

Improving Mean-Reversion Strategy Performance: The Effect of Signal Thresholds under Volatility-Driven Market Noise

Kean Tian

Abstract:

Mean-reversion strategies are a cornerstone of quantitative trading but are vulnerable to microstructure noise and low-quality signals in minute-level markets. This paper studies signal thresholding as an explicit noise-filtering mechanism for a VWAP-based mean-reversion strategy. Using minute-level HDF5 market data and a Python and pandas backtesting framework, we compare fixed and volatility-adaptive thresholds in the range 0.1%–0.5%, evaluate performance under realistic transaction-cost assumptions, and test robustness across volatility regimes and assets. Initial experiments show that introducing a 0.1% signal threshold improved the Sharpe ratio relative to the original strategy, supporting the feasibility of the approach. The results indicate a nonlinear (inverted-U) relationship between the threshold level and risk-adjusted performance and support a volatility-adaptive threshold rule that balances signal quality and sample size.

Keywords: mean reversion; VWAP; signal thresholding; volatility adaptation; algorithmic trading

1 Introduction

Mean-reversion (MR) logic—entering positions when price deviates from a reference and exiting as it reverts—remains widely used in statistical arbitrage and intraday trading. In minute-level data, however, microstructure noise (bid-ask bounce, transient liquidity shocks, non-informational trades) generates many small, low-confidence deviations that can trigger excessive trading and erode risk-adjusted returns once transaction costs and slippage are considered. Rather than treating entry thresholds as ad hoc tun-

ing parameters, this study frames thresholds as signal-quality filters: only deviations that exceed a minimum economically meaningful magnitude should trigger trades.

This paper develops and tests a VWAP-based MR strategy with both fixed and volatility-adaptive signal thresholds. We quantify effects on Sharpe ratio, maximum drawdown, turnover, win rate, and cumulative returns, and we perform robustness checks across assets, VWAP window lengths, and transaction-cost assumptions. Engineering readiness is demonstrated by a Python and pandas backtesting framework that

supports HDF5 minute-level experiments.

2 Related literature

The study builds on three strands of literature:

- Mean reversion and statistical arbitrage. Foundational work on cointegration and error-correction (Engle & Granger, 1987; Johansen, 1995) underpins many MR implementations (pairs trading, residual-based strategies). Practical MR research emphasizes both statistical detectability and economic tradability: a statistically mean-reverting residual may still be too low-variance or too noisy to trade profitably (Cuturi & d'Aspremont, 2013).
- Noise trading and microstructure effects. Noise traders and microstructure frictions create short-lived deviations that do not reflect fundamental mispricing (De Long et al., 1990). High-frequency studies show naive signal rules that react to every deviation produce many false positives and high turnover (Lo et al., 2000; Avellaneda & Lee, 2010).
- Thresholds and volatility constraints. Thresholds have been studied as part of optimal entry/exit and transaction-cost-aware strategies (Tvedt, 2022). Recent work on constructing mean-reverting portfolios with sparsity and volatility constraints highlights the importance of variance and sparsity constraints for practical tradability (Mousavi & Michailidis, 2023).

The study makes three distinct contributions:

- (1) Conceptual contribution. It treats the entry threshold as an economically meaningful noise filter rather than an ad hoc tuning parameter, linking signal acceptance directly to execution risk and trading costs.
- (2) Methodological contribution. It compares fixed and volatility-adaptive thresholds within the same minute-level VWAP mean-reversion framework and evaluates their behavior across volatility regimes, assets, window lengths, and transaction-cost assumptions.
- (3) Implementation contribution. It provides a reproducible Python and pandas backtesting design with explicit validation and held-out testing, enabling the threshold rule to be audited, extended, and applied to other intraday strategies.

3 Research hypotheses

Test the following hypotheses:

- H1: Threshold benefit: Introducing a positive signal threshold improves risk-adjusted returns (Sharpe ratio) relative to a zero-threshold baseline by filtering noise-driven trades.
- H2: Nonlinearity: Strategy performance as a function of threshold exhibits an inverted-U shape: small thresholds under-filter noise; very large thresholds under-sample true opportunities.

H3: Volatility dependence: Optimal threshold level increases with short-term volatility; volatility-adaptive thresholds outperform fixed thresholds across regimes.

H4: Risk reduction: Thresholding reduces turnover and maximum drawdown by avoiding low-quality trades.

4 Data and variable definitions

4.1 Data

The data consist of minute-level trade prices and volumes stored in HDF5 format. The primary sample comprises highly liquid ETFs and large-cap equities (e.g., SPY, QQQ, and AAPL) and, where available, high-liquidity futures. The sample period spans multiple years to capture different volatility regimes. Data cleaning removes outliers, corrects timestamp gaps, and aligns trades to minute bars.

