Design and Analysis of an Innovative Arbitrage Strategy: Bridging Stock Index Futures and Cross-border ETFs

Longxi Liang [1st]

School of Automation, Guangdong University of Technology, Guangzhou, 510006, China, lianglongxi@me.com

Shaoshan Guo [1st]

School of Automation, Guangdong University of Technology, Guangzhou, 510006, China, 2055829622@qq.com

Zhuoming Xie *

School of Integrated Circuits, Guangdong University of Technology, Guangzhou, 510006, China, 280645220@qq.com

Abstract

This study explores cross-market arbitrage opportunities, focusing on stock index futures and cross-border ETFs, amid increasing global capital liquidity and market interconnectivity. By developing a mathematical model that incorporates "inter-asset ratio coefficient" and "market volatility index," it introduces innovative arbitrage strategies. These strategies account for price discrepancies, volatility, and correlations between these financial instruments, determining trading strategy parameters through price correlation, regression analysis, and volume assessment. Verification through simulation tests on data from the China Financial Futures Exchange and related cross-border ETFs in the U.S. and Hong Kong markets identifies ten optimal portfolios. Utilizing Sharpe Ratio and Markowitz Mean-Variance model, the study optimizes return and risk control across larger financial portfolios. Empirical results affirm the proposed algorithm's substantial arbitrage potential and profitability, offering a practical approach to cross-market arbitrage.

CCS CONCEPTS

Applied computing ~ Law, social and behavioral sciences ~ Economics

Keywords

Stock Index Futures and Cross-Border ETFs, Cross Market Arbitrage, Trading Strategies and Mathematical Models

1 INTRODUCTION

Research on cross-market arbitrage reveals that interconnectivity among international financial markets plays an indispensable role in promoting the internationalization of securities markets. With the increasing quota under the Qualified Domestic Investor (QDII) and Qualified Foreign Investor (QFII) systems, domestic and foreign capital liquidity has been significantly enhanced, thus promoting the strengthening of market connectivity. As an important component of the financial futures market, stock index futures use stock indices as their underlying for trading. Given the homogeneity or deep connection between the underlying assets of stock index futures and cross-border exchange-traded funds (ETFs) in different markets, the differences and fluctuations between their market prices provide investors with cross-market arbitrage opportunities. In order to explore the feasibility of cross-market arbitrage between stock

[1st] Statement: Longxi Liang and Shaoshan Guo jointly made equal contributions to this research and jointly assume the responsibilities of the first author.

index futures and cross-border ETFs, this paper will propose a set of specific arbitrage strategies by building a mathematical model.

Several studies have significantly advanced the understanding of stock index futures and cross-border ETF trading methodologies. He Nie et al. uncover both linear and two-way nonlinear long-term causality between US stock market volatility and China's ETF market [1]. Zhao et al. demonstrate that ETF holdings enhance the pricing efficiency of the A-share market, with efficiency increasing post-ETF entry and decreasing upon exit [2]. Barbon and Gianinazzi identify a unitary price elasticity in the long-term stock demand curve influenced by index-linked ETFs [3]. Harada and Okimoto's analysis reveal improved stock performance within the Nikkei 225 following the Bank of Japan's ETF purchase program [4]. Hurlin et al. find higher counterparty risk exposure in swap funds compared to physical funds, though with risk compensation for investors [5]. These findings provide critical insights and data for research of this paper.

This paper firstly introduces the concept of "inter-asset scaling factor" to ensure the consistency of the underlying assets and the number of indicators in the cross-market trading of stock index futures and cross-border ETFs. Subsequently, by defining the "market volatility index", the regression prediction of the price difference between inter-market stock index futures and cross-border ETFs is conducted through a time series prediction network. Next, price correlation, regression slope, and volume ratio are evaluated to determine the volume impact factor in the trading strategy. Meanwhile, this paper proposes an improved trading strategy for cross-market trading and a DBP position covering strategy. Combining the above proposed and introduced methods, this paper completely constructs a cross-market trading arbitrage algorithm for stock index futures and cross-border ETFs.

