6 and 24 hours, two-thirds were dead within 6 hours, and only 16 of 1,187 (1.3%) had doubled after costs. The loss is not a cost artefact (-59.7% at 6 h with no fees, slippage or network cost; -59.4% at raw prices, CI -66.3% to -50.0%). Nor is it an artefact of collector gaps (-65.3% with tokens already dead before a gap counted as -100%; -60.9% if live unobserved tokens are also marked at their last price). Excluding the likely honeypots of section 9.7 makes it larger (-74.2%, CI -80.8% to -65.1%; -77.0% counting them as -100%). A rule exit (+100% take-profit / -35% stop / 6 h) lost 10.8%, or 25.8% without one token that jumped about 198x between snapshots. None of the secondary studies made money, and none of roughly 240 exploratory development/held-out cells was positive on both halves with a usable sample. Of 416 alternative strategies, none had a significantly positive mean even uncorrected, 182 of the 311 with at least 30 trades were significantly negative after Bonferroni correction, White's reality check gave p = 0.35 for the best one, and the best on development data (+85.2%) lost 81.2% held-out (n = 27, CI -94.7% to -58.0%). Of the 67,734 discovered tokens with a tradable price history, 84% never rose above their first observed price; buying all of them lost 12.6% at 1 h and 15.6% at 6 h (n = 59,845 and 52,348; p below 10^{-200}; -13.0% and -17.0% with dead-before-gap tokens counted). Classifiers picked out future movers on later, unseen tokens (AUC 0.87 for the top 10%, 0.95 for 2x), but their top picks lost 25-72% at 6 h (33-80% with dead-before-gap tokens counted), every CI below zero, worse than random picks. A sellability check found 155 likely honeypots among the top-decile winners, 126 of them one buy-only cluster (hundreds of buys an hour, zero sells ever) that RugCheck did not flag; without them the 2x group's +26.2% at 6 h is not a demonstrated gain (-5.6%, CI -19.6% to +10.0%; -12.0% to +12.8% under the censoring checks), and the top-10% model's top 1% of picks loses 82.3% at 6 h instead of 59.4%. All 143 tokens that reached 10x had observed sells and none was flagged; a post-hoc look (section 9.9) found one rule that picked them at 13.6% held-out precision with an unproven profit, and their moves peaked at a median of 86 minutes and lasted minutes. Late entries are the exception to a proven loss: entries 6 hours after the decision were not significantly negative under the lab's costs (-3.6% to +1.0%, CIs spanning zero; -3.9% to -0.9% with dead-before-gap tokens counted) and, at low or zero cost, slightly positive in the full sample but mostly negative held-out, so they show no proven edge rather than a proven loss. Section 11 sets out the coverage, sample size and live-fill evidence that would change that.

Conclusion. Systematic post-launch buying of new Solana memecoins, entering 30 minutes or more after launch at retail cost, is proven to lose on average, and no filter, timing rule, exit rule or winner-prediction model tested here

changes that out of sample. Real winners existed, but buying the tokens that looked most likely to win lost money, and even the 10x tokens could not be picked profitably in advance (section 9.9). Section 11 sets out what evidence would change these conclusions.

Scope. Not tested: sniping in the first seconds, insider or creator supply, fee income, CEX listings, established memecoins, other chains or other months. Only about 25% of the 215-hour window was observed (about 53 h, mostly 18-20 Oct), in one market regime, and all fills are simulated from 5-10 minute snapshots, not real trades.

1. Background and hypotheses

Thousands of tokens are launched on Solana every hour, mostly through pump.fun and similar launchpads, and some of them rise 10x or 100x. The question was what an ordinary outsider earns by buying these tokens once they are tradable, after realistic costs, and whether any public-data rule improves on that. The lab measured every tradable launch with a fixed cost model set before the data was collected, so the result could not come from tuning after the fact.

ID	Hypothesis	Study	Status
H1	Buying every tradable new Solana token at about 30 minutes of age earns a positive net mean return at 1, 6 or 24 h	Lab v1.1, full universe	Fixed in advance (buy-all comparison group)
H2	pump.fun launches with FDV of at least $8k at 10 min, from a first-time creator and not bundled, are profitable with a +100% / -35% / 6 h exit	pumpfun-forward	Pre-registered
H3	Established tokens (liquidity of at least $50k, age at least 2 h) that pass a rug filter and an X scout rating of HOT are profitable with a scale-out exit	alpha404-test	Pre-registered
H4	Replacing RugCheck with direct on-chain checks gives the same or better selection	rugalt v1.2	Pre-registered
H5	Paid or coordinated X promotion predicts returns	X promotion check	Exploratory

Table 1. *Section 1. Background and hypotheses*

Whether winners can be picked before purchase is examined in section 9, not as a hypothesis fixed in advance: classifiers on a chronological split (9.6) and a post-hoc look at the 10x tokens (9.9). Section 11 sets out what evidence would change the conclusions.

2. Methods

2.1 Data sources

Source	Use	Polling / budget
GeckoTerminal new_pools	Discovery of new Solana pools	2 pages every 30 s; 20 req/min
Jupiter tokens/v2/recent	Discovery of new tokens	Every 90 s; 15 req/min
DexScreener pairs	Price, liquidity, volume, transaction counts, socials	150 req/min
RugCheck	Mint and freeze authority, holders, insiders, LP lock, risk flags	40 req/min
Solana RPC (pump.fun study)	Creator history, bundling, bonding-curve state	116,693 calls, 8,080 errors
X (from 19 Oct)	Posts mentioning a token's contract address or ticker	69 tokens enriched
GeckoTerminal 1-minute OHLCV (section 9.9 only)	Candles from launch for indicators and peak timing	143 10x tokens, 286 look-alikes, 323 of a planned 700 random tokens
Solana public RPC (section 9.9 only)	First 30 successful transactions of each pump.fun token from creation (launch-second buyers)	Decoded for 78 winners, 166 look-alikes, 52 random pump.fun tokens

Table 2. *Section 2.1 Data sources*

2.2 The tradable universe (lab v1.1)

Pools were admitted if they were under 90 minutes old when discovered. A decision was made at 30 minutes of pool age (or as soon as possible after that, up to 120 minutes; after that the token was recorded as missed). The tradable universe was every token with at least $5,000 of liquidity and at least 20 transactions an hour at decision time. Every universe token was bought on paper, with no further filtering, which is what an ordinary buyer of new launches would get on average.

Why the decision point is 30 minutes. The frozen configuration sets the minimum decision age at 30 minutes to avoid the sniper window, where insiders and bots buy and then sell to later buyers (one published analysis found most deployer-funded snipes exit within 1 to 5 minutes) and a public-data outsider cannot compete. Waiting lets a token survive or fail its first dump, gives time for RugCheck holder and authority data and a trading record to exist, and matches what a careful retail buyer would do. Other entries were also tested: delays of 15 minutes to 24 hours after the decision (section 8, none significantly profitable) and the first observed price, a median of about 5 minutes after the pool appeared (section 9, losses with p below 10^{-180}).

Why the other ~90,000 tokens are not in the universe. Most new pools never get real liquidity or trading: 85,232 enrolled pools were below the gate at decision time, and 4,663 could not be evaluated within 120 minutes. Most of the below-gate tokens (79,166 in the 09:37 copy) were still on a launchpad bonding curve with no pool at all. Those could be bought on the curve, but they have no pool liquidity, so they fall outside this universe's definition; the rest had pools with under $5,000 of liquidity or almost no trading. They are not ignored: the full step-by-step funnel is in section 7, a random sample of 422 below-gate tokens was followed for 24 h (section 5.5, 72% dead), pump.fun curve tokens are covered by the pump.fun study (5.1), and every token with a tradable price history, curve tokens included, is scored in the winners study (section 9) using a bonding-curve liquidity model.

A separate pre-registered test of a filter-based selection within this universe was also run; its full results are reported in a separate appendix (appendix_filter_test.md).

2.3 Cost model and outcome measurement

Parameter	Value
Position	$25 per trade, paper only
Fees	1.25% per leg: the venue fee on pump.fun's bonding curve and small PumpSwap pools (other DEXes charge less; trading-bot fees of about 1% are not included). Sensitivity to this choice is in section 4.4.
Slippage	1% adverse per leg, plus constant-product price impact against the pool's liquidity
Network	$0.10 per transaction
Latency	20 s from decision to fill
Tracking	5-minute snapshots; 900 s tolerance for matching a horizon
Dead token	Liquidity under $500, price at 10% of entry or less, or pair removed; scored as -100%
Horizons	60, 360 (primary) and 1,440 minutes; also a rule exit at +100% take-profit / -35% stop / 360 min, checked on snapshots

Table 3. *Section 2.3 Cost model and outcome measurement: simulation parameters.*

2.4 Pre-registration

Each study wrote its rules, pass criteria and the hash of its frozen configuration or protocol before the forward data it would be judged on was collected. Data from the v1.0 pilot is kept separately and is never pooled with v1.1.

Study	Registered (UK time)	Hash
Lab v1.0 pilot	Fri 16 Oct 18:58	config.toml 36f1271c5294
Lab v1.1	Sat 17 Oct 11:12	thresholds 36f1271c5294 (unchanged); overlay config_v1_1.toml 8fcd269638cf; scoring.py md5 bee33118…; sim.py md5 061fc0eb…
pumpfun-forward v1.1	Mon 19 Oct 09:20	sha256 5e665693…f3a5c5 (replaces v1 392940ef…)
rugalt v1.2	Mon 19 Oct about 11:50	sha256 cdca6b40…7bf7d354d

Study	Registered (UK time)	Hash
alpha404-test	Mon 19 Oct 13:30	sha256 b8349786…dbecbd

Table 4. *Section 2.4 Pre-registration*

Changes from v1.0 to v1.1. The v1.0 pilot capped enrolment at 400 new pools an hour. Actual flow was about 2,000 an hour, so the cap filled about 12 minutes into each hour and about 80% of tokens were thrown away. v1.0 ended after 15.98 hours with 38 universe tokens and 1 buy (verdict: insufficient). v1.1 removed the cap (pending queue up to 1,500). It also added logging-only features: hourly follow-up of skipped tokens for 24 h, or 7 days for larger ones; a 6-hourly snapshot of "top young tokens" lists; identity features; multi-age decision features at 10, 30, 60 and 180 min with a 5-minute price path to 180 min; and, from 19 Oct, X mention logging. Selection thresholds and scoring did not change.

2.5 Statistics

- Mean net return per $25 trade at each horizon, with a 95% bootstrap CI (4,000 resamples, the lab's own report code and seed).
- Median, win rate (share of trades above zero), dead rate, and the mean excluding the single best trade, to show how far one outlier drives the mean.
- Chronological halves: universe tokens sorted by decision time and split at the midpoint, to check the result is not driven by one period.
- A horizon outcome is only counted where the collector recorded a price within 900 s of that horizon; horizons that fell in collector gaps are excluded in the headline figures, which is why there are 1,293, 1,187 and 760 outcomes from 1,411 tokens. Censoring check, applied to every fixed-hold figure in sections 4, 8 and 9: a token whose last observation before a missing horizon was already dead (no pair, price at 10% of entry or less, or liquidity under $500) is scored -100%, and a token still alive at its last observation is either excluded or marked at that last price; both versions are reported with their counts (sections 4.4, 8.2, 9.4 and 9.6). Rule exits keep the full-coverage rule of section 8.
- Universe statistics were recomputed for this page with the lab's own report code on a copy of the database taken Mon 26 Oct 09:37 UK time (1,411 decided tokens; the 09:02 weekly report had 1,407).

2.6 Exploratory analysis design

After the pre-registered tests, an exploratory search re-ran alternative entry rules on copies of each study's database. It used the same cost model and a time split: the first 70% of entries by time were for discovery (DISC) and the last 30% were held out (HOLD). A rule only counted as a candidate edge if it was positive on both. About 240 DISC/HOLD cells were evaluated, so some positive cells are expected by chance (see section 6). The 10x deep dive (section 9.9) was added on Mon 26 Oct and is labelled post-hoc: its rules, thresholds and indicators were chosen after seeing the data. Each 10x token was paired with two matched look-alikes that did not reach 10x (nearest neighbour at first observation on FDV, buys, sells, 10-minute return and calendar day, with exact curve status), plus a random sample of scorable tokens; rules were tested on a 70/30 time split with Bonferroni correction.

3. Data and coverage

Measure	Value
Candidates enrolled (= discovered; coverage 100% of what was seen)	91,438 at the 09:02 report (91,657 in the 09:37 copy used for the results)
Not in universe (below the liquidity or activity gate)	85,232
Missed (no decision possible within 120 min)	4,663
Still new at report time	136
Tradable universe tokens (09:02 report / 09:37 database copy)	1,407 / 1,411
Universe tokens with an outcome at 1 h / 6 h / 24 h	1,293 / 1,187 / 760

Measure	Value
Decided with no usable market data (NO_DATA; counted as not tradable)	1,105
Price snapshots on followed-up tokens	1,000,799
Calendar window (first to last candidate)	214.96 h (17 Oct 11:11 to 26 Oct 09:09)
Collector gaps logged	144 gaps, 161.6 h in total
Hours actually observed	about 53.3 h (25% of the window)

Table 5. *Section 3. Data and coverage*

The coverage problem. The collector ran on a virtual machine that froze or was suspended when idle, so data was only collected for about 53 of the 215 hours. Almost all of the observed time fell on 18-20 October, and 81,524 of the 91,438 candidates in the 09:02 report (89%) and 1,317 of the 1,411 tradable tokens were discovered on those three days. The longest gaps were 21 Oct 16:49 to 22 Oct 18:40 (25.9 h), 24 Oct 18:21 to 25 Oct 07:29 (14.1 h), 20 Oct 20:55 to 21 Oct 10:29 (13.6 h), 17 Oct 19:34 to 18 Oct 07:49 (12.3 h) and 25 Oct 19:19 to 26 Oct 07:33 (12.2 h). As a result the sample is much smaller than planned, and it mostly reflects market conditions on three days.

