Sample Expansion Strategy for StatArb Engine

This document outlines a staged framework for expanding the sample set of the StatArb Engine while maintaining rigorous constraints on internal and external validity. The approach emphasizes rejecting the null hypothesis only after crossing defined acceptance gates.

Stage 0 — Baseline (current)

• Universe: SPY20 / NQ20 / core ETFs. • Split: rolling walk-forward (train N days → test M days). • Gate G0: Stable monotonicity of signal→return, IR>0.3, bootstrap p<0.05.

Stage 1 — Time expansion (same names, more regimes)

• Add: pre-COVID, COVID shock, post-2022 hikes, 2024–2025 AI regime. • Gate G1: Effect size roughly non-decaying across regimes, similar drawdowns, turnover stable.

Stage 2 — Cross-section expansion

• Add: sector-balanced sample (8 sectors × 10 names), cap buckets (mega/large). • Gate G2: Sharpe within 80–120% of baseline, hit rate persists in ≥60% of sectors.

Stage 3 — Market-structure diversity

• Add: mid-caps, ADRs with liquidity, sector ETFs. • Gate G3: Net alpha survives realistic frictions; edge not concentrated in one corner.

Stage 4 — Alternative constructions

• Build pairs via cointegration, economic linkage, factor-neutral residuals. • Gate G4: Rank-order consistency across constructions; Spearman rho ≥ 0.25.

Stage 5 — Out-of-domain (optional)

• Add: EU/JP equities, sector futures vs ETF proxies. • Gate G5: Require directional validity, positive IR after frictions.

Practical Guardrails

Practical Guardrails: • ≥1,000 independent OOS trades per stage (≥200 per regime slice). • Control multiple testing via FDR. • Use walk-forward only (no global refit). • Check robustness to events (earnings, halts). • Test with pessimistic slippage (2–5×). • Log stability metrics: IR, Q5–Q1 spread, turnover, hit rate, max DD, tail ratio, regime dispersion.

Conclusion

We restrict inferential claims to US, liquid equities/ETFs. Expansion proceeds only after gates (G0–G3) are passed. This staged approach maximizes internal validity before pursuing external validity.