In addition, in order to verify the reliability and application value of the cross-market trading arbitrage algorithm proposed in this paper, this paper utilizes the algorithm to simulate test and analyze the stock index futures data of the four Chinese financial futures exchanges and the corresponding or related cross-border ETF data in the U.S. and Hong Kong markets. At the application value level, this paper simulates and selects 10 portfolios of stock index futures and cross-border ETFs that maximize portfolio returns and combine them into a larger financial portfolio. These 10 portfolios are weighted proportionally in the larger portfolio by Sharpe's ratio and Markowitz mean-variance model to maximize portfolio returns while controlling risks. Regarding the analysis and validation of the reliability of the algorithm, this paper empirically analyzes the 10 portfolios of stock index futures and cross-border ETFs using the trading data of 7 trading days in November 2023 to verify the arbitrage potential and profitability of the proposed algorithm in the real trading environment.

2 METHODS

In this paper, in terms of research methodology, a time series forecasting network is used and a "market volatility index" is introduced to forecast the price fluctuations of stock index futures and cross-border ETFs. The inter-asset scaling factor ensures the consistency between stock index futures and cross-border ETFs in terms of the number of underlying assets or indicators. The paper deeply analyzes the price correlation and covariance between different markets as well as the volume ratio and regression slope to decide the values of trading decision parameters. In addition, this paper proposes a complete algorithm of cross-market trading arbitrage strategy, applied with DBP strategies for volatile markets.

In order to simplify the model, this paper performs a two-by-two combination of stock index futures X and cross-border ETF Y, and processes the time-series data of these combinations to obtain the uniform market price P and trading volume Q for each trading time point t. The market price is calculated by taking the arithmetic average of the opening price O, the closing price C, the highest price H, and the lowest price L of the stock index futures and cross-border ETFs for each trading point in time to arrive at the market price of the stock index futures and cross-border ETFs, as described in the formula below:

$$P_{x,y}(t) = O_{x,y}(t) + C_{x,y}(t) + H_{x,y}(t) + L_{x,y}(t)/4 \quad (1)$$

In this equation, the subscript x represents the variable corresponding to stock index futures and y represents the variable corresponding to cross-border ETFs. For the trading volume, this paper uses the trading volume data of each trading time point t, which is denoted as the trading volume of stock index futures $Q_x(t)$ and the trading volume of cross-border ETFs $Q_y(t)$, respectively. This study aims to reduce data noise and extract stable trends through this process, and ensure that all factor values are within the same numerical range through scaling to facilitate the stability and accuracy of trading decisions.

2.1 Determination of inter-asset scaling factors, market volatility indices

This paper proposes to introduce the concepts of "inter-asset proportionality coefficient" and market volatility index δ_{mkt} into the model, and define the inter-asset proportionality coefficient in terms of the starting point of the time series $t = 0$:

$$\beta = P_{futures}/P_{ETF} \quad , t = 0 \quad (2)$$

In this equation, $P_{futures}$ and P_{ETF} represent the unit prices of domestic stock index futures and offshore cross-border ETFs respectively. This coefficient is designed to ensure that the underlying assets of domestic stock index futures and overseas cross-border ETFs are the same in trading by adjusting the trading volume, so as to ensure the logical consistency of the arbitrage strategy and the effectiveness of the cross-market hedging strategy in actual operation.

Further, this paper defines the market volatility index as:

$$\delta_{mkt}(t+1) = P_x(t) - \left(P_y(t)/\beta\right) \quad (3)$$

Where $P_x(t)$ and $P_y(t)$ are the market prices of stock index futures and cross-border ETFs at time point t, respectively. The index provides a basis for identifying arbitrage opportunities by measuring the price difference between the foreign market and the domestic market, which helps to maximize arbitrage gains while controlling risks.

2.2 Predicting future market volatility indices using LSTMs

In order to accurately predict indices of market volatility in the future, this paper proposes a prediction model based on Long Short-Term Memory Network (LSTM). LSTM, as a kind of recurrent neural network that can efficiently process time series data, solves the gradient vanishing problem faced by traditional RNNs through the information transfer mechanism of its internal states [8], enabling it to capture the long-term dependencies in historical data, and then predict future market prices [9].

The model proposed in this paper adopts a simplified sequential structure, which mainly consists of an LSTM layer containing 50 LSTM units for learning patterns in time-series data, a 20-neuron fully-connected layer for feature mapping, and a 3-neuron output layer for predicting the prices of the next three time points. The total number of model parameters is 11,483, all of which are trainable. The input data is a sequence of prices at six consecutive time points, and the output is the price prediction at three future time points.