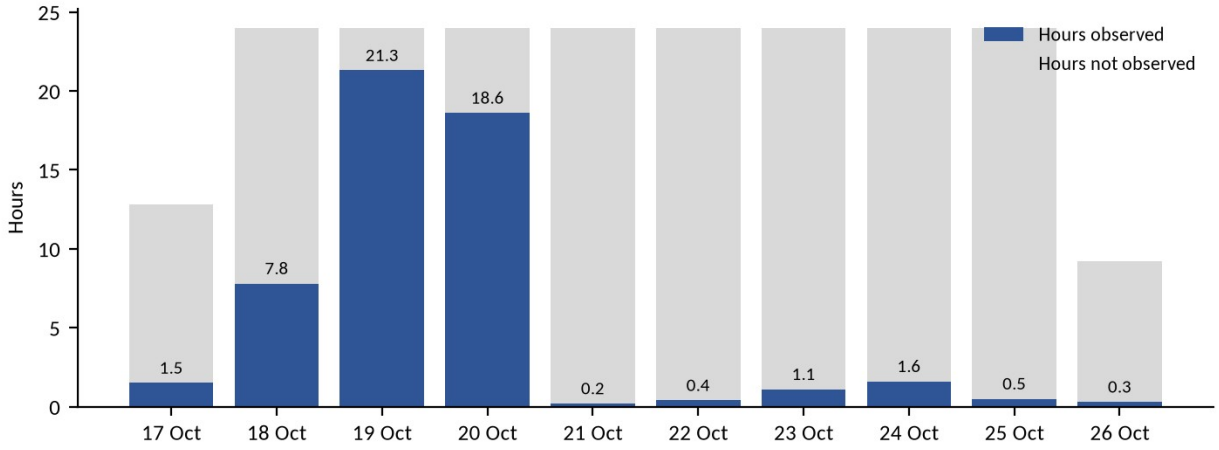


Figure 1. *Hours observed and not observed by the collector on each day, 17-26 Oct 2026 (UK time). Values as shown on the study Page.*

New candidates discovered per day (UK days):

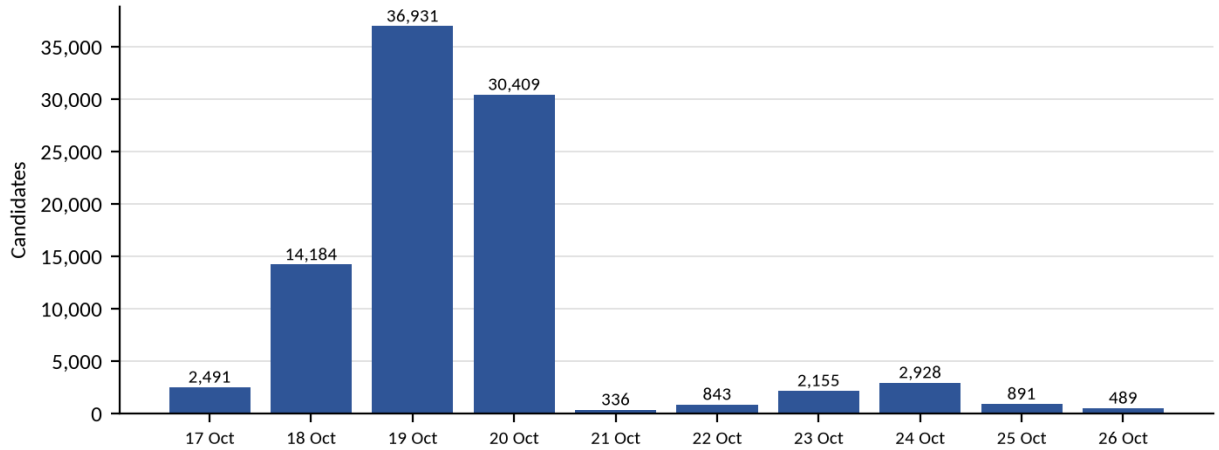


Figure 2. *New candidates (pools) discovered per day, 17-26 Oct 2026. Values as shown on the study Page.*

Day	Candidates	Tradable universe tokens	Hours observed
17 Oct	2,491	30	1.5
18 Oct	14,184	190	7.8

Day	Candidates	Tradable universe tokens	Hours observed
19 Oct	36,931	612	21.3
20 Oct	30,409	515	18.6
21 Oct	336	0	0.2
22 Oct	843	3	0.4
23 Oct	2,155	28	1.1
24 Oct	2,928	23	1.6
25 Oct	891	5	0.5
26 Oct	489	5	0.3

Table 6. *Section 3. Data and coverage*

Other data-quality events. RugCheck could not be reached from about 23:00 on 18 Oct (completely from 01:27 on 19 Oct) until about 10:56 on 19 Oct. Across the run, 17,337 RugCheck lookups due at 30 minutes were skipped because of backlog. The log has 34 error events, all "database is locked" in the rug worker. On tokens flagged for long follow-up, the recorded FDV values are implausible (trillions of dollars), which points to data artefacts, so those values were not used.

4. Results: buying every tradable new token

Recomputed with the lab's report code on a database copy taken Mon 26 Oct 09:37 UK time

4.1 Returns by horizon

Horizon	n	Mean	95% CI	Median	Win	Dead	Mean ex-top trade	Halves (1st / 2nd)
1 h	1,293	-31.4%	-37.1 to -24.9	-81.4%	28%	49%	-33.3%	-35.5% / -27.3%
6 h	1,187	-61.8%	-68.0 to -53.4	-100%	18%	67%	-65.1%	-54.2% / -69.4%
24 h	760	-55.7%	-68.4 to -38.9	-100%	16%	69%	-61.6%	-38.2% / -73.3%

Table 7. *Section 4.1 Returns by horizon*

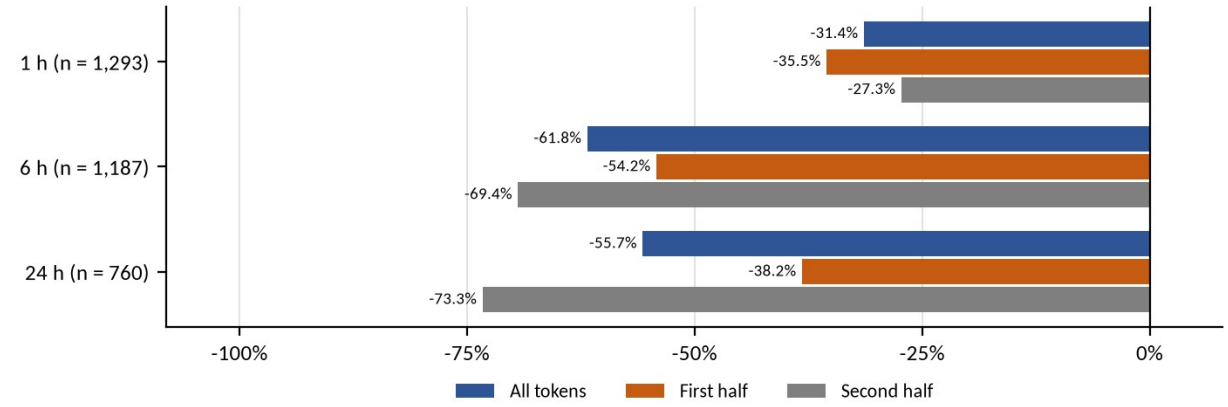


Figure 3. *Mean net return by horizon for all tradable tokens and for each half of the study period. Values as shown on the study Page.*

Removing the ten best trades gives -36.9%, -68.1% and -71.7%. In money terms, $25 on each of the 1,187 tokens with a 6 h outcome is $29,675 staked and a net loss of $18,349 after 6 hours. Fewer 24 h horizons fell inside observed hours, hence the lower n. On the 720 tokens with both outcomes the mean is -57.7% at 6 h and -54.3% at 24 h (a few late winners) and the median is -100% at both, so holding longer does not rescue the typical token.

4.2 What the distribution looks like

Net return at 6 h, number of tokens in each band (n = 1,187):

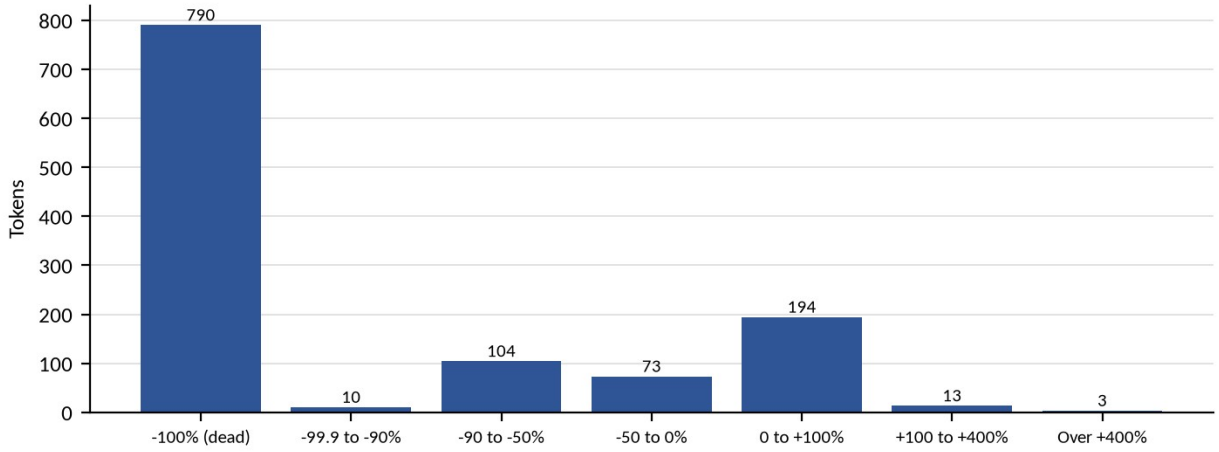


Figure 4. *Distribution of net returns at 6 h for the tradable universe (number of tokens per return band, n = 1,187). Values as shown on the study Page.*

Net return band	1 h	6 h	24 h
-100% (dead)	631	790	526
-99.9% to -90%	2	10	2
-90% to -50%	92	104	78
-50% to 0%	203	73	35
0% to +100%	279	194	107
+100% to +400%	79	13	5
Over +400%	7	3	7
Total	1,293	1,187	760
At least 2x after costs	86 (6.7%)	16 (1.3%)	12 (1.6%)
At least 10x after costs	1	1	4
Best single trade	+2,408%	+3,775%	+4,351%

Table 8. *Section 4.2 What the distribution looks like*

This is a lottery-ticket distribution: most tokens go to zero, a minority drift around break-even, and a handful of real but rare large winners cannot pay for the losses.

4.3 Using a take-profit and stop-loss instead

Rule exit (+100% / -35% / 6 h)	Value
Trades	1,315
Mean	-10.8% (95% CI -29.2 to +22.3)
Median	-54.3%
Win rate	33.7%
Mean excluding best trade / best 3	-25.8% / -26.3%

Table 9. *Section 4.3 Using a take-profit and stop-loss instead*

A stop caps many losses near -35%, so the rule exit loses less than a fixed hold. Its CI reaches above zero only because of one trade, which jumped from below the +100% target to about +19,700% between two 5-minute snapshots (19 Oct, about 20:07 UK time) while the pool's liquidity drained from about $37k to $10k; 8 minutes later liquidity was $0.03. It was a price spike during a rug that a real order would almost certainly not have filled. Real stop fills in a fast rug would probably also be worse than snapshot fills. With the stricter coverage rule of section 8 (a path exit only counts if the snapshots cover the whole 6 h window), the rule exit gives n = 1,176, mean -9.4% (CI

-30.1 to +26.8), -26.2% without the best trade, and -24.2% on the held-out last 30% (n = 353, CI -33.4 to -14.9, entirely below zero).

Main result. Buying tradable new Solana tokens at about 30 minutes of age loses at every horizon (-31% at 1 h, -62% at 6 h, -56% at 24 h), with every CI far below zero and in both halves of the period.

4.4 Sensitivity to costs, trade size, the dead-token rule, honeypots and collector gaps

These checks were added after an internal critical review on Mon 26 Oct, on the same 09:37 database copy. They test whether the losses are an artefact of the cost model, of scoring dead tokens as -100%, or of horizons lost in collector gaps. The cost, size, dead-rule and honeypot rows use the independent snapshot-matching code of section 8 (so the lab-cost row differs slightly from 4.1; 2,000 resamples); the two gap rows use the lab's report code, as in 4.1 (4,000 resamples).