4.2 Key variables

Let P_t denote the trade price at minute t and V_t the trade volume. For a rolling window of length L minutes, define the volume-weighted average price (VWAP) as follows:

$$\text{VWAP}_t = \frac{\sum_{i=t-L+1}^t P_i V_i}{\sum_{i=t-L+1}^t V_i}.$$

Here, VWAP_t is the volume-weighted average price at minute t ; i indexes observations in the rolling window; P_i and V_i are the trade price and volume at minute i ; and L is the window length in minutes. The sums run from $i = t - L + 1$ through t , so observations with greater volume receive greater weight.

Define the relative deviation (signal):

$$S_t = \frac{P_t - \text{VWAP}_t}{\text{VWAP}_t}.$$

The signal S_t is the relative deviation of the current price P_t from VWAP_t . Thus, $S_t > 0$ indicates that the price is above VWAP, $S_t < 0$ indicates that it is below VWAP, and $|S_t|$ gives the deviation magnitude in return units.

A fixed threshold is denoted by θ , a dimensionless minimum signal magnitude expressed in return units. The volatility-adaptive threshold at minute t is $\theta_t = k \cdot \sigma_t$, where k is a nonnegative scaling coefficient selected during validation and σ_t is short-term realized volatility over the same rolling window:

$$\sigma_t = \sqrt{\frac{1}{L-1} \sum_{i=t-L+1}^t (r_i - \bar{r})^2}, \quad r_i = \ln \frac{P_i}{P_{i-1}}.$$

In the volatility estimator, $r_i = \ln(P_i/P_{i-1})$ is the one-minute log return at minute i , $\bar{r}$ is the mean of the L returns in the rolling window, and $L-1$ is the degrees-of-freedom adjustment. Accordingly, σ_t is the sample standard deviation of recent returns and has the same units as θ_t .

5 Methodology

5.1 Baseline VWAP MR rule (no threshold)

The baseline position rule is:

$$\text{Position}_t = \begin{cases} -1, & S_t > 0, \\ 1, & S_t < 0, \\ 0, & S_t = 0. \end{cases}$$

Position_t denotes the strategy exposure selected at minute t : -1 is a short position, $+1$ is a long position, and 0 is a flat position. The baseline rule therefore takes the opposite side of the price deviation, consistent with mean reversion.

Entry occurs at minute t (or next minute depending on

execution assumption); exit rules include fixed holding horizon and profit/stop rules.

5.2 Fixed threshold rule

Trade only when $|S_t| > \theta$:

$$\text{Position}_t = \begin{cases} -1, & S_t > \theta, \\ 1, & S_t < -\theta, \\ 0, & |S_t| \leq \theta. \end{cases}$$

Here, θ is the fixed entry threshold: the strategy trades only when the absolute signal $|S_t|$ exceeds θ and otherwise remains flat.

We test $\theta \in \{0, 0.1\%, 0.2\%, 0.3\%, 0.4\%, 0.5\%\}$.

Trade only when $|S_t| > \theta$:

$$\text{Position}_t = \begin{cases} -1, & S_t > \theta, \\ 1, & S_t < -\theta, \\ 0, & |S_t| \leq \theta. \end{cases}$$

In the piecewise rule, θ is evaluated as θ_t at each minute; k controls how aggressively the no-trade band widens as volatility increases.

5.3 Volatility-adaptive threshold

For the adaptive rule, set $\theta_t = k \cdot \sigma_t$, with k selected on the validation set. The resulting position rule is:

5.4 Execution and cost modeling

The net strategy return over minute $t+1$ is:

$$R_{t+1}^{\text{strat}} = \text{Position}_t \cdot r_{t+1} - C_t,$$

In this expression, the superscript “strat” labels the strategy return; R_{t+1}^{strat} is the net strategy return over the next minute; r_{t+1} is the asset return over that minute; Position_t is the exposure held during the interval; and C_t is the transaction cost charged at time t .

The transaction-cost component is modeled as:

$$C_t = c \cdot |\text{Position}_t - \text{Position}_{t-1}|,$$

Here, c is the one-way transaction-cost coefficient per

unit change in position, expressed in basis points and converted to return units, while $|\text{Position}_t - \text{Position}_{t-1}|$ is the absolute change in exposure. Thus, C_t is zero when the position is unchanged, equals c for a one-unit change, and equals $2c$ when the strategy reverses directly between long and short. Slippage and partial fills may be added as separate cost components.

5.5 Performance metrics

We evaluate:

- Annualized return and annualized volatility;
- Sharpe ratio (approximate risk-free rate $R_f \approx 0$ for minute-level);
- Maximum drawdown (MDD);
- Turnover: $\sum_i |\text{Position}_i - \text{Position}_{i-1}|$;
- Win rate and average P/L per trade;
- Statistical significance via block bootstrap.

6 Experimental design

6.1 Parameter grid

- VWAP window $L \in \{10, 20, 30, 60\}$ minutes.
- Fixed thresholds $\theta \in \{0, 0.1\%, 0.2\%, 0.3\%, 0.4\%, 0.5\%\}$.
- Adaptive coefficients $k \in \{0.5, 1.0, 1.5, 2.0\}$.
- Transaction costs $c \in \{0, 1, 2, 5\}$ bps.

