

Time-Series Analysis of Stock Spread Dynamics: ARIMA–GARCH Modeling and Backtesting Evidence

Abstract

This project examines whether modeling the relationship between two economically related stocks provides additional insights beyond traditional single-asset time-series analysis. Using The Coca-Cola Company and PepsiCo as an empirical example, we construct a log-price spread and analyze its dynamics within a unified time-series framework. We first fit ARIMA models to capture linear dependence and mean-reverting behavior, and further improve the fit by adopting Student-t innovations to better accommodate heavy-tailed financial data. Diagnostic analysis shows that although the conditional mean is well modeled, the residuals exhibit clear volatility clustering, motivating the use of a GARCH(1,1) specification to capture time-varying conditional variance. Based on the estimated mean and volatility dynamics, we implement a simple spread-based trading strategy and evaluate its performance through backtesting. The results show that combining ARIMA and GARCH models provides a practical framework for translating statistical structure into trading signals and for understanding relative price dynamics in financial time series.

1 Introduction and Motivation

Most empirical time-series studies in finance focus on a *single asset*, such as modeling the return or volatility of one stock. However, this approach may overlook information contained in relationships between assets. Economically related stocks often move together, so analyzing their **relative behavior** can provide additional insights into co-movement, diversification, and risk. A common approach is **pairs trading**, which constructs a price spread between two assets and studies its dynamics. Prior work suggests that spread-based methods can reveal meaningful patterns, although these relationships may vary across time and market conditions (Zhu 2024; Lu et al. 2021; Ti et al. 2024).

Motivated by this literature, this project models **two stocks jointly** rather than separately by focusing on the spread between their log prices. Studying the spread shifts attention from absolute price movements to relative valuation dynamics, allowing us to examine whether deviations between related firms exhibit systematic mean and volatility behavior. We select **The Coca-Cola Company** and **PepsiCo** as our research pair. Both firms operate in the global non-alcoholic

beverage industry and share similar market exposures, which increases the likelihood of long-run co-movement, while differences in business structure and strategic positioning may generate temporary divergences. In addition, both stocks are large-cap, highly liquid, and have long trading histories, making them suitable for reliable estimation within a unified regression, ARIMA, and GARCH modeling framework.

1.1 Research Questions

Using two selected stocks (Stock A and Stock B), we construct a spread and ask:

1. Does the spread exhibit systematic time-series behavior beyond what we see from single-stock modeling?
2. Is the relationship between the two stocks stable over time?
3. Can the mean-reverting dynamics of the spread be effectively translated into a robust algorithmic trading strategy?

2 Method

2.1 Cointegration Regression Analysis

To implement a pairs trading strategy, we first establish the long-term equilibrium relationship between the two securities. We assume that the log-prices of Asset A (PEP) and Asset B (KO) are cointegrated. We perform an Ordinary Least Squares (OLS) regression:

$$\ln(P_{PEP,t}) = \beta \ln(P_{KO,t}) + \alpha + \varepsilon_t$$

Where: β is the hedge ratio (cointegration coefficient). α is the intercept (constant drift). ε_t represents the residuals, which we define as the **Spread**.

The regression results indicate a significant relationship between PEP and KO, whose coefficient equals to 1.26 with a p-value of 0.000. The price spread is shown in fig1.