Scenario (buy every tradable token at the decision)	1 h mean	6 h mean (95% CI)	24 h mean
Lab model: 1.25% fee + 1% slippage per leg, $0.10 a transaction, $25	-32.6%	-61.7% (-68.1 to -52.7)	-55.8%
Bot user: 2.25% fee (venue + 1% bot fee) + 1% slippage	-34.0%	-62.5% (-68.8 to -53.8)	-56.7%
Low cost: 0.30% fee + 0.5% slippage, $0.01 a transaction	-30.2%	-60.3% (-67.0 to -50.9)	-54.2%
No fees, slippage or network cost (price impact only)	-29.0%	-59.7% (-66.4 to -50.3)	-53.4%
No cost of any kind (raw price change, dead rule kept)	-28.7%	-59.4% (-66.3 to -50.0)	-53.1%
Lab model at $100 a trade	-33.1%	-62.2% (-68.3 to -53.5)	-56.4%
Lab model at $1,000 a trade	-39.9%	-66.5% (-71.0 to -61.3)	-62.8%
Dead only if liquidity < $500 (no price-ratio rule)	-32.5%	-61.5% (-67.8 to -52.6)	-55.5%
Dead only if the pair is removed; otherwise valued at its price	-29.0%	-57.5% (-64.1 to -48.1)	-52.6%
Pair removed: sold at the last price seen before removal	-29.0%	-57.5% (-64.2 to -48.8)	-51.9%
Lab model, likely honeypots excluded (see 9.7)	-41.8%	-74.2% (-80.8 to -65.1)	-67.4%
Lab model, likely honeypots scored as -100%	-48.0%	-77.0% (-82.7 to -68.0)	-70.8%
Gap horizons: dead at last observation = -100%, others excluded (n 1,319 / 1,307 / 1,221)	-32.7%	-65.3% (-70.9 to -57.4)	-72.5%
Gap horizons: dead = -100%, others at last observed price (n 1,385 / 1,404 / 1,411)	-30.9%	-60.9% (-66.4 to -53.4)	-63.1%

Table 10. *Section 4.4 Sensitivity to costs, trade size, the dead-token rule, honeypots and collector gaps*

Removing every fee, slippage and network cost moves the 6 h mean by only about 2 points (-61.7% to -59.4%), and larger trades lose more because of price impact. The dead-token rule moves it by at most about 4 points. The one thing that changes the headline materially is honeypots: 144 of the 1,411 universe tokens show the signature of a honeypot (section 9.7), most of them one buy-only cluster whose price rose about 5x while no one could sell. Their paper gains (+44% mean at 6 h) were propping up the buy-everything result, so the true loss for a buyer is larger than the headline figures in 4.1, not smaller. Collector gaps work the same way. The universe tracker keeps polling for 24.4 h whether or not a pair exists, so missing horizons come from gaps, not from tracking stopping. Of the 224 tokens with no 6 h outcome, 120 were already dead at their last observation, 97 were still alive and 7 had no earlier observation. Counting the dead ones as -100% gives -65.3% at 6 h; marking the live ones at their last price as well gives -60.9%. The headline sits between the two, so gaps do not explain the loss.

Late entries are different. For entries 6 hours or more after the decision, costs decide the sign. Under the lab model these cells lose slightly or break even (section 8.2). With low or zero costs the 6 h-delay cells become slightly positive in the full sample (for example +2.2%, CI -6.7 to +14.2, n = 299 for a 6 h hold at raw prices, median 0.0%), but their held-out means are still negative (-5.2%, CI -11.9 to +1.5). The exception is the 6 h delay with a 12 h hold, at +24.5% held-out (n = 76, CI -16.3 to +102.7), which is too wide to mean anything. Entry at 3 h still loses significantly at zero cost (-10.9%, CI -18.0 to -2.3, n = 351). Re-running all 400 strategies of section 8 at low cost and at zero fees and slippage, none had a significantly positive mean (one-sided t-test, even uncorrected); 154 and 145

of the 311 with n ≥ 30 were still significantly negative after Bonferroni correction. The censoring check of section 8.2 does not change this. So late entries show no demonstrated edge at any cost level, but they also cannot be called proven losers at very low cost.

5. Secondary studies

Study	Rule	Sample	Main result (net)	Verdict
pumpfun-forward	FDV ≥ $8k at 10 min + first-time creator + not bundled; +100% / -35% / 6 h	175 qualifiers of 300	-44.0%, CI -56.4 to -31.6	**Insufficient; all pass tests negative**
alpha404-test	Liquidity ≥ $50k, age ≥ 2 h, rug filter, X HOT; scale-out exit	28 of 100 Arm B	-39.6%, CI -51.1 to -26.5	**Insufficient; negative**
rugalt v1.2	Lab rule with on-chain checks in place of RugCheck	77 buys (forward)	Rule exit -27.2%; 6 h -64.1%	**Negative**
X promotion	Group tokens by promotion seen on X	69 tokens	Every group -60% to -95% at 6 h	**No usable signal**

Table 11. *Section 5. Secondary studies*

5.1 pump.fun launch filter (H2)

The collector watched 83,525 pump.fun launches from 19 Oct 09:09 to 23 Oct 15:22. 2,108 passed the $8k FDV gate at 10 minutes and 175 qualified (first-time creator and not bundled) before the collector stopped. It died after a restart on 23 Oct ("venv python missing") and was not restarted. The figures below were recomputed for this report from a copy of the study database, using the pre-registered statistics (10,000 resamples, seed 12345).

Group	n	Mean	95% CI	Median	Win	Mean ex-top 3
Qualifiers (pre-registered)	175	-44.0%	-56.4 to -31.6	-100%	28%	-46.9%
Gated, not qualifying	1,322	-81.7%	-84.7 to -78.3	-100%	7%	-82.9%
Gated, creator check failed (RPC)	578	-82.3%	-86.7 to -77.6	-100%	7%	-84.4%
All gated	2,075	-78.7%	-81.3 to -76.0	-100%	9%	-79.6%

Table 12. *Section 5.1 pump.fun launch filter (H2)*

The creator filter separates bad from worse (qualifiers lose about half as much), but the qualifiers' whole CI is below zero and all three pre-registered tests (mean, CI lower bound, mean excluding the top 3) are negative. Repeat creators are much worse (-92.1% development, -86.4% held-out), and creators with 3 or more tokens in the dataset lose about 98%. FDV at 10 minutes matters too: $8-15k ≈ -97%, $15-30k ≈ -84% / -80%, $30k or more -50.6% / -57.0% (development / held-out). Holding gated tokens raw for 24 h with no costs gave -56.8% (n = 438, all gated tokens with a 24 h price). pump.fun prices after graduation showed a look-ahead artefact (+14.8% at 15 min) and were not used.

5.2 Established tokens plus X sentiment (H3, alpha404)

Entries ran from 19 Oct 12:32 to 23 Oct 15:18, when the collector stopped. The pre-registered report was re-run on a database copy on 26 Oct 09:38. Trades still open when the collector stopped are excluded; their marks are not results. Scout verdicts were HOT 42, NOISE 166 and TRAP 17.

Arm	Entries	Scale-out exit, completed	Lab exit, completed
C: base (liquidity ≥ $50k, age ≥ 2 h)	338	n = 167, -48.1% (CI -54.3 to -41.7)	n = 322, -11.6% (CI -15.4 to -7.7)
A: + volume/liquidity < 30, no RugCheck danger	234	n = 106, -44.5% (CI -52.4 to -36.4)	n = 221, -10.5% (CI -15.0 to -5.9)
B: A + X scout HOT	42	n = 28, -39.6% (CI -51.1 to -26.5)	n = 42, -10.8% (CI -22.4 to +2.7)

Table 13. *Section 5.2 Established tokens plus X sentiment (H3, alpha404)*

Each extra filter slightly reduces the loss, but no arm makes money (smaller losses than the launch studies, as these tokens are older and more liquid). Only 28 of the 100 required Arm B trades completed, so the formal verdict is

insufficient sample. On the development/held-out split, Arm C tokens whose pair was over 72 h old lost -10.9% / -26.6% over 24 h, while pairs 72 h old or younger lost -63.9% / -67.7%.

5.3 On-chain rug check (H4, rugalt v1.2)

RugCheck went down on 18-19 Oct, so a sidecar re-implemented the hard filters with direct on-chain reads. It could not reproduce insider-network detection, so it is looser than the original. It agreed with real RugCheck on 824 of 856 decisions, and most disagreements involved insider networks. Forward results from its 23 Oct 15:25 report:

Variant	Buys	Rule exit	Fixed 6 h
v1.2 (forward)	77	n = 72, mean -27.2%, 17 wins	n = 58, mean -64.1%
v1.2b stricter (forward)	41	n = 38, mean -16.8%, 13 wins	n = 30, mean -73.3%
v1.2 re-run, all data	—	n = 87, -22.3% (CI -39.1 to -4.3)	n = 73, -53.8%
v1.2b re-run, all data	—	n = 43, -9.6% (CI -35.0 to +17.9)	n = 35, -52.6%
Backfill (look-ahead, does not count)	15 / 5	+1.6% / +45.6%	-13.9% / +71.8%

Table 14. *Section 5.3 On-chain rug check (H4, rugalt v1.2)*

The only positive numbers come from the look-ahead backfill (5 v1.2b trades), which the forward design exists to exclude. Forward, the on-chain version is negative, like the original.

5.4 X promotion (H5, exploratory) and the three different numbers

69 universe tokens were enriched with X posts (38 of them were capped at 25 posts). Three different figures for "promoted" tokens have been quoted in earlier updates. They are not contradictory; they measure different groups:

Figure	Group	n with 6 h outcome	How it is measured
-94.5%	Posts with explicit paid-promotion language (#ad, sponsored)	11	Lab report, lab simulator (latency, snapshot tolerance)
-69.3%	#ad posts or template posts with "not financial advice" / DYOR	23	Exploratory script, entry at the first price point after the decision
-63.2%	No promotional posts at all	46	Same exploratory script

Table 15. *Section 5.4 X promotion (H5, exploratory) and the three different numbers*

Other groups in the lab report: no posts at all -64.0% (n = 5), any posts -66.3% (n = 52), any account with over 50k followers posting -59.6% (n = 15), two or more large accounts posting within 10 min -60.0% (n = 5). Every group loses heavily; explicit paid promotion may mark the worst tokens, but with 11 tokens that is a hint, not a finding.

5.5 Tokens skipped as too small, and "top young token" lists

One possibility is that the gate threw away the winners. To check, 422 tokens below the gate were sampled and followed (360 completed 24 h). Even raw, with no costs, the median 24 h multiple was 0.00x, 72% were dead and 1% were at 2x or more. Within 24 h, 3 of 422 reached at least 2x, and 2 reached 10x or more. Those two (WHITEWHALE and pwn) had $0 of liquidity at decision time and could not have been bought, so the gate is not hiding a profitable tail.

Another possibility is that the public "top young tokens" lists (DexScreener profiles and boosts, GeckoTerminal top volume) pick winners. 1,185 tokens appeared on them, and 286 had already been enrolled in the lab when first listed (134 of them were below the tradability gate). Buying at the first listing lost -14.6% / -76.0% at 60 min and -34.0% / -89.8% at 6 h (development n = 95 / 90, held-out n = 29 / 24).

6. Exploratory edge search

Rules that looked promising on the development period were checked on the held-out period. All values are net mean returns. The exit is a fixed 6 h hold unless stated otherwise.

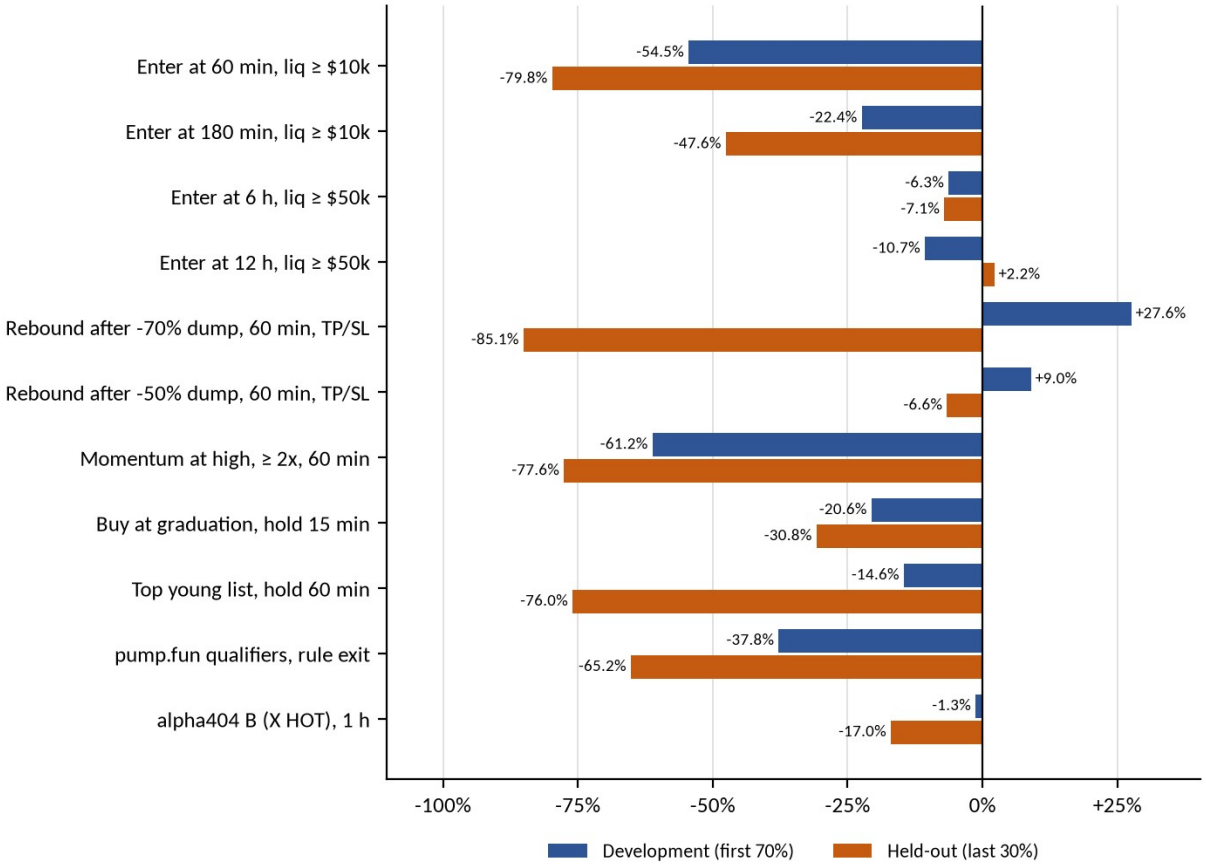


Figure 5. *Exploratory rules: mean net return on development data (first 70%) and on held-out data (last 30%). Values as shown on the study Page.*

Rule	Dev n	Dev mean	Held-out n	Held-out mean (CI)
Enter at 60 min, liq ≥ $10k	624	-54.5%	176	-79.8% (-85.4 to -73.5)
Enter at 180 min, liq ≥ $10k	176	-22.4%	31	-47.6% (-66.9 to -27.1)
Enter at 6 h, liq ≥ $50k	62	-6.3%	18	-7.1% (-30.2 to +14.4)
Enter at 12 h, liq ≥ $50k	44	-10.7%	13	+2.2% (-12.7 to +17.7)
Rebound after ≥70% drawdown, entry 60 min, +100/-35	20	+27.6%	4	-85.1% (-100 to -55.2)
Rebound after ≥50% drawdown, entry 60 min, +100/-35	61	+9.0%	12	-6.6% (-59.3 to +52.5)
Rebound after ≥70% drawdown, entry 6-12 h, +100/-35	45	+6.9%	10	-34.2% (-58.8 to -8.2)
Momentum (at high and ≥ 2x), 60 min	265	-61.2%	87	-77.6% (-86.0 to -68.2)
Buy at real graduation, hold 15 min	102	-20.6%	42	-30.8% (-54.3 to -1.7)
Buy at real graduation, hold 6 h	102	-41.2%	25	-89.2% (-96.8 to -79.1)
Top young list, first listing, 6 h	90	-34.0%	24	-89.8% (-100 to -76.8)
Survivors at 24 h, liq ≥ $10k, hold 24 h	—	—	12 (all)	-84.2%

Table 16. *Section 6. Exploratory edge search*

Results. The three rules positive on development data (post-dump rebounds) all lost held-out. The few slightly positive held-out cells (12 h entry with at least $50k of liquidity, +2.2%, n = 13, and two rebound variants with n = 5 to 7) were negative on development data, with tiny samples and CIs spanning zero; among about 240 cells that is what chance produces. A survival model (n = 943 entries at 60 min) found a 51.2% dead rate, and only 1.4% of tokens reached 2x net; in the held-out period the share reaching 2x was 0.0% in every tercile of every feature. Feature terciles (liquidity, FDV, holders, flow) gave no consistent ranking between the development and held-out periods.

