

Performance of Pairs Trading Strategies with Machine Learning and non-Machine Learning Methods: Comprehensive Empirical Evidence from the Chinese Stock Market

Introduction and Motivation

The Chinese stock market, characterized by its rapid development and significant volatility, presents a unique environment for exploring trading strategies. Among these, pairs trading—a market-neutral strategy that capitalizes on relative price movements between two correlated assets—has garnered increasing attention from both academics and practitioners. This study focuses on evaluating the profitability of pairs trading strategies in the context of the Chinese stock market, with a particular emphasis on different market conditions.

The CSI 300 Index, as a benchmark of the Chinese A-share market, has exhibited distinct phases of sharp price fluctuations over the years. As depicted in Figure 1, three prominent periods of price surges followed by sharp declines are evident: the pre-2008 global financial crisis phase, the 2015 stock market bubble, and the 2020 post-COVID recovery. These periods, as shown in Figure 1, marked by heightened market volatility and investor sentiment, provide a fertile ground for assessing the performance of pairs trading strategies under extreme market conditions.

In contrast, the "normal" market phases—characterized by relatively stable price movements—offer an opportunity to benchmark the strategies' profitability in less volatile environments. By examining both extreme and normal phases, this study aims to uncover whether pairs trading strategies exhibit consistent performance across varying market conditions.

Research Objectives

1. **Market Phase Analysis:** To analyze the profitability of pairs trading strategies during the three significant market phases (pre-2008, 2015 bubble, and post-2020 recovery) compared to normal phases.
2. **Machine Learning Integration:** To assess whether machine learning-based pairs trading strategies outperform traditional non-machine learning methods in the Chinese stock market.
3. **Market Neutrality and Risk:** To evaluate the risk-adjusted returns of these strategies across different market conditions, ensuring their effectiveness as market-neutral approaches.

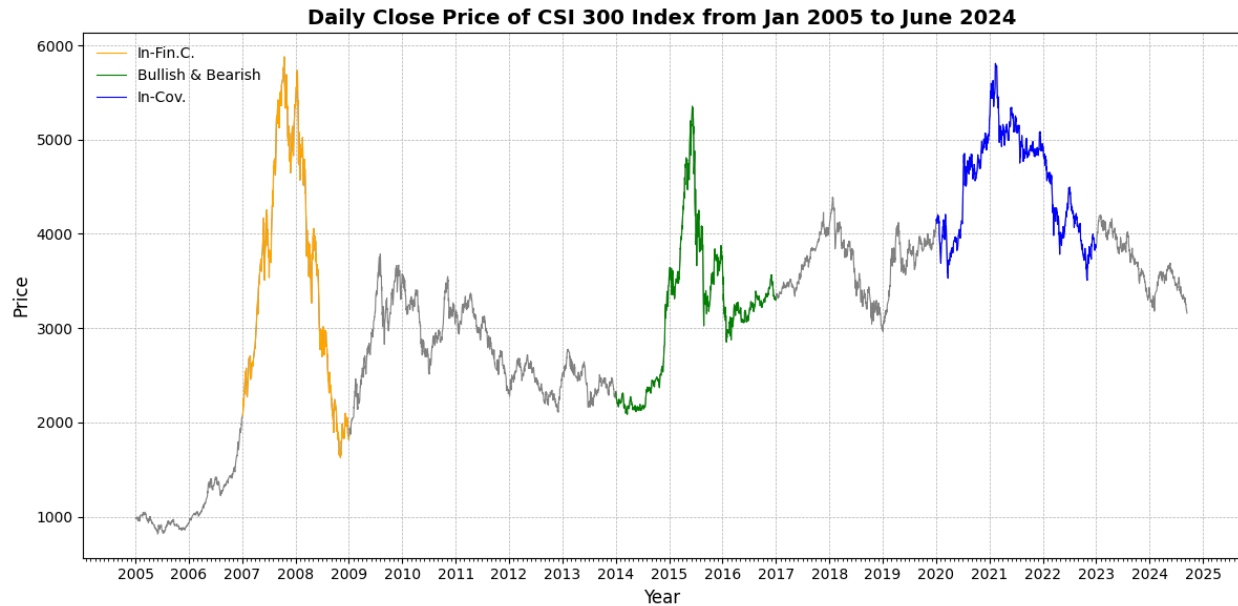


Figure 1. Daily close price of the CSI 300 index.

This figure shows the daily close price trend of the CSI 300 Index from January 2005 to June 2024, highlighting three key periods: the Financial Crisis (January 2007-December 2008), the Bullish and Bearish Market Phase (January 2014-December 2016), and the COVID-19 Pandemic Period (January 2020-December 2022).

Methodology Framework

The research methods can be divided into two parts, the first part is non-machine learning methods, and the second part is machine learning methods.

First, I will introduce the methods of the non-machine learning part, combined with my current work progress.

This part of the work began with a literature study ([A survey of statistical arbitrage pairs trading strategies with non-machine learning methods, 2016-2023](#)). This literature review summarizes statistical methods for pairs trading. See my paper for details.

My second paper used the distance method from non-machine learning methods, which is also the most basic method ([Does Pairs Trading Still Work Evidence from Chinese Stock Market](#)).

Next, my third paper used the Copula method from non-machine learning methods ([Performance of Pairs Trading Strategies Based on Various Copula Methods](#)).

Next, my fourth paper used the Renko and Kagi chart method from non-machine learning methods ([Performance of Pairs Trading Strategies Based on Renko and Kagi Charts](#)).

Second, I will introduce the methods of the machine learning part, combined with my current work progress.

For this part of the research, I first wrote the fifth paper, which is a literature review summarizing the machine learning methods used in pairs trading ([A survey of statistical arbitrage pair trading with machine learning, deep learning, and reinforcement learning methods](#)).

Next, I will complete the sixth and final paper, and I will choose the PCA method (or other methods).

Framework Diagram