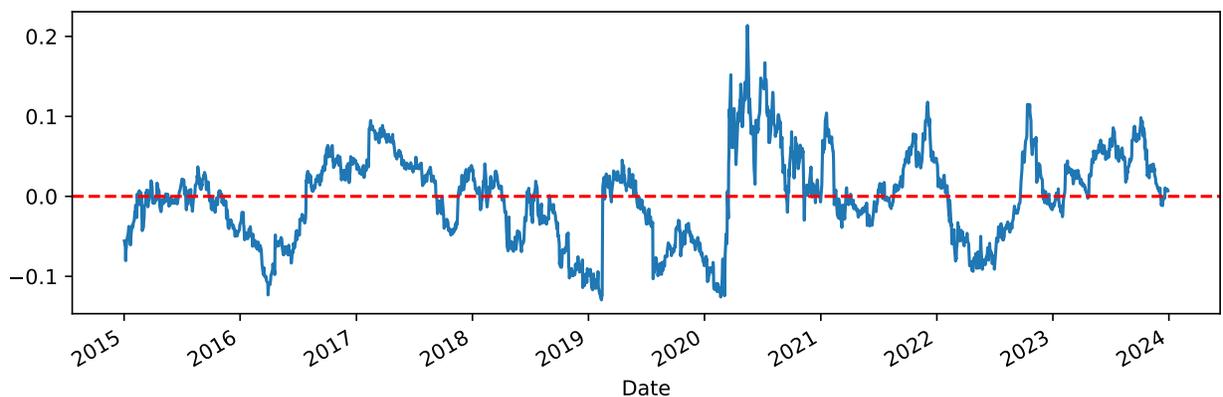


Figure 1: Log-Price Spread (Residuals)(2015-2023)

2.1.1 ACF OF OLS Model

Following the construction of the static spread via the Ordinary Least Squares (OLS) regression, diagnostic checks of the residuals revealed a slow, gradual decay in the Autocorrelation Function (ACF). This high persistence indicates a large serial correlation, showing that the spread possesses significant historical ‘memory’ rather than behaving as independent white noise. To account for temporal dependence, we model the spread using ARIMA. This removes autocorrelation and helps estimate the mean-reversion half-life for trading signals.

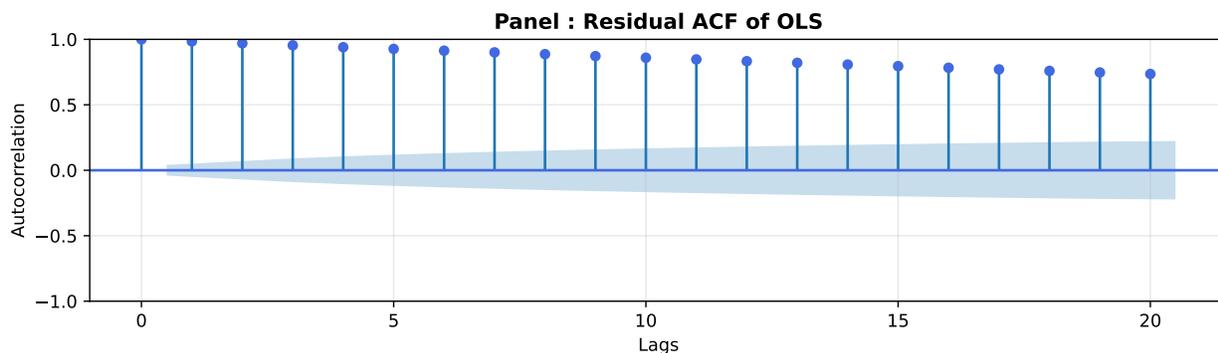


Figure 2: Log-Price Spread (Residuals)(2015-2023)

Following the stationarity confirmation from the ADF test, we proceed to determine the appropriate mean specification for the spread. To this end, we estimate a range of ARMA(p, q) models and compare their performance using the Akaike Information Criterion (AIC).

Among the candidate specifications, the AR(1) model yields the lowest AIC value. Introducing higher-order autoregressive or moving-average terms does not lead to further improvement in the information criterion. This suggests that the first-order autoregressive structure sufficiently captures the primary linear dependence in the spread dynamics.

Therefore, we adopt AR(1) as the mean specification for the subsequent analysis.

The estimation results of the AR(1) model indicate that the autoregressive coefficient is highly significant and close to unity. This suggests strong persistence in the spread dynamics, implying that current deviations are largely influenced by their previous values. Although the coefficient is close to one, it remains below unity, which is consistent with the earlier ADF result indicating stationarity.

The constant term is statistically insignificant and economically small, implying that the spread fluctuates around a stable long-run mean close to zero. This aligns with the theoretical expectation that the relative valuation between the two stocks does not exhibit a systematic drift over time.