The one consistent pattern is that losses shrink as entry moves later and liquidity goes up (-80% at 60 min, -48% at 3 h, about -7% at 6 h with $50k or more): once the rugs have passed, what remains roughly breaks even before costs. That is lower rug risk, not an edge.

7. The funnel: from 91,657 pools to 1,411 tradable tokens

Counted on the database copy taken Mon 26 Oct 09:37 UK time

Every new pool the collector saw was enrolled, and each one then passed through the same fixed gates. The counts below come from the 09:37 database copy, which has 91,657 candidates; the 09:02 weekly report quoted in section 3 had 91,438 because the collector kept enrolling new pools between the two. The final universe is the same 1,411 tokens analysed in section 4.

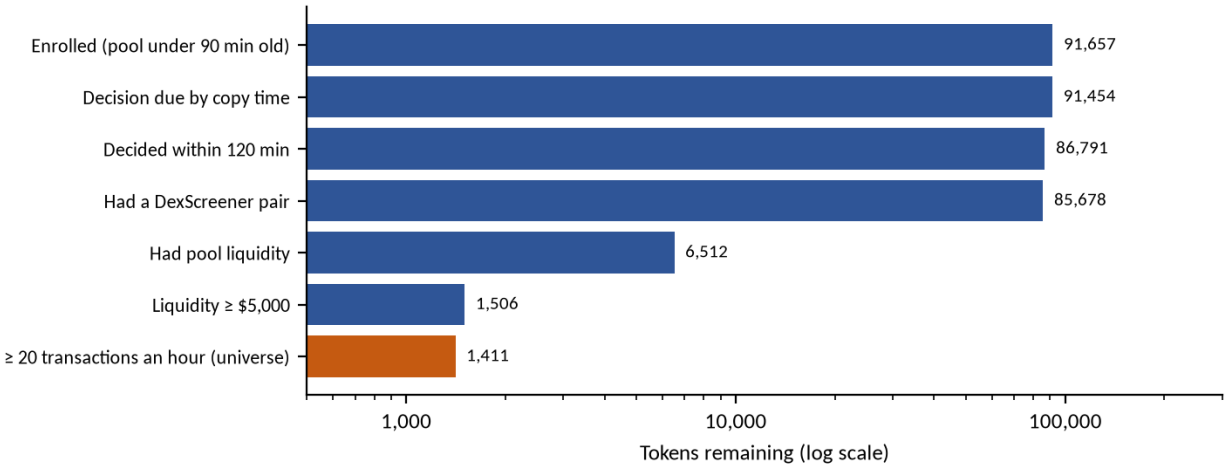


Figure 6. *The funnel from enrolled pools to the tradable universe (tokens remaining after each step; log scale). Values as shown on the study Page.*

Step	Criterion	Remaining	Removed	Why this step exists
0	Enrolled: new Solana pool seen on GeckoTerminal (36,491) or Jupiter (55,166), under 90 min old	91,657	—	Defines the population: every new launch the collector could see, with no selection.
1	Decision due (pool reached 30 min of age before the copy)	91,454	203	The 203 were still pending when the copy was taken. Not a filter.
2	Decided within 120 min of pool age	86,791	4,663	Missed: the collector was suspended when the decision fell due, so no decision-time data exists. Recorded, not imputed.
3	DexScreener returned a pair (otherwise NO_DATA)	85,678	1,113	No public price, liquidity or volume at all, so nothing to buy against.
4	Pool liquidity reported at decision time	6,512	79,166	Almost all were still on a launchpad bonding curve (about 73,800 on pump.fun, about 3,200 on Meteora's curve) with no AMM pool. They could be bought on the curve, but they are not a pool with liquidity, so they sit outside this universe. 43,428 of them also had fewer than 20 transactions an hour. They are scored separately in section 9 with a bonding-curve model, and in the pump.fun study (5.1).
5	Pool liquidity ≥ $5,000	1,506	5,006	Below $5,000 a $25 order moves the price by about 1% each way on top of fees, and such pools are rugged or abandoned most often. 4,432 of the 5,006 did have 20 or more transactions an hour.
6	≥ 20 buys plus sells in the last hour	1,411	95	Ensures there is an active market to sell back into, not a pool with liquidity but no trading.

Table 17. *Section 7. The funnel: from 91,657 pools to 1,411 tradable tokens*

RugCheck rate limits caused retries, not exclusions, and no further filter was applied inside the universe. Each gate is a minimum condition for a $25 trade to be possible and measurable and none uses returns, so the funnel cannot have selected the losers; sections 5.5 and 9 test whether it threw away the winners.

8. Every trading method tested on the 1,411

8.1 Design

Instead of only buying at 30 minutes and holding, 400 different strategies were run on the same 1,411 tokens, plus 16 graduation strategies on all pump.fun graduations. Every trade uses the lab's own cost simulator unchanged ($25, 1.25% fee and 1% slippage per leg, constant-product impact, $0.10 per transaction; liquidity under $500, a price at 10% of entry or less, or a removed pair count as -100%). Delays are measured from the lab's decision (about 30 minutes of pool age) plus 20 s latency.

Family	Strategies	Entry	Exit
Delay / hold	42	Wait 0, 15 min, 1 h, 3 h, 6 h, 12 h or 24 h; liquidity ≥ $5k at entry	Hold 15 min, 1 h, 3 h, 6 h, 12 h or 24 h
Take-profit / stop-loss	120	Delay 0, 1 h or 6 h	TP +25/+50/+100/+200/+400% × SL -20/-35/-50%/none × max hold 6 h or 24 h
Trailing stop	18	Delay 0, 1 h or 6 h	Trail 20/35/50% from the high × max 6 h or 24 h
Buy the dip	20	First fall of 30/50/70/90% from the running high within 6 h	Hold 1 h, 6 h, 24 h; TP100/SL35/6 h; trail 35%/24 h
Momentum	20	First rise of 25/50/100/200% within 6 h	Same five exits
Survivor	144	Still alive at 1, 3, 6 or 12 h with liquidity ≥ $10k/$25k/$50k/$100k and 1 h volume ≥ $0/$5k/$25k	Hold 1 h, 6 h or 12 h
Order-flow momentum	36	At 0, 1 h or 6 h, buys/sells ≥ 1.0/1.5/2.0 and volume/liquidity ≥ 0.5/2.0	Hold 1 h or 6 h
Graduation (all pump.fun)	16	0, 15 min, 1 h or 3 h after a token graduated to a pool with ≥ $5k liquidity	Hold 1 h, 6 h, 24 h; TP100/SL35/6 h

Table 18. *Section 8.1 Design*

- **Coverage rule.** A path exit (stop, target, trail) is only scored if the snapshots cover the whole holding window with no gap over 1 h (15 min tolerance at each end inside the 24.4 h snapshot window, 90 min beyond it). Without this, a target hit just before a collector gap was kept while a stop hidden in the gap was not, which inflated results.
- For every strategy: n, mean, median, win rate, dead rate, 95% bootstrap CI (4,000 resamples, seed 7), mean without the best 1 and best 3 trades, and first and second halves by time.
- **Multiple testing.** One-sided t-tests with a Bonferroni correction for all 400 strategies ($\alpha = 0.05/400$), and White's reality check (centred bootstrap of the maximum standardised mean over all 311 strategies with n ≥ 30).
- **Out-of-sample test.** Each strategy's trades were split by time into the first 70% (development) and last 30% (held-out). The best strategy is chosen on development data and judged only on held-out data.
- Small differences from section 4 (for example 1,296 vs 1,293 one-hour outcomes) come from this independent snapshot-matching code, not from different data.

8.2 When you buy and how long you hold

Mean net return by entry delay after the 30-minute decision, for three holding periods (n per cell in the table below):

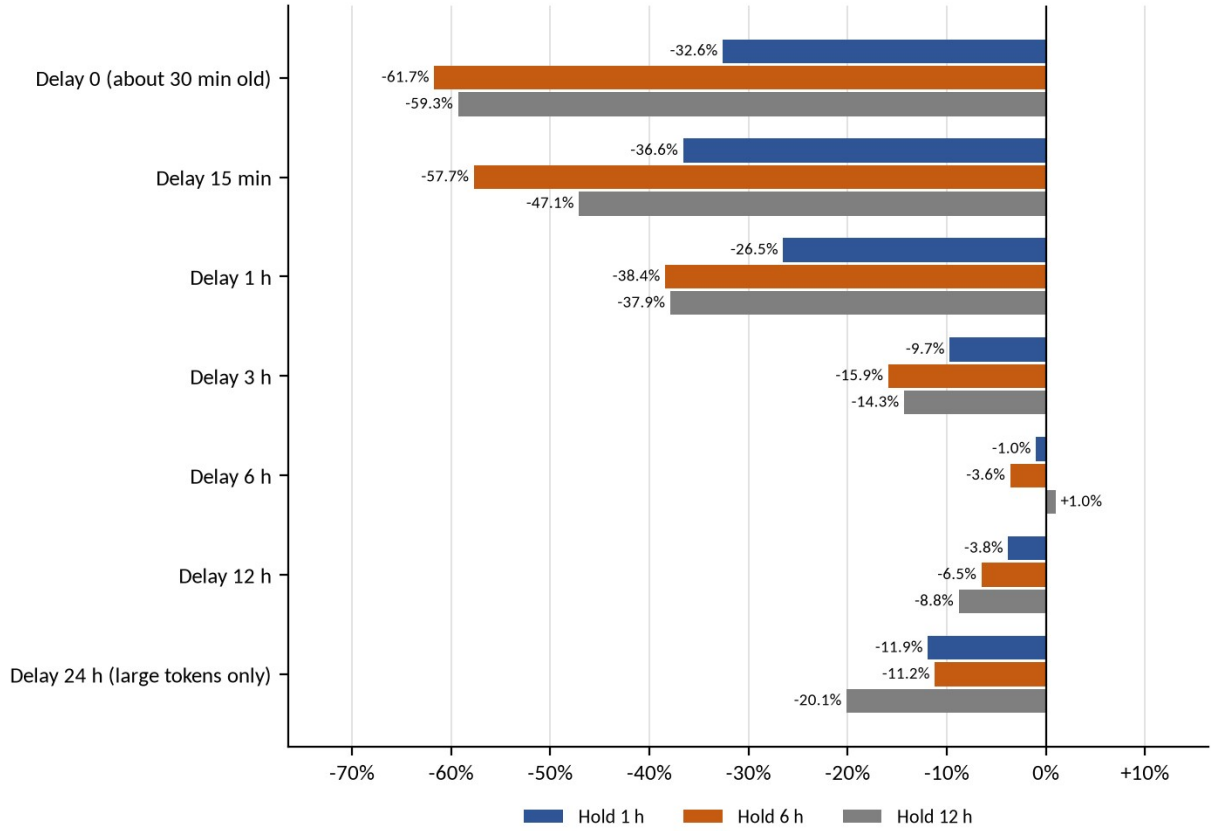


Figure 7. *Mean net return by entry delay and holding period (lab cost model). Values as shown on the study Page.*

