Statistical Arbitrage in Cryptocurrency Markets:

Cointegration-Based Pairs Trading on BTC/ETH

Colin Whetstone

Texas State University

2026

Abstract

This paper presents a statistical arbitrage system for the Bitcoin (BTC)–Ethereum (ETH) cryptocurrency pair, developed and implemented by the Quant Finance Collective (QFC) at Texas State University. We apply the Engle-Granger two-step cointegration test to establish a long-run equilibrium relationship between BTC/USD and ETH/USD price series. The spread between the two assets is modeled as an Ornstein-Uhlenbeck (OU) process, from which mean-reversion speed, half-life, and equilibrium level are estimated. A signal-generation framework based on z-score thresholds is developed, and the strategy is backtested under realistic transaction-cost assumptions. Permutation testing (1,000 shuffles) is employed to evaluate signal significance. Backtest results yield an annualized Sharpe ratio of 1.48, a total return of approximately 22%, and a maximum drawdown of 8.3%, with a permutation test p-value of 0.062. These findings suggest a statistically marginal but economically meaningful mean-reversion signal, warranting further investigation with a longer data history and live deployment.

Contents

1. Introduction

Statistical arbitrage is a class of quantitative trading strategies that exploit temporary deviations from a theoretical equilibrium between related financial instruments. Unlike pure arbitrage, statistical arbitrage does not guarantee riskless profit but instead profits from the tendency of historically correlated assets to revert toward a stable long-run relationship.

Pairs trading, the most common implementation of statistical arbitrage, involves identifying two assets whose prices are cointegrated, constructing a hedge ratio, and trading the resulting spread when it deviates significantly from its historical mean. The strategy profits from the spread converging back toward equilibrium.

Bitcoin and Ethereum represent the two largest cryptocurrency assets by market capitalization, and their prices exhibit co-movement driven by shared market sentiment, macroeconomic sensitivity, and overlapping investor bases. While the relationship is not guaranteed to persist, it provides a natural laboratory for testing cointegration-based strategies in a high-volatility, continuously traded market.

This paper makes the following contributions. First, we test the BTC/ETH pair for cointegration using the Engle-Granger methodology and estimate the parameters of the resulting spread process. Second, we develop a signal generation framework based on z-score deviations from the estimated equilibrium. Third, we conduct a backtest with transaction costs and evaluate statistical significance through permutation testing. Fourth, we discuss the limitations of the current implementation and directions for improvement.

2. Theoretical Background

2.1 Cointegration

Two time series X_t and Y_t are said to be *cointegrated* if each is individually non-stationary (integrated of order 1, or $I(1)$), but a linear combination

$$Z_t = Y_t - \beta X_t$$

is stationary. The parameter β is the *hedge ratio* or cointegrating coefficient. Intuitively, cointegration captures the idea that two series share a common stochastic trend: they may wander individually, but they cannot drift too far apart without eventually converging.

The Engle-Granger two-step procedure tests for cointegration as follows. In the first step, we regress Y_t on X_t via OLS to estimate the hedge ratio $\hat{\beta}$. In the second step, we test the residuals $\hat{Z}_t = Y_t - \hat{\beta} X_t$ for stationarity using the Augmented Dickey-Fuller (ADF) test. If the residuals are stationary, the pair is cointegrated.

2.2 The Ornstein-Uhlenbeck Process

Once cointegration is established, the spread Z_t is modeled as an Ornstein-Uhlenbeck (OU) process, the continuous-time analog of a mean-reverting AR(1) process:

$$dZ_t = \kappa(\mu - Z_t)\, dt + \sigma\, dW_t,$$

where κ is the mean-reversion speed, μ is the long-run equilibrium level, σ is the diffusion coefficient, and W_t is a standard Brownian motion. The half-life of mean reversion or the expected time for the spread to revert halfway to the mean is given by

$$\tau_{1/2} = \frac{\ln 2}{\kappa}.$$

A short half-life implies rapid mean reversion and greater trading frequency.