Residual diagnostics further indicate that the AR(1) specification adequately captures the linear dependence structure in the spread. However, subsequent diagnostic checks reveal deviations from normality, particularly in the tails of the distribution, which motivates further refinement of the innovation distribution in the next section.

2.2 Model Improvement

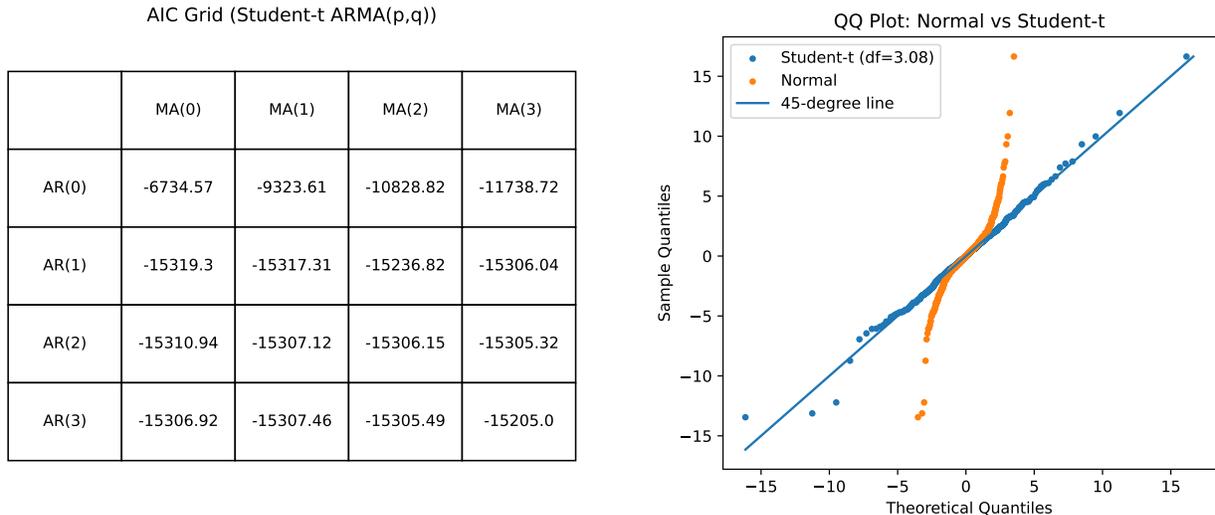


Figure 3: AIC Grid(Student-t ARMA(p,q)) & QQ Plot:Normal vs Student-t

While the AR(1) model adequately captures the linear dependence in the spread, residual diagnostics reveal significant deviations from normality. In particular, the QQ-plot under the Gaussian assumption indicates heavy-tailed behavior, suggesting that extreme deviations occur more frequently than predicted by a normal distribution.

To address this issue, we retain the same mean structure but relax the Gaussian innovation assumption by allowing the residuals to follow a Student’s t distribution. We re-estimate the model and compare fit using AIC.

The Student’s t specification yields a substantial improvement in fit: the AIC decreases from approximately -14648 under the Gaussian AR(1) model to about -15319 under the Student’s t specification. The estimated degrees of freedom is $\nu \approx 3.08$, indicating pronounced heavy tails. The QQ-plot comparison further confirms that Student’s t aligns more closely with the empirical residual quantiles, particularly in the tails, whereas the Gaussian assumption systematically understates extreme outcomes.

Finally, we recompute the AIC grid across ARMA(p,q) candidates under the Student’s t innovation assumption. The lowest AIC still occurs at $(p, q) = (1, 0)$, indicating that AR(1) remains the preferred mean specification even after relaxing the distributional assumption. Overall, the evidence suggests that the primary improvement comes from modeling heavy-tailed innovations rather than increasing ARMA order.

2.3 Stationarity Testing (ADF Test)

The success of a spread-based trading framework relies on the mean-reverting property of the residual series ε_t . To verify this property, we conduct the Augmented Dickey–Fuller (ADF) test

for unit roots. The null hypothesis states that the spread follows a random walk (non-stationary), while the alternative hypothesis implies stationarity.