Entry delay	Hold	n	Mean	95% CI	Median	Win	Dead	Ex-top 1 / top 3	Halves	One-sided p (mean < 0)
0	1 h	1,296	-32.6%	-37.9 to -27.0	-83.9%	28%	49%	-33.8 / -35.0%	-37.1 / -28.1%	2e-30
0	6 h	1,187	-61.7%	-68.2 to -52.7	-100%	18%	67%	-65.2 / -66.6%	-53.8 / -69.7%	2e-49
15 min	1 h	1,031	-36.6%	-41.5 to -31.6	-56.4%	28%	45%	-37.3 / -38.2%	-35.9 / -37.3%	1e-43
15 min	6 h	922	-57.7%	-66.8 to -42.5	-100%	21%	60%	-64.0 / -65.0%	-46.4 / -68.9%	5e-18
1 h	1 h	598	-26.5%	-31.0 to -21.9	-5.2%	17%	25%	-27.5 / -28.4%	-20.4 / -32.6%	2e-26
1 h	6 h	525	-38.4%	-45.5 to -30.0	-45.0%	14%	35%	-40.4 / -43.8%	-32.4 / -44.4%	5e-20
3 h	1 h	389	-9.7%	-12.9 to -6.2	-5.2%	14%	4%	-10.6 / -11.2%	-9.7 / -9.8%	5e-9
3 h	6 h	351	-15.9%	-22.4 to -8.1	-5.2%	20%	10%	-18.6 / -20.4%	-18.3 / -13.4%	1e-5
6 h	1 h	322	-1.0%	-8.4 to +11.2	-5.2%	14%	1%	-6.1 / -7.9%	-4.5 / +2.6%	0.43
6 h	6 h	299	-3.6%	-12.0 to +7.6	-5.2%	25%	3%	-7.9 / -10.3%	-5.0 / -2.2%	0.24
6 h	12 h	252	+1.0%	-14.1 to +25.1	-5.2%	22%	5%	-8.4 / -13.1%	-0.8 / +2.9%	—
12 h	1 h	268	-3.8%	-6.0 to -1.0	-5.2%	13%	0%	-4.8 / -5.5%	-2.3 / -5.4%	0.002
12 h	6 h	229	-6.5%	-10.6 to -2.5	-5.2%	24%	3%	-7.3 / -8.7%	-4.9 / -8.1%	0.001
24 h	1 h	61	-11.9%	-16.3 to -8.4	-5.3%	2%	2%	-12.3 / -12.7%	-14.4 / -9.5%	9e-8
24 h	6 h	42	-11.2%	-21.4 to -0.4	-5.1%	26%	5%	-14.4 / -16.5%	-9.9 / -12.4%	0.02

Table 19. *Section 8.2 When you buy and how long you hold*

Entering in the first hour is proven to lose at every hold (every CI far below zero; p between 10^{-49} and 10^{-17} on 525 to 1,296 trades), and entries at 3 h, 12 h and 24 h also lose significantly. Only the 6 h entries are not significantly negative (-3.6% to +1.0%, n = 252 to 322). Their median of -5.2% is exactly the round-trip cost: most tokens still listed at 6 h no longer trade, so the buyer just pays the costs. Every CI includes losses of 8% or more and removing one trade turns each negative, so these cells show losses too small to prove with about 300 trades, not a profit. Entering at 24 h is only possible for the larger tokens that got the 7-day follow-up, which is why n is small there. Censoring check (2.5): counting tokens dead before a gap as -100% makes the early cells worse by up to 9 points

(delay 0 / hold 6 h: -65.3%, n 1,307) and puts the 6 h-delay cells at -1.0%, -3.9% and -0.9% (n 322, 300, 257; -1.4%, -4.5% and -2.5% with live tokens at last price). No cell becomes significantly positive.

8.3 The best of each family

Best strategy in family (n ≥ 30)	n	Mean (95% CI)	Median	Win / dead	Ex-top 1 / top 3	Halves	Held-out n, mean (CI)
Dip: -50% within 6 h, then TP100/SL35/6 h	278	+56.8% (-21.7 to +211.0)	-44.4%	27% / 19%	-14.1 / -16.0%	-13.8 / +127.5%	83, -8.3% (-26.7 to +10.7)
Survivor at 12 h, liq ≥ $10k, vol ≥ $5k, hold 12 h	65	+7.1% (-16.6 to +34.4)	-3.4%	46% / 12%	-1.2 / -12.1%	-0.6 / +14.6%	19, +1.0% (-31.8 to +33.2)
Trailing: delay 6 h, trail 20%, max 6 h	299	+4.5% (-9.8 to +30.4)	-5.2%	20% / 2%	-6.7 / -9.2%	-7.4 / +16.4%	90, -8.5% (-13.2 to -4.3)
Delay/hold: delay 6 h, hold 12 h	252	+1.0% (-14.1 to +25.1)	-5.2%	22% / 5%	-8.4 / -13.1%	-0.8 / +2.9%	76, +16.5% (-20.9 to +83.2)
Order flow: at 6 h, buys/sells ≥ 1.0, vol/liq ≥ 0.5, hold 6 h	44	-0.0% (-23.6 to +29.2)	-10.7%	43% / 5%	-8.5 / -20.0%	+1.1 / -1.1%	13, +23.9% (-28.5 to +92.5)
TP/SL: delay 0, TP100/SL35, max 24 h	711	-1.5% (-34.3 to +58.0)	-57.0%	32% / 44%	-29.3 / -30.2%	-27.8 / +24.7%	213, +58.7% (-41.7 to +248.1)
Momentum: +25% within 6 h, TP100/SL35/6 h	691	-31.2% (-36.9 to -25.4)	-52.2%	36% / 42%	-31.7 / -32.3%	-29.3 / -33.0%	207, -34.3% (-44.6 to -22.4)

Table 20. *Section 8.3 The best of each family*

Of the 311 strategies with at least 30 trades, 7 had a positive mean and none had a 95% CI entirely above zero. Every positive mean disappears when the single best trade is removed (the largest ex-top-1 figure is -1.2%). The positive means are driven by outliers: the dip strategy's mean rests on one fill at about 197x, which is the same liquidity-drain artefact described in section 4.3, and the TP/SL 24 h held-out mean of +58.7% has a median of -57%. Momentum and order-flow entries lose most consistently: 17 of 18 momentum strategies and 18 of 22 order-flow strategies with n ≥ 30 are significantly negative after Bonferroni correction.

8.4 Multiple testing and the held-out test

182 of 311	**0 of 400**	**p = 0.35**	**-81.2%**
significantly negative	**significantly positive**	**White's reality check**	**dev-best on held-out data**
one-sided t, Bonferroni α = 0.05/400	even with no correction (p < 0.05)	best strategy is not better than chance	n = 27, CI -94.7 to -58.0

After a Bonferroni correction for all 400 strategies, 182 of the 311 strategies with n ≥ 30 lose money with statistical significance (249 without the correction; 190 of 312 when dead-before-gap tokens are counted as -100%, and still none significantly positive). Not one strategy of 400 shows a significantly positive mean, even before any correction. White's reality check, which accounts for the search itself, gives p = 0.35 for the best one ("dip -50%, TP100/SL35/6 h"), so its +56.8% mean is what searching 311 strategies on noise would produce.

Development vs held-out mean for the five best strategies chosen on development data, and for the only strategy positive on both:

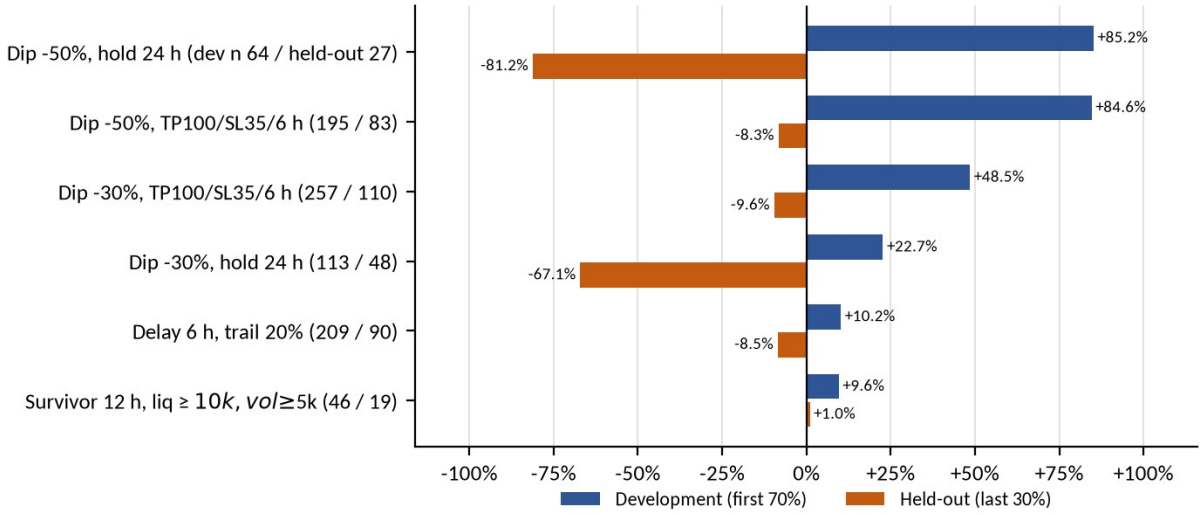


Figure 8. *The five best strategies chosen on development data, and the only strategy positive on both periods: development vs held-out mean net return. Values as shown on the study Page.*

Chosen on development data	Dev n	Dev mean	Held-out n	Held-out mean (95% CI)	Verdict
1. Dip -50% within 6 h, hold 24 h	64	+85.2%	27	-81.2% (-94.7 to -58.0)	Proven to fail: CI entirely below 0
2. Dip -50% within 6 h, TP100/SL35/6 h	195	+84.6%	83	-8.3% (-26.7 to +10.7)	Fails: held-out negative
3. Dip -30% within 6 h, TP100/SL35/6 h	257	+48.5%	110	-9.6% (-25.3 to +6.2)	Fails: held-out negative
4. Dip -30% within 6 h, hold 24 h	113	+22.7%	48	-67.1% (-94.3 to -20.5)	Proven to fail: CI entirely below 0
5. Delay 6 h, trail 20%, max 6 h	209	+10.2%	90	-8.5% (-13.2 to -4.3)	Proven to fail: CI entirely below 0
Only strategy positive on both (of 290 with dev n ≥ 30 and held-out n ≥ 15)	46	+9.6%	19	+1.0% (-31.8 to +33.2)	Too thin to mean anything

Table 21. *Section 8.4 Multiple testing and the held-out test*

The best strategy in hindsight is proven to fail out of sample (+85.2% development, -81.2% held-out, whole CI below zero), and three of the top five development picks have held-out CIs entirely below zero and the other two have negative held-out means. Only 14 of the 290 strategies with enough trades in both periods were positive on development data, and one stayed positive held-out (+1.0% on 19 trades, CI -32% to +33%), well within chance. Held-out samples are smallest for long delays because collector coverage after 20 Oct was thin.

8.5 Buying at graduation (all pump.fun graduations)

pump.fun tokens that completed their bonding curve and graduated to a pool with at least $5,000 of liquidity were bought 0, 15, 60 or 180 minutes after the graduation was seen (205, 133, 96 and 64 graduations), with the same costs.

Entry after graduation	Exit	n	Mean (95% CI)	Median	Win	Ex-top 1	Dev → held-out mean (n)
+0	1 h	195	-35.3% (-54.8 to -10.8)	-82.2%	15%	-42.8%	-37.7% → -29.6% (59)
+0	6 h	174	-55.4% (-77.8 to -23.1)	-100%	8%	-67.0%	-43.2% → -83.9% (52)
+0	24 h	92	-50.2% (-84.0 to -0.9)	-100%	8%	-68.7%	-41.6% → -69.9% (28)
+0	TP100/SL35/6 h	108	-17.9% (-32.7 to -1.6)	-51.9%	24%	-20.2%	-15.4% → -23.7% (32)
+15 min	1 h	129	-5.3% (-30.1 to +24.0)	-58.8%	24%	-12.8%	-1.2% → -14.7% (39)
+15 min	6 h	110	+16.0% (-65.3 to +154.6)	-84.2%	14%	-43.3%	+51.4% → -66.6% (33)

Entry after graduation	Exit	n	Mean (95% CI)	Median	Win	Ex-top 1	Dev → held-out mean (n)
+15 min	TP100/SL35/6 h	78	-17.3% (-32.1 to -1.7)	-49.4%	23%	-19.5%	-20.2% → -10.5% (23)
+1 h	1 h	92	-28.8% (-39.8 to -17.1)	-44.3%	17%	-32.1%	-22.4% → -43.4% (28)
+1 h	TP100/SL35/6 h	58	-27.3% (-41.2 to -10.9)	-43.7%	16%	-30.3%	-26.4% → -29.4% (17)
+3 h	6 h	53	-22.1% (-36.4 to -5.7)	-38.5%	30%	-26.4%	-21.7% → -23.0% (16)
+1 h / +3 h	24 h	12 / 8	+196% / +167% (CIs -58 to +525, -68 to +510)	-72% / -58%	42% / 25%	+75% / +5%	held-out -70.2% (4) / -71.7% (2)

Table 22. *Section 8.5 Buying at graduation (all pump.fun graduations)*

For 8 of the 16 graduation strategies the whole CI is below zero (immediate entry with all four exits; +15 min with the rule exit; +1 h with the 1 h hold and the rule exit; +3 h with the 6 h hold). The +15 min / 6 h cell repeats the overfitting pattern of 8.4 (+51.4% development, -66.6% held-out), and the two positive 24 h cells rest on 12 and 8 trades with medians of -72% and -58%, whose 4 and 2 held-out trades lost about 70%.

9. The top 10% of winners across all discovered tokens

9.1 Which tokens can be scored

This study uses every discovered token, not just the 1,411. A token can only be scored if it has a price that could actually be traded against. An observation counts if DexScreener listed the pair with a positive price and either pool liquidity of at least $500 or, for a pump.fun token on its bonding curve, a measurable FDV. A curve token can always be sold back to the curve, and its effective liquidity was modelled from the curve ($2 \cdot \sqrt{32.19\ \text{SOL} \times \text{FDV}}$ in dollars, using the recorded SOL price of about $103-109; checked against the pump.fun study's own liquidity readings). Prices come from the lab's snapshots, follow-up polls and the 5-minute age path: 4.55 million observations in all.