2.3 Signal Generation

Standardizing the spread into a z-score

$$z_t = \frac{Z_t - \mu}{\sigma_Z}$$

allows for regime-independent signal thresholds. We enter a *long-spread* position (long Y, short X) when $z_t < -2$, and a *short-spread* position (short Y, long X) when $z_t > +2$. Positions are closed when the z-score reverts to within 0.5 standard deviations of the mean.

3. Data and Methodology

3.1 Data

Price data for BTC/USD and ETH/USD were collected at 10-second intervals using the Kraken exchange API via the CCXT library. The collection script yielded approximately 120,000 synchronized observations per asset. Log prices were used throughout to ensure additive returns and stabilize variance.

3.2 Cointegration Testing

Log price series for BTC and ETH were first tested individually for non-stationarity using the ADF test with automatic lag selection (BIC). Both series failed to reject the unit root null at the 5% significance level, confirming $I(1)$ behavior. The Engle-Granger procedure was then applied: log(ETH) was regressed on log(BTC) via OLS to estimate the hedge ratio, and the residuals were tested for stationarity.

3.3 Spread Modeling

The estimated residual series was fit to a discrete-time OU process using OLS regression on the lagged spread. Mean-reversion speed, equilibrium level, and process volatility were estimated from the fitted AR(1) coefficients. The half-life of mean reversion was derived analytically from the estimated κ parameter.

3.4 Backtest Design

The backtest was designed to reflect realistic trading conditions:

- **Entry threshold:** $|z_t| \geq 2.0$ standard deviations from the mean.
- **Exit threshold:** $|z_t| \leq 0.5$ standard deviations from the mean.
- **Transaction costs:** 0.1% per trade applied to both legs of each spread trade.
- **Position sizing:** fixed notional of $500 per trade, rebalanced at each entry.
- **No look-ahead bias:** parameters estimated on a rolling basis using only past data.

3.5 Permutation Testing

To assess whether the observed performance could be attributable to chance, we conducted a permutation test with 1,000 shuffles. In each iteration, the z-score signal series was randomly shuffled while the price series were kept intact, and the strategy was re-run on the shuffled signal. The empirical p-value was computed as the fraction of permuted Sharpe ratios exceeding the observed Sharpe ratio.

4. Results

4.1 Cointegration Test Results

The ADF test on the residual spread series yielded a test statistic of -3.42, rejecting the unit root null at the 5% significance level (critical value: -3.34). This confirms cointegration between log(BTC) and log(ETH) over the sample period. The estimated hedge ratio $\hat{\beta} = 0.847$, implying that a 1% move in BTC is associated with an 0.847% move in ETH on average.

4.2 Ornstein-Uhlenbeck Parameter Estimates

Table 1: Estimated Ornstein-Uhlenbeck parameters for the BTC/ETH spread.

Parameter	**Symbol**	**Estimate**	**Interpretation**
Mean-reversion speed	κ	0.0041 / tick	Moderate reversion
Half-life	$\ln(2)/\kappa$	$\sim$169 ticks ($\sim$28 min)	Intraday signal
Long-run mean	μ	0.00092	Near-zero spread
Process volatility	σ	0.00031	Low noise level

The estimated half-life of approximately 28 minutes is consistent with an intraday mean-reversion signal, suggesting the spread tends to revert to equilibrium within a single trading session.

4.3 Backtest Performance

Table 2: Backtest performance summary.

Metric	**Value**
Total Return	+22.3%
Annualized Sharpe Ratio	1.48
Maximum Drawdown	-8.3%
Number of Trades	47 round trips
Average Hold Time	$\sim$42 minutes
Win Rate	61.7%
Permutation Test p-value	0.062

The strategy generated a Sharpe ratio of 1.48 over the backtest period, with a total return of 22.3% and a maximum drawdown of 8.3%. The permutation test yielded a p-value of 0.062, which falls marginally outside the conventional 5% significance threshold. While this does not permit rejection of the null hypothesis at the 5% level, it provides suggestive evidence of a signal that warrants further investigation with a longer data history.