In practice, a series of consecutive market volatility indices $\delta_{mkt}(n), n = t-6, t-5, \ldots, t$ totaling 7 data points are used as inputs and outputs the predicted values of $\delta_{mkt}(n), n = t+1, \ldots, t+3$ for three future time points. For value corrections, the exponential decay weighted average method is used, which is calculated by the following formula:

$$\begin{cases} \delta'_{mkt}(t+n) = \sum_{n=1}^{3} w_n \delta_{mkt}(t+n) / \sum_{n=1}^{3} w_n \\ w_n = e^{-\lambda n} \end{cases} \quad (4)$$

In this equation, w_n is the weight, λ is the decay rate parameter, $\delta_{mkt}(t+n)$ is the value at the moment t+1 to t+3 of the LSTM network output, $\delta'_{mkt}(t+n)$ is the value at the moment t+1 of the final prediction after correction, in order to reflect the proximate causation effect of the time series, and at the same time to balance the weight of the near-term and far-term prediction, to improve the prediction accuracy and robustness.

2.3 Volume impact factors

This paper introduces a series of metrics to quantify the trading relationship between stock index futures and cross-border ETFs, with a focus on price correlation c_1, regression slopes c_5, volume ratios c_6 and their significance assessment. The core mathematical expressions in this subsection are shown below:

$$c_1 = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) / \sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (5)$$

$$c_5 = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) / \sum_{i=1}^{n}(x_i - \bar{x})^2 \quad (6)$$

$$c_6 = \sum_{i=1}^{n} Q_{y(t)} / \sum_{i=1}^{n} Q_{x(t)} \quad (7)$$

Where equation (5) is the formula for price correlation c_1, equation (6) is the formula for regression slope c_5, and equation (7) is the formula for volume ratio c_6. x_i and y_i represent the prices of stock index futures and cross-border ETFs at the i-th point in time, respectively, and x and y represent the average price of the corresponding time series, respectively. This indicator reflects the synchronization of the price movements of the two-time series, and the closer its value is to 1, the more synchronized the price movements of the two markets are, which provides a high confidence level for the price correlation-based arbitrage strategy.

In addition, the significance of the cointegration relationship c_3 and its significance c_4, as well as the significance of the price correlation c_2, provide statistical significance measures of the synchronization of price movements and the long-term relationship between the two markets, as assessed by the Engle-Granger test [6] and the probability value of the correlation. Finally, calculate the arithmetic mean of the above indicators c_1, c_2, c_3, c_4, c_5, and c_6 after taking their absolute values to obtain the multiplication factor used for the trading strategy, f. The quantitative analytical framework constituted by these indicators utilizes the linkage between cross-border ETFs and the international market and its impact on the Chinese A-share market to provide a scientific and practical quantitative basis for the trading strategy, which is aimed at capturing market volatility and liquidity and accurately reflect inter-market correlations.

2.4 Improved trading strategies for cross-market trading

In this paper, an improved trading strategy is proposed for cross-border market trading arbitrage strategy with adaptive settings in three aspects: hyperparameter updating, trading strategy updating and position covering strategy updating. In order to simulate the actual trading environment, a series of hyper-parameters are introduced to reflect the specific trading costs, tax rates, market shock costs and liquidity constraints of the domestic and foreign markets considered in the trading strategy, taking into account the trading costs and liquidity constraints of the domestic and foreign markets, which ensures the practical applicability of the model and compliance with the current financial market regulations.

Table 1. Explanation of the hyperparameters.

Hyperparameter Symbol	Numerical Value	Description
$eff1\_dom$	0	Fixed costs in domestic markets, assuming no fixed transaction costs
$eff2\_dom$	0.1%	Domestic market stamp duty rates
$eff3\_dom$	0.2%	Domestic market trading commissions

Hyperparameter Symbol	Numerical Value	Description
$eff4\_dom$	0.05%	Domestic market shocks costs
$trade\_regulation\_dom$	70%	Indicators of liquidity constraints in domestic markets
$eff1\_abr$	CNY ¥0.72	Fixed cost in foreign markets, converted at exchange rates in October, 2023
$eff2\_abr$	0.1%	Offshore market ETF management fee
$eff3\_abr$	0.1%	Commission on foreign market transactions
$eff4\_abr$	0.02%	Cost of offshore market spreads
$trade\_regulation\_abr$	70%	Indicators of liquidity constraints in foreign markets
δ_{mkt-th}	1e-5	Market volatility index thresholds, for guiding trading decision points