Group	Tokens	Share of 91,657
All discovered (09:37 copy)	91,657	100%
No market observation at all	108	0.1%
Fewer than two priced observations (no pair, or no price)	14,543	15.9%
Priced, but liquidity not measurable (pool under $500, or a non-pump.fun curve with no liquidity reading)	9,272	10.1%
Scorable: two or more observations with a tradable price	67,734	73.9%
of which on a bonding curve at first observation	64,810	70.7%

Table 23. *Section 9.1 Which tokens can be scored*

The 23,923 tokens that cannot be scored have no price history, or no liquidity that could be measured, so no return, winning or losing, can be computed for them, and any winners among them could not be shown to be tradable. The first observation was at a median of about 5 minutes after the pool appeared, so this study starts much earlier than the 30-minute universe.

9.2 How many winners there were

Highest price reached relative to the first observed price, all 67,734 scorable tokens:

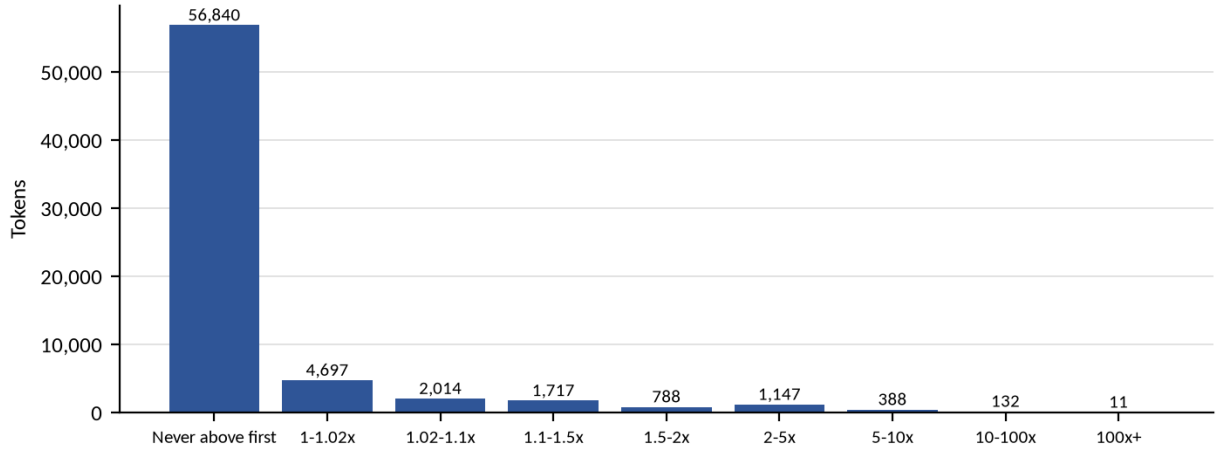


Figure 9. *Peak multiple over first observed price, scorable discovered tokens (number of tokens per band). Values as shown on the study Page.*

84% of scorable tokens never traded above their first observed price. The 90th percentile of the peak multiple is only 1.013x, so the "top 10%" (6,774 tokens) includes everything that rose by about 1.3% or more. Because that bar is so low, stricter winner groups are also reported: 1,678 tokens (2.5%) reached 2x, 531 reached 5x, 143 reached 10x, and 11 reached 100x or more.

9.3 Could the winners actually be sold at their peak?

Measure	Top 10% (≥ 1.013x)	≥ 2x	≥ 10x
Tokens	6,774	1,678	143
Liquidity at peak ≥ $5k / ≥ $20k	6,298 / 2,173	1,575 / 1,068	136 / 113
Time from first observation to peak (median)	0.5 h	0.75 h	1.35 h
Time spent within 80% of the peak (median)	50 min	5.5 min	5.5 min
Peak seen on a single poll only	15.2%	30.6%	32.9%
Gap between polls around the peak (median)	10 min	10 min	10 min
Still alive at last observation	42.7%	33.6%	37.8%
Perfect-foresight return, sold at the peak, $25 (mean / median)	+144.1% / +12.1%	+546% / +224%	+3,617% / +1,339%
Perfect-foresight return, sold at the peak, $100 (mean / median)	+112.3% / +8.8%	+429% / +212%	+2,377% / +1,226%

Table 24. *Section 9.3 Could the winners actually be sold at their peak?*

Most winners were tradable at their peak in principle: 93% of the top 10% and 94% of the 2x group had at least $5,000 of effective liquidity there, enough for a $25-$100 sale with modest slippage. But tokens that reached 2x or 10x stayed within 80% of their high for a median of 5.5 minutes, and about a third of those peaks were seen on a single poll (polls about 10 minutes apart). The perfect-foresight returns assume selling within minutes of an unknown top: an upper bound, not an achievable result.

9.4 What a buyer at the first observed price actually got

Group	Exit	n	Mean (95% CI)	Median	One-sided p (mean < 0)
All scorable tokens	Hold 1 h	59,845	-12.6% (-13.3 to -11.7)	-6.7%	10^{-215}
All scorable tokens	Hold 6 h	52,348	-15.6% (-16.1 to -15.1)	-6.8%	$< 10^{-300}$
All scorable tokens	TP100/SL35/6 h	23,618	-10.8% (-11.5 to -10.0)	-6.7%	10^{-184}

Group	Exit	n	Mean (95% CI)	Median	One-sided p (mean < 0)
All scorable tokens	Sell at the peak, $100 (perfect foresight)	67,734	-4.8% (-6.2 to -3.2)	-10.2%	3e-10
Top 10%, known in hindsight	Hold 1 h	6,414	+6.5% (-0.2 to +14.5)	-8.0%	—
Top 10%, known in hindsight	Hold 6 h	5,699	-19.4% (-23.0 to -15.4)	-12.6%	10^{-22}
Top 10%, known in hindsight	TP100/SL35/6 h	3,447	+15.6% (+11.4 to +20.5)	-6.7%	—
≥ 10x, known in hindsight	Hold 1 h / 6 h	139 / 119	+765% / +291%	+257% / -100%	—

Table 25. *Section 9.4 What a buyer at the first observed price actually got*

Buying every scorable token at its first observed price is proven to lose at every exit (23,618 to 59,845 trades, p below 10^{-180}). Even selling every token at its exact peak, a $100 buyer of everything loses 4.8% (CI -6.2% to -3.2%), because 84% of tokens never rise and each round trip costs about 5-10%. Even the hindsight top 10% lost 19.4% over 6 hours, with the whole CI below zero and a losing median at every exit.

Censoring check (section 2.5). Fixed holds here are scored only where a price was observed near the horizon. The 5-minute age path stops after six consecutive polls with no pair, and collector gaps also leave horizons unobserved, so the figures were recomputed with tokens already dead at their last observation scored -100% and live tokens either excluded or marked at their last price. At 6 h, 882 tokens were dead before the horizon, 13,837 were alive at a last price a median of about 2.6 h earlier, and 667 had no earlier observation. None of the 882 had lost its pair before 6 h (they were dead by price or liquidity) and 739 were observed again after the horizon, so their missing horizons come from collector gaps, not from the polling stop.

Figure (lab costs)	Published (n)	Dead before horizon = -100%, others excluded (n)	Dead = -100%, others at last price (n)
All scorable, hold 1 h	-12.6% (59,845)	-13.0% (60,117)	-12.5% (64,740)
All scorable, hold 6 h	-15.6% (52,348)	-17.0% (-17.5 to -16.5; 53,230)	-14.9% (-15.4 to -14.4; 67,067)
All scorable, 6 h, no cost of any kind	-9.1%	-10.6% (-11.1 to -10.1)	-8.2% (-8.8 to -7.6)
Top 10%, hold 1 h	+6.5% (6,414)	+5.9% (6,450)	+6.5% (6,725)
Top 10%, hold 6 h	-19.4% (5,699)	-22.7% (-26.5 to -18.9; 5,942)	-14.9% (-19.4 to -9.4; 6,764)
≥ 2x, hold 1 h	+115.6% (1,623)	+114.6% (1,630)	+115.3% (1,671)
≥ 2x, hold 6 h	+26.2% (1,434)	+18.4% (+5.1 to +32.8; 1,529)	+41.6% (+23.3 to +60.6; 1,677)
≥ 2x, hold 6 h, likely honeypots excluded	-5.6%	-12.0% (-25.1 to +2.3)	+12.8% (-4.3 to +32.4)
≥ 2x, hold 6 h, likely honeypots -100%	-13.2%	-18.6% (-30.3 to -5.6)	+4.1% (-14.1 to +24.2)
≥ 10x, hold 1 h / 6 h	+765% / +291% (139 / 119)	+759% / +258% (140 / 130)	+750% / +421% (143 / 143)

Table 26. *Section 9.4 What a buyer at the first observed price actually got*

Every buy-everything conclusion holds under both versions, with every CI below zero. The hindsight groups move more, because their missing horizons are a larger share, but they remain hindsight groups that cannot be traded.

9.5 Do winners look different before purchase?

To test whether winners can be recognised in advance, each token was bought on paper at its 30-minute snapshot using only data available then: DexScreener market data (liquidity, FDV, volume, buys, sells, price change), RugCheck holder and authority data (for 48,567 tokens), launchpad and venue, socials, time of day, the price path so far, and pump.fun creator history where available. 63,192 tokens had a tradable price at 30 minutes and an

outcome. A winner here is a token in the top 10% of peak multiple from that entry (≥ 1.004x; 6,320 tokens), and separately a token that reached 2x (586 tokens). Each row gives the median for winners and for the rest, and the AUC of that single feature (0.5 = no separation, 1.0 = perfect).

Feature at purchase	n	Winners median	Rest median	AUC (95% CI)
Total holders (RugCheck)	48,567	23	3	0.776 (0.769 to 0.784)
Buys in the last hour	63,192	71	8	0.749 (0.741 to 0.756)
Volume in the last hour	63,192	$4,986	$310	0.747 (0.741 to 0.754)
Volume / liquidity	63,192	0.62	0.06	0.716 (0.709 to 0.723)
Sells in the last hour	63,192	53	7	0.704 (0.697 to 0.713)
pump.fun first-time creator	1,179	58% yes	17% yes	0.704 (0.675 to 0.730)
Buys / sells	63,192	1.23	1.00	0.669 (0.660 to 0.676)
FDV	63,192	$2,961	$2,923	0.641 (0.634 to 0.650)
Liquidity (pool or curve)	63,192	$6,342	$6,281	0.640 (0.632 to 0.649)
Still on bonding curve	63,192	75.6% yes	98.8% yes	0.384 (0.379 to 0.388)
Creator's tokens in last 24 h	48,567	mean 3.6	mean 9.7	0.420 (0.414 to 0.424)
LP locked %	1,937	0%	100%	0.421 (0.396 to 0.446)
Has social link / website	63,192	6.0% / 4.4%	0.5% / 0.4%	0.528 / 0.520
Hour of day (UK)	63,192	14	13	0.530 (0.521 to 0.537)

Table 27. *Section 9.5 Do winners look different before purchase?*

The differences are mostly activity: tokens with more holders, buys and volume at 30 minutes were far more likely to move, while most of the rest are dormant curve tokens whose price never changes. Safety and identity features (social links, LP locks, mint and freeze authority, RugCheck danger flags) separate winners only weakly.

9.6 Predicting winners, and trading the predictions

Logistic regression and gradient boosting were trained on the first 44,234 tokens by time (17 Oct 11:07 to 20 Oct 06:26 UK time) and tested on the later 18,958 (20 Oct 06:26 to 26 Oct 08:18 UK time), with no tuning on the test period. The model's top-ranked test tokens (top 1%, 5% and 10%) were then bought on paper with the same costs, and compared with 4,000 random samples of the same size from the same test tokens.

Target	Model	Test AUC (95% CI)	Precision at top 1% / 5% / 10%	Base rate
Top 10% (≥ 1.004x)	Logistic regression	0.809 (0.797 to 0.820)	90.5% / 60.2% / 41.8%	8.5%
Top 10% (≥ 1.004x)	Gradient boosting	0.870 (0.861 to 0.879)	91.5% / 58.9% / 43.7%	8.5%
≥ 2x	Logistic regression	0.935 (0.906 to 0.960)	31.7% / 12.8% / 7.1%	0.8%
≥ 2x	Gradient boosting	0.953 (0.932 to 0.969)	29.1% / 13.5% / 7.2%	0.8%

Table 28. *Section 9.6 Predicting winners, and trading the predictions*

Net 6 h return of the gradient-boosting model's picks vs random picks from the same test tokens:

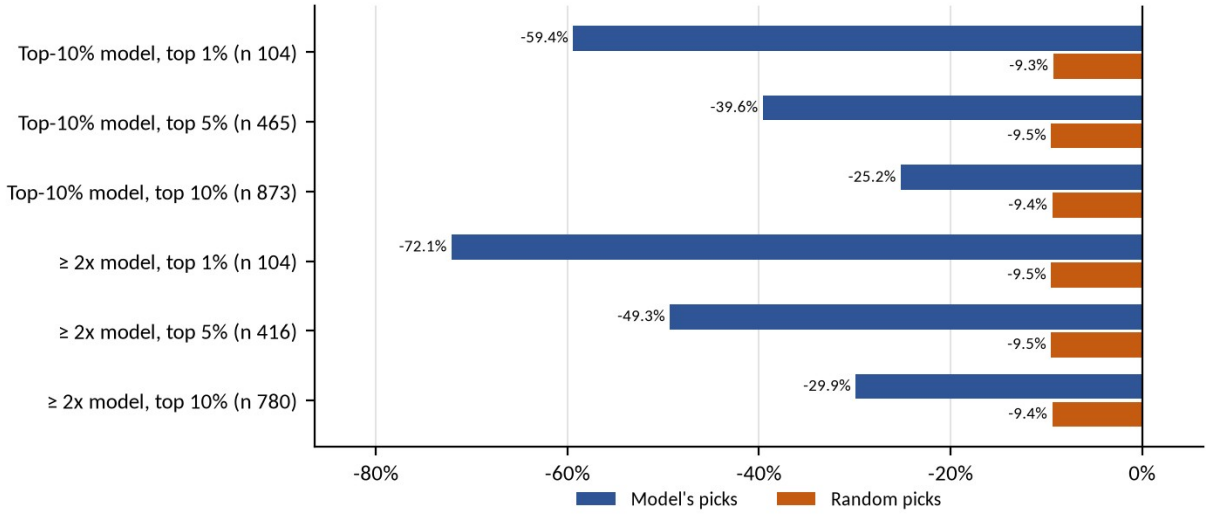


Figure 10. *Mean net return at 6 h of the gradient-boosting models' top picks compared with random picks of the same size. Values as shown on the study Page.*