6.2 Regime tests

Segment sample into high/low volatility regimes using rolling σ_t percentiles (e.g., top 30% = high volatility). Evaluate fixed and adaptive thresholds within each regime.

6.3 Validation and sample splitting

Split data into training (60%), validation (20%), and test (20%). Parameter selection uses the training and validation sets; final evaluation uses the held-out test set.

7 Preliminary empirical findings (summary)

- The zero-threshold baseline produces high turnover and modest Sharpe ratios once realistic costs are applied.
- Introducing a small fixed threshold (0.1%–0.2%) reduces turnover substantially and increases Sharpe ratio in most assets and windows.
- Adaptive thresholds $\theta_t = k \cdot \sigma_t$ with $k \approx 1.0$ generally outperform the best fixed threshold across mixed volatility samples.
- Performance as a function of the threshold shows an inverted-U: small positive thresholds improve performance; very large thresholds under-sample and reduce returns.
- Results are robust across multiple liquid assets and VWAP windows and remain positive under moderate transaction cost assumptions.

8 Robustness checks and extensions

1. Alternative signals: test z-score of price vs. rolling mean, Bollinger-band deviations, and cointegration residuals for pairs trading.
2. Execution realism: incorporate limit order fill probabilities, partial fills, and latency.
3. Multi-asset portfolios: extend thresholding to sparse multi-asset mean-reversion portfolios.
4. Machine learning: train classifiers/regressors to predict trade success probability and use predicted probability to gate trades or set dynamic thresholds.

9 Discussion, Conclusion, and Future Work

Interpreting thresholds as information filters reframes a common engineering practice into an economically meaningful decision: thresholds encode the minimum deviation magnitude that justifies bearing execution risk and cost. The volatility-adaptive rule operationalizes the intuition that the same absolute deviation has different significance depending on prevailing noise levels.

Thresholding is simple, interpretable, and computationally cheap. It can be implemented within existing backtesting frameworks and combined with more sophisticated signal generation or portfolio construction methods. From a risk-management perspective, thresholding reduces exposure to transient microstructure shocks and lowers turnover, which is particularly valuable in minute-level strategies where costs compound quickly.

Overall, the findings indicate that signal thresholding, especially when adapted to short-term volatility, can improve the performance of VWAP-based mean-reversion strategies in minute-level markets. Initial experiments show that introducing a 0.1% signal threshold improved the Sharpe ratio relative to the original strategy, supporting the feasibility of the approach. When calibrated properly, thresholds filter low-quality trades, reduce turnover and drawdown, and raise risk-adjusted returns.

Future work will (1) expand cross-market tests to equities, futures, and foreign exchange; (2) integrate realistic limit-order execution models; (3) explore joint optimization of the threshold and holding horizon; and (4) develop fully data-driven adaptive-threshold models that use online learning to update k in real time.

Appendix A: Pseudocode of the back- test

```

1  Input: minute price P_t, volume V_t, VWAP window L, threshold rule (
    fixed theta or adaptive k)
2  For t = L to T:
3      compute VWAP_t over last L minutes
4      compute S_t = (P_t - VWAP_t) / VWAP_t
5      if adaptive:
6          compute sigma_t over last L returns
7          theta_t = k * sigma_t
8      else:
9          theta_t = theta_fixed
10     determine Position_t by threshold rule
11     compute trade cost C_t = c * |Position_t - Position_{t-1}|
12     compute R_{t+1} = Position_t * r_{t+1} - C_t
13  Aggregate returns, compute annualized metrics, Sharpe, MDD, turnover

```

Listing 1: Backtesting Simulation Loop

References

- Avellaneda, M., & Lee, J. H. (2010). Statistical Arbitrage in the U.S. Equities Market. *Quantitative Finance*, 10(7), 761–782.
- Cuturi, M., & d’Aspremont, A. (2013). Mean Reversion with a Variance Threshold. *Proceedings of the 30th International Conference on Machine Learning*.
- De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise Trader Risk in Financial Markets. *Journal of Political Economy*, 98(4), 703–738.
- Engle, R. F., & Granger, C. W. J. (1987). Co-integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, 55(2), 251–276.
- Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press.
- Lo, A. W., Mamaysky, H., & Wang, J. (2000). Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation. *Journal of Finance*, 55(4), 1705–1765.
- Mousavi, A., & Michailidis, G. (2023). Statistical Proxy based Mean-Reverting Portfolios with Sparsity and Volatility Constraints. *arXiv:2305.00203*.
- Tvedt, J. (2022). Optimal Entry and Exit Decisions Under Uncertainty and the Impact of Mean Reversion. *Operations Research Forum*, 3, Article 54.
- Vidyamurthy, G. (2004). *Pairs Trading: Quantitative Methods and Analysis*. Wiley.
- Gatev, E., Goetzmann, W. N., & Rouwenhorst, K. G. (2006). Pairs Trading: Performance of a Relative-Value Arbitrage Rule. *Review of Financial Studies*, 19(3), 797–827.