The test produces an ADF statistic of -4.0764 with a p-value of 0.0011 , which is well below the 5% significance level. We therefore reject the null hypothesis and conclude that the spread is stationary. This justifies modeling the spread directly using an ARMA/ARIMA framework with $d = 0$.

2.4 Conditional Heteroskedasticity and GARCH Modeling

After fitting the ARIMA model, we examine whether the residual variance remains constant over time. Although ARIMA captures linear dependence in the conditional mean, financial spreads often display volatility clustering, where large shocks tend to be followed by large shocks.

To test for conditional heteroskedasticity, we apply the ARCH-LM test. The null hypothesis assumes constant variance (no ARCH effect). The test yields an LM statistic of about 2008.80 with a p-value close to zero, so we strongly reject the null and conclude that volatility clustering is present in the spread series.

2.4.1 GARCH(1,1) Model Specification

To model time-varying volatility, we fit a standard GARCH(1,1) model:

$$X_t = \mu + \varepsilon_t, \quad \varepsilon_t = \sigma_t z_t,$$

where z_t is an i.i.d. innovation with unit variance. The conditional variance follows

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2.$$

Here, $\omega > 0$ controls the long-run variance level, α measures the immediate shock impact (ARCH effect), and β captures volatility persistence (GARCH effect). The sum $\alpha + \beta$ determines how quickly volatility shocks decay; values close to one indicate highly persistent volatility, which is common in financial time series.

2.4.2 Estimated Conditional Volatility

The estimated conditional volatility from the fitted GARCH(1,1) model is shown below.

The volatility process exhibits clear clustering, with extended periods of high and low volatility. In particular, volatility spikes are observed during market stress periods (e.g., around 2020), indicating that spread risk increases substantially during turbulent market conditions.

Furthermore, the persistence measure $\alpha + \beta$ is close to one, implying that volatility shocks decay gradually rather than immediately. This confirms that the spread variance is highly persistent and time-varying, justifying the use of a GARCH framework instead of a constant-variance assumption.

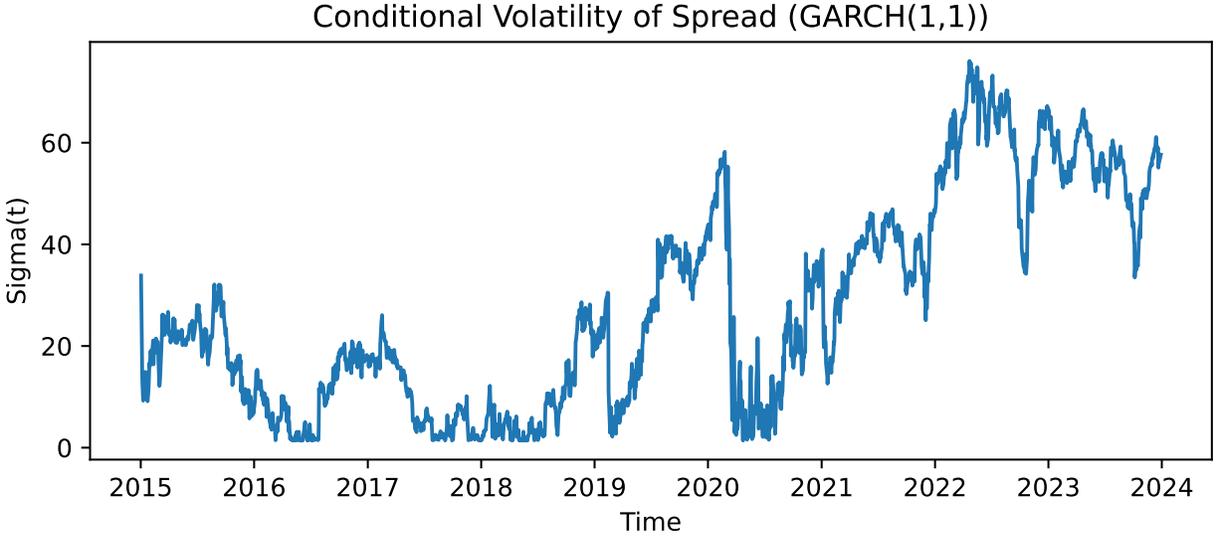


Figure 4: Conditional Volatility of the Spread from GARCH(1,1)

2.4.3 Diagnostic Check

As a diagnostic step, we apply the ARCH-LM test to the standardized residuals from the fitted GARCH model.