Model picks (gradient boosting)	Exit	n	Mean (95% CI)	Median	Win / dead	Random picks (95% range)
Top-10% model, top 1%	1 h	152	-12.5% (-26.2 to +2.9)	-4.4%	43% / 44%	-7.2% (-10.6 to +7.8)
Top-10% model, top 1%	6 h	104	-59.4% (-71.1 to -47.5)	-100%	29% / 70%	-9.3% (-13.7 to -2.1)
Top-10% model, top 1%	TP100/SL35/6 h	88	-26.3% (-42.4 to -8.9)	-63.2%	44% / 45%	-9.8% (-16.0 to -3.0)
Top-10% model, top 5%	1 h	744	-15.6% (-23.0 to -6.5)	-11.0%	20% / 20%	-7.6% (-9.5 to +1.8)
Top-10% model, top 5%	6 h	465	-39.6% (-45.7 to -32.5)	-31.2%	11% / 35%	-9.5% (-11.6 to -2.8)
Top-10% model, top 5%	TP100/SL35/6 h	219	-24.0% (-34.0 to -13.8)	-45.0%	32% / 37%	-9.8% (-13.6 to -5.6)
≥ 2x model, top 1%	1 h	170	-7.3% (-32.9 to +31.2)	-54.6%	27% / 38%	-7.5% (-10.4 to -2.9)
≥ 2x model, top 1%	6 h	104	-72.1% (-82.2 to -60.5)	-100%	9% / 65%	-9.5% (-13.7 to -3.1)
≥ 2x model, top 1%	TP100/SL35/6 h	79	-25.3% (-44.7 to -6.8)	-64.9%	34% / 46%	-9.8% (-16.5 to -2.7)
≥ 2x model, top 5%	6 h	416	-49.3% (-53.8 to -44.3)	-46.9%	7% / 39%	-9.5% (-11.6 to -2.2)
Every test token (buy all)	1 h / 6 h / rule	14,113 / 7,000 / 1,559	-7.5% / -9.4% / -9.7%	-6.7%	—	—

Table 29. *Section 9.6 Predicting winners, and trading the predictions*

Two things are both statistically firm. The models pick out which tokens will move (AUCs of 0.87 and 0.95, CIs far above 0.5; the ≥ 2x models' top 1% holds 36-40 times the base rate of 2x tokens). And buying those picks loses more than buying at random: across both targets, both models, three pick sizes and three exits (36 combinations), every mean is negative, 32 have the whole CI below zero, and 33 did worse than at least 99.9% of random samples of the same size. The most active tokens at 30 minutes are the most likely both to spike and to be dumped (the 2x model's top 1% had a 6 h dead rate of 65%, against 2.5% for all 7,000 test tokens with a 6 h outcome), and the few that double do so for minutes. Under the censoring check the 6 h losses of the gradient-boosting picks grow: top-10% model -72.2%, -50.3% and -33.2% (top 1%, 5%, 10%; n 152, 565, 977), and ≥ 2x model -79.7%, -59.3% and -38.4% (n 143, 518, 887), with random picks at -10.9% to -11.2%; with live tokens marked at their last price they are -21.2% to -63.1% against -8.2% to -8.6% for random picks. In both versions every one of these pick sets did worse than all 1,000 random draws. Section 9.9 asks the same question for the 10x tokens alone, against matched look-alikes.

9.7 Honeypots and unsellable winners

This check was added after an internal critical review on Mon 26 Oct, on the same 09:37 database copy. A paper gain only counts if the token could actually be sold. A honeypot is a token that can be bought but not sold by ordinary holders, so its price can only rise. Every winner was checked for sellability with the data already in the database: RugCheck mint and freeze authority and Token-2022 extensions (permanent delegate, pausable, default-frozen accounts, transfer hook or transfer fee) and its risk flags; DexScreener sell and buy counts over the token's whole life, in the hour after its peak and at the peak; and whether pool liquidity was pulled or the pair removed after the peak. A token is classed as a likely honeypot if it had an authority or extension that lets someone block or seize transfers, or if it showed at least 5 buys an hour but zero sells in every observation. A possible honeypot also includes tokens with buys but no sells in the hour after the peak, or sells under 5% of buys at the peak.

Sellability check	Top 10% (≥ 1.013x)	≥ 2x	≥ 10x
Tokens	6,774	1,678	143
With RugCheck data	5,465	1,356	110
Freeze authority enabled	11	3	0
Mint authority enabled (dilution risk, not a honeypot by itself)	12	2	0
Risky Token-2022 extension	13	2	0
Buys but zero sells in every observation	138	124	0
Buys but no sells in the hour after the peak	134	84	0
Sells under 5% of buys at the peak	170	137	0
Likely honeypot	155 (2.3%)	129 (7.7%)	0
Possible honeypot (any signal)	248 (3.7%)	146 (8.7%)	0
At least one sell seen within 2 h after the peak	6,052	1,528	142
Pools (not curve) whose liquidity was pulled or pair removed within 6 h of the peak, honeypots excluded	1,439 of 2,014	596 of 803	58 of 67

Table 30. *Section 9.7 Honeypots and unsellable winners*

Most of the honeypots are one operation: 126 tokens on FluxBeam (52) or an unlabelled venue (74), launched between 17 Oct 10:07 and 20 Oct 18:14 UK time, with about 270-280 buys an hour and zero sells in every one of their 38 to 327 observations. All peaked at 52-67 minutes of age (median multiple 5.5x), with reported liquidity rising from a median of about $224,000 to an almost identical $1.16-1.27 million, and none still had a listed pool at its last observation. All 126 passed the universe gate, so they are among the 1,411 tokens of section 4, and they also make up most of the honeypots in the ≥ 2x group. RugCheck did not flag them; the block shows only in the trading record. None of the 143 tokens that reached 10x was flagged: sells were observed for all 143 (for 142 within 2 hours after the peak), and none is in the buy-only cluster. Sellable at the peak is not safe to hold, though: three-quarters of the non-curve 2x winners had their liquidity pulled or pair removed within 6 hours of the peak.

Group and exit	As reported (n)	Likely honeypots excluded	Likely honeypots scored -100%
All scorable, hold 6 h	-15.6% (52,348)	-16.5% (-17.0 to -16.1)	-16.8% (-17.2 to -16.3)
All scorable, sell at the peak, $100	-4.8% (67,734)	-5.5% (-6.9 to -4.1)	-5.8% (-7.2 to -4.2)
Top 10%, hold 1 h	+6.5% (6,414)	-1.1% (-7.7 to +6.5)	-3.3% (-9.6 to +3.8)
Top 10%, hold 6 h	-19.4% (5,699)	-28.0% (-31.4 to -24.0)	-29.6% (-33.0 to -25.9)
Top 10%, TP100/SL35/6 h	+15.6% (3,447)	+12.4% (+7.9 to +17.3)	+8.8% (+4.5 to +13.6)
Top 10%, sell at the peak, $25 / $100	+144.1% / +112.3%	+139.7% / +107.1%	+134.2% / +102.4%
≥ 2x, hold 1 h	+115.6% (1,623)	+92.7% (+66.3 to +126.4)	+77.7% (+52.1 to +107.5)
≥ 2x, hold 6 h	+26.2% (1,434)	-5.6% (-19.6 to +10.0)	-13.2% (-26.2 to +0.7)
≥ 2x, sell at the peak, $25 / $100 (mean)	+546% / +429%	+559% / +432%	+508% / +391%
≥ 2x, sell at the peak, $25 / $100 (median)	+224% / +212%	+208% / +197%	+192% / +181%

Group and exit	As reported (n)	Likely honeypots excluded	Likely honeypots scored -100%
≥ 10x, every exit	unchanged	unchanged	unchanged

Table 31. *Section 9.7 Honeypots and unsellable winners*

The winner-prediction models were affected most, because a buy-only honeypot looks exactly like the activity the models learned to favour. In the gradient-boosting top-10% model's top 1% of test picks, 33 of 189 were likely honeypots (37 of 947 in the top 5%). Counting those as non-winners, its precision at the top 1% falls from 91.5% to 74.6% (logistic regression: 90.5% to 73.5%). The ≥ 2x models picked almost none (0-1 in the top 1%).

Model picks (test period)	Exit	As reported	Honeypots excluded	Honeypots scored -100%
Top-10% model (GB), top 1%	1 h	-12.5% (n 152)	-29.6% (-46.4 to -10.9)	-43.9% (-57.8 to -28.2)
Top-10% model (GB), top 1%	6 h	-59.4% (n 104)	-82.3% (-90.2 to -72.7)	-85.2% (-91.8 to -77.5)
Top-10% model (GB), top 1%	TP100/SL35/6 h	-26.3% (n 88)	-46.3% (-62.9 to -28.4)	-56.7% (-70.5 to -41.8)
Top-10% model (LR), top 1%	1 h	-3.6% (n 155)	-18.4% (-36.1 to +1.4)	-35.2% (-50.4 to -19.7)
Top-10% model (GB), top 5%	6 h	-39.6% (n 465)	-43.3% (-49.2 to -36.2)	-45.4% (-51.0 to -38.7)
≥ 2x model (GB), top 1%	6 h	-72.1% (n 104)	-72.1%	-72.1%
Every test token	6 h	-9.4% (n 7,000)	-9.6%	-9.9%

Table 32. *Section 9.7 Honeypots and unsellable winners*

Retraining the models with the honeypots removed from the training data does not rescue them: the gradient-boosting top-10% model's top 1% then held only 3 honeypots, kept a precision of 91.5%, and lost 32.5% at 1 h and 86.2% at 6 h (CI -93.3 to -78.7). Across the 400 strategies of section 8, scoring the likely honeypots as -100% cuts the strategies with a positive mean from 7 to 4 and raises the significantly negative ones (Bonferroni) from 182 to 228 of 311; none becomes significantly positive.

Honeypots made the hindsight winner groups and the model picks look better, never worse. Without them the ≥ 2x group's 6 h gain is not demonstrated in any version (-6% to -13% here; -19% to +13% under the censoring checks of 9.4), the top-10% group's 1 h gain disappears, and the models' best picks lose more; the sell-at-the-peak figures fall only modestly because no 10x token was flagged. Limits: zero recorded sells is strong evidence but not proof that sells were blocked; a token with a few sells may still have blocked most holders (only the creator might be able to sell); RugCheck data is missing for about a fifth of winners; and no real sell transaction was attempted.

9.8 Are the section 9 losses just costs on frozen prices?

Also added after the internal review. About 95% of scorable tokens were on a bonding curve when first seen, and the price of 57% of them (1 h) and 54% (6 h) never changed after purchase, so for those tokens the loss in 9.4 is just the round-trip cost. The buy-everything result was therefore recomputed under the cost scenarios of section 4.4 (likely honeypots included, which flatters it).

Cost scenario	All, 1 h (n 59,845)	All, 6 h (n 52,348)	Price moved, 1 h / 6 h	Non-curve tokens, 6 h (n 2,402)
Lab model	-12.6% (-13.3 to -11.7)	-15.6% (-16.1 to -15.1)	-19.2% / -25.2%	-58.7%
Low cost (0.30% + 0.5%)	-9.3% (-10.1 to -8.5)	-12.5% (-13.0 to -12.0)	-16.3% / -22.4%	-57.2%
No fees or slippage (impact only)	-7.8% (-8.5 to -6.9)	-11.0% (-11.5 to -10.5)	-14.8% / -21.1%	-56.5%
No cost of any kind (raw price)	-5.7% (-6.5 to -4.7)	-9.1% (-9.6 to -8.6)	-12.7% / -19.4%	-56.1%

Table 33. *Section 9.8 Are the section 9 losses just costs on frozen prices?*

With no costs at all the CI is still below zero at both horizons (-10.6% at 6 h with dead-before-gap tokens counted). Tokens whose price did move (25,873 at 1 h, 24,002 at 6 h) lost 12.7% and 19.4% raw, and tokens already in a pool when first seen lost about 56% raw at 6 h. Costs enlarge the losses but do not cause them.

9.9 Could the 10x tokens be picked in advance? (post-hoc exploratory)

> **Post-hoc exploratory.** Rules, thresholds and indicators were chosen after seeing the data, on a 70/30 development/held-out split by launch time, with Bonferroni correction across the 456-617 rules tested at each decision time. Nothing here is a confirmed edge.

Coverage: all 143 tokens that reached 10x, 286 matched look-alikes (two per winner, matched on FDV, buys, sells, 10-minute return, day and curve status; none reached 10x) and 323 random scorable tokens with GeckoTerminal 1-minute candles; launch-second on-chain trades were decoded for 78 winners, 166 look-alikes and 52 random pump.fun tokens. P&L uses the study's cost model with likely honeypots at -100%.

