

Shadow Entry-Rules Study (entry_rules.py) — 2026-09-26

What we did

The new shadow recorder (your entry_rules.py) was connected to the bot. It watches every stock, writes down what it sees each time the bot's decision changes, and **never buys or sells anything**. Afterwards, a second script (shadow_outcomes.py) checks what the price actually did next: did it hit the plan's target or the plan's stop first, and where was it after 60 minutes?

We ran it on the data we already had:

- **20-day bar data** (8/27 – 9/24): setups J, A (old live) and H. This data has 1-minute bars only, **no individual trades or quotes**, so the order-flow parts (buy/sell imbalance, spread, and therefore the VAS score and the PASS verdict) cannot be calculated here.
- **Tick replay** (9/18 – 9/25, setup J): full trade-by-trade data, so every shadow feature works.

About 20,000 snapshots were recorded per run.

Setup J results (same as before, confirmed)

Data	Trades	Wins	P/L
20-day bars, J	10	6	+\$102.75
20-day bars, A (old live)	44	20	-\$98.21
20-day bars, H	88	40	-\$169.43
Tick replay 9/18-9/25, J	7	2	-\$6.27
Tick replay 9/22, J	1	1	+\$63.83

What the shadow features told us

1. Shadow verdict PASS vs FAIL (tick data only)

Verdict	Bot entries	Win rate	P/L
PASS	4	75%	+\$110
FAIL	4	0%	-\$52

This is the most interesting result — every losing J trade was one the shadow would have rejected. **But it is only 8 trades**. Across all 19,000 tick snapshots the edge is smaller: PASS hit target 47% / stop 27% (+0.47% after 60 min) vs FAIL 45% / 31% (+0.27%). Promising, not proven.

2. Distance above VWAP (strongest signal, both data sets)

Price vs VWAP	Stop hit first (bars)	Avg 60-min move (bars, setup A)
below VWAP	31%	+0.09%
0 – 1% above	28%	+0.04%
1 – 2% above	31%	+0.07%
2 – 4% above	40%	–0.19%
4%+ above	48%	–1.25%

Buying a stock stretched 2%+ above VWAP loses. This confirms the idea behind the J rule (max 3–5% above the open) — the chase is the problem.

3. Beats SPY & IWM (relative strength)

- Bar data: **no difference** (target 41% vs 40%; stop actually slightly worse when "stronger than the market").
- Tick data: a small plus (target 48% vs 43%, 60-min +0.38% vs +0.18%).

Mixed — not good enough to become a rule yet.

4. Volume trend, bar RVOL, imbalance, VAS score

No useful separation. Higher VAS scores did **not** give better outcomes across all snapshots (8-10 band: target 42%, stop 33% — slightly worse than low scores). Wide spreads (0.2%+) had more stop hits (38-42% vs 27%).

Conclusions

1. **Nothing here is ready to turn on.** The shadow stays shadow-only.
2. The one feature worth watching is the **shadow PASS verdict** — it needs live ticks, which Monday gives us.
3. The **VWAP-stretch finding supports setup J** and suggests a possible future rule: "don't buy more than ~2% above VWAP". Not enabled — idea only.
4. Relative strength vs SPY/IWM and the volume-trend features did not earn their place so far.

Monday 9/28 — what is running

- **Setup J live:** max 3/4/5% above the open (by volume pace and buy pressure) + minimum reward/risk 2.0. Everything else as last week.
- **Shadow recorder ON**, recording only (writes to logs/shadow/). SPY and IWM are streamed as 1-minute bars only, so they don't slow the bot's data feed.
- After the close: label Monday's shadow file with shadow_outcomes.py and add it to this study. Every live day adds real order-flow data, which the 20-day set lacks.