Applying the ARCH-LM test to the standardized residuals gives a p-value of nearly zero, showing that the heteroskedasticity is fully explained by the GARCH model.

Overall, the model substantially improves upon the constant-variance assumption and provides an adequate description of volatility behavior for the purposes of this project.

3 Backtesting and Strategy Implementation

3.1 Signal Modeling & Beta-Neutrality

We modeled the spread as a discrete Ornstein-Uhlenbeck process via an AR(1) regression. The estimated coefficient $\phi = 0.99 (p < 0.01)$ confirms strict stationarity and yields a half-life of $H = -\ln(2)/\ln(\phi) \approx 68.97$ days. Discarding heuristic lookback periods, we strictly set our rolling window to $w = 69$ days to compute the dynamic signal: $Z_t = \frac{\text{Spread}_t - \mu_{t,w}}{\sigma_{t,w}}$

3.2 State-Machine Trading Rules

Based on our residual diagnostics and half-life estimates, we implemented the following bounded rules:

- **Entry:** Go short the spread (short 1 unit PEP, long β units KO) when $Z_t > 2.0$. Go long the spread (long 1 unit PEP, short β units KO) when $Z_t < -2.0$.
- **Take-Profit:** Close positions when Z_t reverts to 0.

- **Stop-Loss (Tail & Time):** Liquidate unconditionally if extreme tail risks trigger $|Z_t| \geq 3.5$, or if the holding period exceeds $2H \approx 89$ days, indicating a potential cointegration breakdown.

3.3 Out-of-Sample Performance

Evaluated out-of-sample from 2024 to present, the pairs strategy delivered a robust cumulative return of 36.40%, trailing the S&P 500's 48.40% amid a strong bull market. However, this performance perfectly aligns with its statistical arbitrage mandate. Crucially, the daily return correlation with the broader market is remarkably low at $\rho = 0.069$. This near-zero correlation validates our ρ -weighted allocation, proving the strategy successfully decoupled from systematic market risk to generate pure, uncorrelated Alpha.