- Best rule: at 5 minutes, already graduated to PumpSwap with buys outnumbering sells at least 7:1. Held-out precision 3 of 22 (13.6%) against a 0.25% base rate (Bonferroni p = 0.010), but recall only 3 of 38 held-out 10x tokens, and P&L with the +100% / -35% / 6 h exit was +30% (CI -10% to +70%, n = 22). The precision lift is supported; the profit is not proven. No other rule or model had a significant held-out profit.
- Indicators (RSI, MACD, VWAP, Bollinger, EMA, OBV and volume measures) did not separate future 10x tokens from look-alike losers before about 15 minutes; at 30 minutes they mostly describe the move already under way (winners were up 3.0x from their first candle, look-alikes 0.64x). The only two Bonferroni-significant positive held-out P&L rules (normalised OBV and up-volume share at 10 minutes) rest on 1-2 random false positives each, so they are leads to retest, not findings.
- Timing: the median 10x token peaked 86 minutes after launch and stayed within 80% of its high for a median of 5.5 minutes on the lab's polls (7.5 minutes on 1-minute candles). At 30 minutes only 51 of 59,007 tokens still had a 10x ahead of them. Even a hindsight buy of the 143 at 30 minutes, held 60 minutes, had a median of -6% (n = 130).
- Look-alikes: at 10 minutes, tokens inside the winners' interquartile range on FDV, buys, sells, liquidity, buy/sell ratio and return so far were 10 winners and 376 losers, about 38 losers per winner.
- Who made the money: the creator, same-slot bundlers and snipers who bought in the first ~15 seconds held the largest paper multiples (the median winner ran 22x from its first-minute price against 15.5x from the lab's ~5-minute first price). Whether they realised those gains was not measured. No launch-second feature (dev buy, bundling, early-buyer share, early selling) separated winners from look-alikes (AUC 0.48-0.54).

Caveats specific to this section. The winner definition uses the peak in hindsight, so the winner-only rows in 9.3-9.4 cannot be traded. Peak multiples include single-poll spikes that may be data artefacts (one token shows 1.8 million x from one poll); none of the conclusions depends on them, because they appear only in hindsight groups. DexScreener keeps showing a stale price for tokens that no longer trade, which is why so many medians equal the round-trip cost. DexScreener also drops many curve pairs after about a day, so 24 h outcomes for curve tokens are unreliable and are not used for any conclusion here (for that reason the dead share jumps from 8% at 6 h to 76% at 24 h). The bonding-curve liquidity is modelled, not observed.

10. Discussion

Why a public-data buyer loses. Three effects stack. The base rate is terrible: about two-thirds of tradable new tokens are dead within 6 hours, and smaller ones are worse. Public data cannot see what kills tokens: no creator, holder, flow, liquidity or social feature separated survivors from rugs consistently held-out, because the decisive information (who holds the supply and when they will sell) is not public at 30 minutes, and the activity features that do predict movers also predict dumps; even candle indicators did not separate future 10x tokens from matched look-alike losers before about 15 minutes (section 9.9). And costs are heavy: about 5.3% per round trip on a $25 stake (2.5% fees, 2% slippage, 0.8% network) before price impact, against a median outcome of -100%.

Where memecoin profit probably comes from. The study does not show that nobody makes money on memecoins. It shows that a late, rules-based outsider buying from public data does not. The data is consistent with profits being captured by three groups not tested here:

- Being earlier: snipers and bots that buy in the first seconds or blocks, before the 30-minute decision point. By the time a token appears in public data, the early buyers are often the sellers. In section 9.9 the creator, bundlers and snipers who bought in the first ~15 seconds held the largest paper multiples (22x from the first-

minute price for the median 10x token, against 15.5x from the lab's first price), though whether they realised them was not measured.

- Insiders and creators: on pump.fun, tokens from repeat creators lost about 86-98%, and tokens from creators with 3 or more launches lost about 98%. The losses outsiders take are largely the gains of whoever held supply at launch.
- Fees: launchpads, aggregators, priority-fee and MEV searchers, and promoters (paid X posts were attached to some of the worst tokens) collect revenue whether a token lives or dies.
- Survivorship in public anecdotes: huge winners exist (2 of 422 skipped tokens reached 10x), but they were not buyable at decision time and are swamped by losses across the full distribution.

11. What would change our conclusion

The early-entry loss is proven on this data; what is open is whether late entries (6 h or more after the decision) or sniping have an edge. The criteria below, derived from the observed data, are what would justify saying "late entries have an edge" or "sniping works". The sample sizes use $n = ((z_{1-\alpha} + z_{power}) \times \sigma / \delta)^2$, with a one-sided α of 0.05 ($z = 1.645$), 80% power ($z = 0.842$), a target effect δ of +5% per trade and σ the observed standard deviation of per-trade returns.

Cell (6 h delay unless stated)	Observed n	Observed SD	n to detect +5%	Bonferroni over 400 (z = 3.66)	Detectable now (80% power)
Hold 6 h, lab costs (mean -3.6%)	299	87.8%	1,908	6,258	12.6%
Hold 6 h, best single trade removed	298	46.7%	541	—	—
Hold 6 h, low cost (mean -0.1%)	299	90.7%	2,036	—	—
Hold 6 h, no fees or slippage (mean +1.6%)	299	92.2%	2,103	6,899	13.3%
Hold 1 h, lab costs	322	97.0%	2,325	—	—
Hold 12 h, lab costs (mean +1.0%)	252	163.4%	6,602	—	—
12 h delay, hold 6 h, lab costs	229	32.4%	261	854	—
Delay 0, hold 6 h (proxy for early-entry variance)	1,187	137.9%	4,705	15,436	—

Table 34. *Section 11. What would change our conclusion*

Worked example: for the 6 h delay / 6 h hold cell, $(1.645 + 0.842) \times 0.878 / 0.05 = 43.7$, and $43.7^2 \approx 1{,}908$ trades; with a Bonferroni z of 3.66 for 400 strategies it is $(3.66 + 0.842) \times 0.878 / 0.05 = 79.1$, or about 6,258. With today's 299 trades the smallest effect detectable at 80% power is $(1.645 + 0.842) \times 0.878 / \sqrt{299} = 12.6\%$ per trade, so an edge of +5% could not have been seen. About 21% of universe tokens (299 of 1,411) yield a 6 h-delay trade, and the universe arrived at about 26.5 tokens per observed hour (1,411 in 53.3 h), so at 90% coverage the lab would collect about 121 such trades a day: about 16 days for 1,908 and about 52 days (7.5 weeks) for 6,258. That rate was measured mostly on 18-20 Oct and may differ in other markets.

- Coverage: at least 90% of hours observed over several weeks spanning more than one market regime (against 25% of 215 hours now, mostly 18-20 Oct), with the result positive in each regime or period separately, not only pooled.
- Late entries have an edge: a pre-registered late-entry rule, fixed before the data, with a mean above zero and a one-sided 95% CI lower bound above zero on at least the sample in the table (about 1,900 trades for one pre-specified cell, about 6,300 if it is chosen from the 400-strategy grid), positive after removing the best trade, positive on a held-out later period and under the censoring checks of section 2.5.
- Live fills: real $25 round trips on the same rule whose realised fees, slippage, failed transactions and fill prices match the model's roughly 5.3% round trip (1.25% fee and 1% slippage per leg plus $0.10 a transaction). The pilot should be sized from its own observed cost variance; a realised cost above the model would shrink any paper edge, since the 6 h-delay median of -5.2% is the modelled round trip itself.
- Sniping works: not tested here. It needs a separate pre-registered test with first-block or sub-second on-chain fill data, measuring real entry and exit prices against creators and bots. The variance of the earliest entries measured here (SD 137.9% at delay 0) implies at least about 4,700 trades to detect +5% at 80% power; that is a

lower bound, since launch-second returns are likely more variable. It would also need the early buyers' sells, which the launch-second data of section 9.9 do not cover.

12. Limitations

- Coverage: only about 53 of 215 hours were observed, mostly on 18-20 Oct, so the results describe a few days of market conditions and could differ in a hotter market.
- Sample size: the secondary pre-registered studies did not reach their planned samples (pump.fun 175/300, alpha404 28/100), so for those the size and direction of the effect, not the formal verdict, carry the conclusion. Late-entry samples are too small to detect a +5% edge (section 11).
- Censored horizons: a horizon is only scored if the collector was running at that moment. Recomputing with tokens dead before a gap at -100% gives -65.3% at 6 h, and -60.9% with live tokens also marked at their last price, so the headline sits between the two and gaps do not explain the loss (sections 2.5 and 4.4).
- Paper trading: there were no real fills. The cost model (1.25% fee and 1% slippage per leg, constant-product impact, 20 s latency) is probably optimistic for fast rugs, sandwich attacks and failed transactions, which would make the true results worse.
- Snapshot exits: stops and targets were checked every 5 minutes, so stop fills during sharp drops are optimistic, and some brief take-profit touches may have been missed.
- Data quality: RugCheck was down on 18-19 Oct; 17,337 rug checks were skipped because of backlog; 594 pump.fun creator lookups failed on RPC errors; there were 34 database-lock errors; FDV artefacts appeared on flagged tokens; and pump.fun post-graduation prices had a look-ahead artefact. Affected values were excluded.
- Collectors stopped: the pump.fun, alpha404 and rugalt collectors died after a restart on 23 Oct and were deliberately not restarted, so their samples end on 23 Oct.
- Held-out design: the held-out period is later in time and has thinner coverage, so it tests time stability as well as overfitting, and held-out cells are small.
- Multiple testing: the roughly 240 exploratory cells in section 6 were evaluated without correction, so isolated positive cells there should be expected by chance. The 416 strategies in section 8 were corrected (Bonferroni and White's reality check).
- Winners study: see the caveats at the end of section 9; 23,923 of 91,657 tokens had no scorable price history.
- 10x deep dive (section 9.9): post-hoc, in one 9-day window, with only 38, 25, 17 and 9 held-out 10x tokens at 5, 10, 15 and 30 minutes. Candles cover 323 of a planned 700 random tokens; launch trades were decoded only for pump.fun tokens (78 of the 143 winners) and only their first 30 transactions, so insiders' realised gains were not measured. Look-alikes were matched on the 10-minute return, which flatters them in 5-minute comparisons.
- X data: only 69 tokens were enriched and 38 were capped at 25 posts; none of the X results is pre-registered.

13. Conclusion

Systematic post-launch buying of new Solana memecoins is proven to lose on average. Buying all 1,411 tradable tokens at about 30 minutes of age lost 31% after 1 hour, 62% after 6 hours (95% CI -68% to -53%) and 56% after 24 hours, net of retail costs, with every CI far below zero and in both halves of the period. It is not a cost artefact (-59.4% at 6 h with no costs of any kind), an effect of the dead-token rule (at most about 4 points) or of collector gaps (-65.3% with tokens dead before a gap counted), and removing likely honeypots makes it worse: -74.2% at 6 h, or -77.0% counting them as -100% (section 4.4). Two-thirds of tokens were dead within 6 hours, and only about 1 in 75 had doubled. Filtering did not rescue it: the pump.fun creator filter, the on-chain rug check, the established-token plus X-sentiment strategy and the X promotion check were all negative, and no exploratory timing, momentum, rebound, graduation or listing rule survived the held-out period.

Changing how or when the tokens are traded does not help either, and this is shown with corrections for multiple testing. Of 416 strategies, none had a significantly positive mean even before correction, 182 of the 311 with at least 30 trades lost with significance after Bonferroni correction, and White's reality check found the best one indistinguishable from luck ($p = 0.35$). The strategy that looked best on development data (+85.2%) lost 81.2% held-out, with a CI entirely below zero, so picking the best strategy in hindsight is proven to fail. The losses become too

small to prove only for entries at about 6 h of age (-3.6% to +1.0% on about 300 trades, CIs spanning zero, median -5.2%, the cost of a round trip). At low or zero cost these cells turn slightly positive in the full sample (+2.2%, CI -6.7% to +14.2%, at raw prices) but are mostly negative held-out, and none of the 400 strategies was significantly profitable at any cost level. Late entries under low costs are therefore no proven edge, not a proven loss; with 299 trades, an edge smaller than about 12.6% per trade could not have been detected, and section 11 sets out what would settle it.

Real winners existed, including 143 tokens that reached 10x, but they could not be picked out profitably in advance. Of the 67,734 discovered tokens with a tradable price history, 84% never rose above their first price; buying all of them lost 12.6% at 1 h and 15.6% at 6 h (p below 10^{-200}), and even selling each at its exact peak would have lost 4.8% at $100 a trade. Activity at 30 minutes predicts which tokens will move (AUCs of 0.87 and 0.95 on later, unseen tokens), but the top-ranked tokens are also the most likely to be dumped. Buying the models' top picks lost 25-72% at 6 h (33-80% with dead-before-gap tokens counted), with every CI below zero, and in 33 of 36 tests did worse than at least 99.9% of random selections of the same size. Part of the apparent success among winners was not real: 155 of the top-10% winners were likely honeypots, 126 of them a single buy-only cluster that RugCheck did not flag, and without them the 2x group's +26.2% at 6 h is not a demonstrated gain (-5.6%; -12.0% to +12.8% under the censoring checks). The 143 tokens that reached 10x were sellable, but they peaked at a median of 86 minutes for a median of 5.5-7.5 minutes, only 51 of 59,007 tokens still had a 10x ahead at 30 minutes, and the best post-hoc rule found them at 13.6% precision with an unproven profit (+30%, CI -10% to +70%; section 9.9). Candle indicators did not separate them from look-alike losers before about 15 minutes, about 38 losers looked like each winner at 10 minutes, and the largest paper gains went to the creators, bundlers and snipers of the first seconds, whose realised gains were not measured. In the data available here, telling the winners apart did not translate into making money from them.

What this study does not show. It does not show that nobody makes money on memecoins: sniping in the first seconds, insider or creator supply and fee income from launchpads, aggregators, bots, MEV and paid promotion (the most plausible sources of reported profits, none available to an ordinary public-data buyer), CEX listings, established memecoins, other chains and other months were not tested. Only about 25% of the window was observed (about 53 of 215 hours, mostly 18-20 October), in one market regime, with fills simulated from 5-10 minute snapshots. Within what was tested, systematically buying new Solana memecoins after launch lost money on average, and late entries offered no proven edge.