4.4 Signal Characteristics

The z-score series spent the majority of the sample period between -2 and $+2$, triggering 47 round-trip trades over the sample period. The signal was most active during periods of elevated cross-asset volatility, when temporary dislocations between BTC and ETH were more frequent. The average hold time of 42 minutes is consistent with the estimated OU half-life, suggesting the mean-reversion model is capturing the dominant dynamic in the spread.

5. Discussion

5.1 Interpretation of Results

The results suggest that a statistically marginal but meaningful mean-reversion signal exists in the BTC/ETH spread over the sample period. The Sharpe ratio of 1.48 is competitive with systematic strategies reported in the academic literature, though the short sample period limits the strength of conclusions that can be drawn.

The permutation test p-value of 0.062 is an reflection of the data. With a limited sample of 10-second data, the effective number of independent signal realizations is constrained, reducing statistical power. A longer data history ideally 6–12 months would help in distinguishing signal from noise.

5.2 Limitations

- **Short sample period:** the data window limits statistical power and out-of-sample generalizability.
- **Stationarity assumption:** the cointegrating relationship may break down during regime changes, stress events, or structural shifts.
- **Execution assumptions:** the backtest assumes market orders at observed prices; slippage and queue position would reduce performance.
- **Parameter stability:** OU parameters may be sensitive to the estimation window length.

- **Single pair:** the strategy is concentrated in one pair; diversification across multiple cointegrated pairs would reduce idiosyncratic risk.

5.3 Extensions and Future Work

Several extensions are planned for future research:

- Live paper trading deployment with real-time z-score monitoring and automated signal generation.
- Expansion to additional cryptocurrency pairs (e.g., SOL/AVAX, BNB/LINK) to build a diversified multi-pair portfolio.
- Johansen cointegration testing for portfolios of three or more assets.
- Dynamic hedge ratio estimation using Kalman filtering to adapt to time-varying relationships.
- Regime-conditional signal filtering: suppressing trades during high-volatility regimes identified by the HALO regime detector.

6. Conclusion

This paper presented a cointegration-based statistical arbitrage system for the BTC/ETH cryptocurrency pair, developed as part of the Quant Finance Collective's research at Texas State University. The Engle-Granger procedure confirmed a cointegrating relationship between log(BTC) and log(ETH), with an estimated hedge ratio of $\hat{\beta} = 0.847$. The spread was modeled as an Ornstein-Uhlenbeck process with a half-life of approximately 28 minutes, consistent with intraday mean reversion.

Backtesting with realistic transaction costs yielded a Sharpe ratio of 1.48 and a total return of 22.3%, with a permutation test p-value of 0.062. These results are promising but require validation over a longer data history and live deployment before strong conclusions can be drawn.

This work represents one component of QFC's broader multi-strategy research program, which also includes the HALO Avellaneda-Stoikov market-making system and a Reddit sentiment-based signal for ALGO/USD.

References

[1] Engle, R.F. and Granger, C.W.J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, 55(2), 251–276.

[2] Ornstein, L.S. and Uhlenbeck, G.E. (1930). On the theory of Brownian motion. *Physical Review*, 36(5), 823–841.

[3] Gatev, E., Goetzmann, W.N., and Rouwenhorst, K.G. (2006). Pairs trading: Performance of a relative-value arbitrage rule. *Review of Financial Studies*, 19(3), 797–827.

[4] Vidyamurthy, G. (2004). *Pairs Trading: Quantitative Methods and Analysis*. John Wiley & Sons.

[5] Avellaneda, M. and Lee, J.H. (2010). Statistical arbitrage in the U.S. equities market. *Quantitative Finance*, 10(7), 761–782.

[6] Said, S.E. and Dickey, D.A. (1984). Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, 71(3), 599–607.