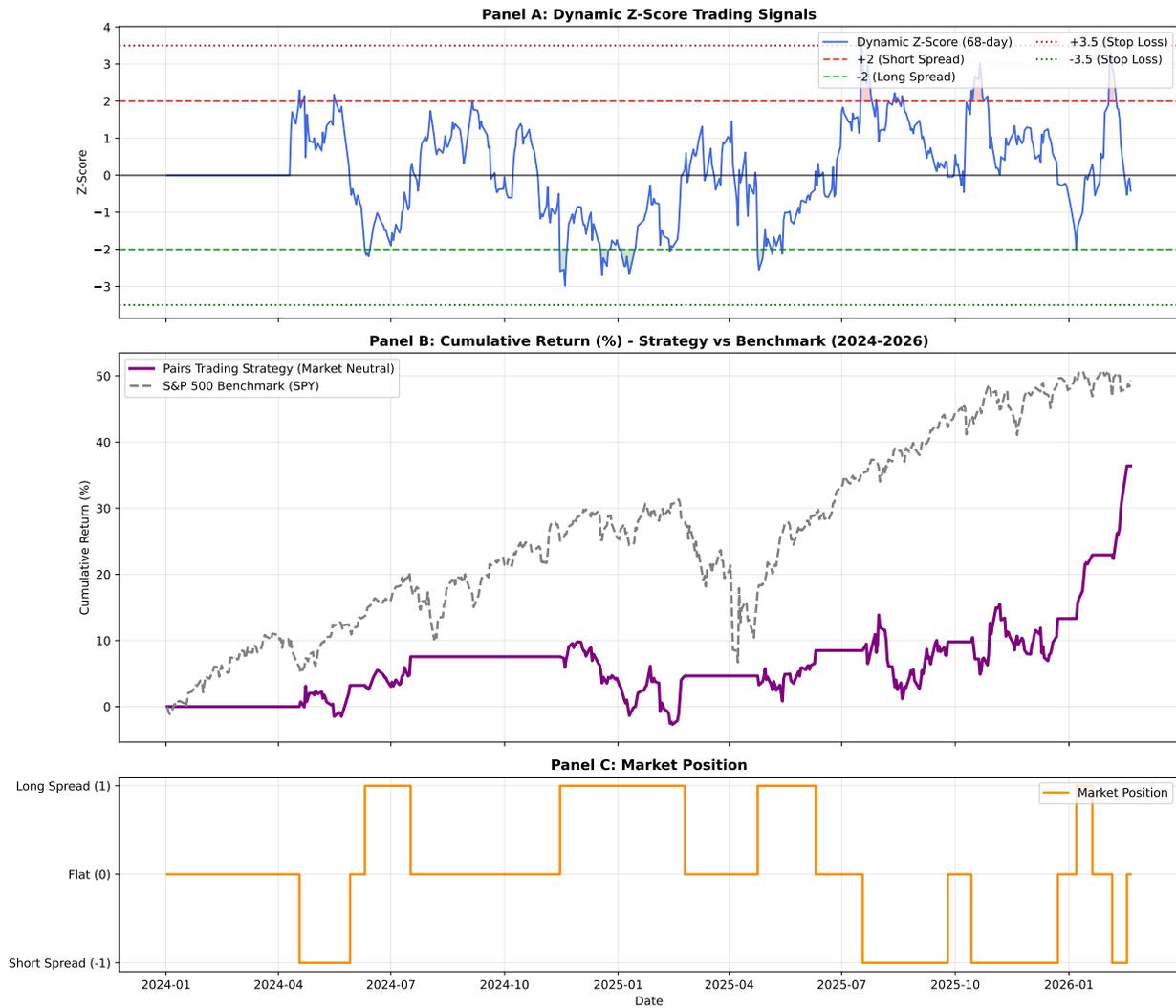


Figure 5: Out-of-Sample Strategy Diagnostics (2024-Present).

4 Conclusion

This project examines the dynamics of a stock spread to evaluate whether a two-asset framework provides richer insights than traditional single-stock modeling. Using Coca-Cola and PepsiCo as a representative pair, we analyze the spread through a sequential ARIMA–GARCH approach, allowing us to study both the conditional mean and conditional volatility.

The ARIMA model captures the linear dependence and mean-reverting behavior of the spread, suggesting that relative price movements contain structured time-series information. However, residual diagnostics reveal significant conditional heteroskedasticity, indicating that volatility changes over time. Incorporating a GARCH(1,1) model addresses this issue by capturing volatility clustering and persistence, producing a more realistic description of the spread dynamics.

Importantly, the analysis is not purely statistical. Based on the estimated mean-reversion and volatility structure, we construct a simple trading strategy and evaluate it through backtesting. The backtest demonstrates the practical value of modeling spread dynamics: statistical signals derived from ARIMA and GARCH models can be translated into economically meaningful trading rules. While the strategy remains simplified, the results highlight how time-series modeling can bridge empirical analysis and real-world decision-making. Future work may extend this framework by considering adaptive spread construction, structural breaks, or more advanced risk-adjusted trading rules.

5 Acknowledgements

We acknowledge the use of AI-assisted tools for language polishing and debugging support in this project. The authors take full responsibility for the model design, empirical analysis, and interpretation of results.

References

- Lu, Jun-Yan et al. 2021. “Structural Break-Aware Pairs Trading Strategy Using Deep Learning.” *The Journal of Supercomputing*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC8369334/>.
- Ti, Yi-Wen et al. 2024. “Improving Cointegration-Based Pairs Trading Strategy with Adaptive Thresholds.” *Computational Economics*. <https://doi.org/10.1007/s10614-023-10539-4>.
- Zhu, Xuanchi. 2024. “Examining Pairs Trading Profitability.” Yale University, Department of Economics. https://economics.yale.edu/sites/default/files/2024-05/Zhu_Pairs_Trading.pdf.