In the trading judgment strategy of this paper, the determination of trading volume follows the formula:

$$Q_{\{buy,sell\},\{1,2\}} = f \cdot d_{\{1,2,3,4\}} \cdot \ln|\delta_{mkt}(t)| + 1 \qquad (8)$$

Where, when cross-border ETFs are at a discount to domestic stock index futures (i.e., $\delta_{mkt}(t) > |\delta_{mkt-th}|$), the buying volume $Q_{buy,1}$ of stock index futures in the domestic market and the selling volume $Q_{sell,2}$ of cross-border ETFs in the international market are calculated in accordance with this formula; Conversely, when there is a premium (i.e., $\delta_{mkt}(t) < -|\delta_{mkt-th}|$), the buying volume $Q_{buy,2}$ of cross-border ETFs in the international market and the selling volume $Q_{sell,1}$ of stock index futures in the domestic market are determined in accordance with the same formula. In this modeling assumption, the constraints on the settlement rules for domestic T+1 trading settlement rule constraints are ignored, as the model aims to capture and utilize cross-market instantaneous price differences for arbitrage, which is usually done on the same trading day, and thus is not substantially affected by the T+1 rule.

For position covering sessions, this paper proposes Dynamic Barrier Positioning (DBP) strategy to optimize the position size in cross-market arbitrage. When the arbitrage indicator $\delta_{mkt}(t)$ crosses the $|\delta_{mkt-th}|$ interval, i.e., changes from the arbitrage interval to the no-arbitrage interval, the position $Q_{adjust,t}$ is adjusted according to the following formula:

$$Q_{adjust,t} = \Theta^* - \Theta_t \qquad (9)$$

Where Θ^* is the target position size and Θ_t is the current position size. This strategy allows immediate adjustment of position size to cope with market fluctuations, reduce position errors, improve arbitrage efficiency, and achieve position optimization.

This paper synthesizes the methods mentioned above, then obtains the following algorithm for cross-market trading arbitrage strategy:

ALGORITHM 1: Cross-Market Trading Arbitrage Strategy

Input: Time series data $x(t)$, $y(t)$, current moment t

Output: Yield curve, total return, maximum retracement data

Set t as the current time, δ_{mkt-th} as the market volatility threshold;

while the final trading time (e.g. closing moment) has not been reached, do

  Calculate the market volatility indices $\delta_{mkt}(t-6)$ to $\delta_{mkt}(t)$;

  Predict $\delta_{mkt}(t+1)$ using sequential regression prediction model based on LSTM network;

  if $\delta_{mkt}(t+1) < -\delta_{mkt-th}$, then

    Execute the arbitrage discount strategy with a trading volume of $Q_{buy,1} + Q_{sell,2}$;

  else if $-\delta_{mkt-th} \le \delta_{mkt}(t+1) \le +\delta_{mkt-th}$, then

    Execute the DBP position covering strategy;

  else

Execute the arbitrage premium strategy with a trading volume of $Q_{buy,2} + Q_{sell,1}$;

Execute a cost-cutting session, including subtracting fixed costs, proportional costs, and enforcing trading ratio limits;

if the final trading time (e.g. closing moment) has been reached, then

Return the yield curve, total return, maximum retracement data, etc., conduct further analysis;

Terminate the while loop;

else

if current position $\Theta > 0$, then

Execute close;

else

Execute cover;

Update time $t = t + 1$, update position Θ, update capitalization, calculate return, update retracement, then go back to the beginning of the loop;

end

3 RESULTS AND ANALYSIS

3.1 Analysis of optimal combination selection of cross-border ETFs and stock index futures

This paper analyzes the historical data and statistical indicators of stock index futures and cross-border ETFs of the four indexes listed on the CICC, namely, SSE 50, CSI 300, CSI 500, and CSI 1000, and considers factors such as price difference, trading time, trading volume, trading cost, and liquidity of funds, and synchronously collects the actual trading data and plots the yield curves, and then selects 10 most suitable cross-border ETF and stock index futures portfolios. On this basis, the Sharpe ratio and Markowitz mean-variance model [7] are used to simplify the implementation and calculate the weighting ratio of the portfolios, aiming to maximize the portfolio returns while controlling the risks. The input data include the portfolio's relative return time series, cumulative trading profit and maximum retracement, and the output is the investment weight, overall risk and expected return, the curve of which is shown in Figure 1, which provides data support for investment decisions.

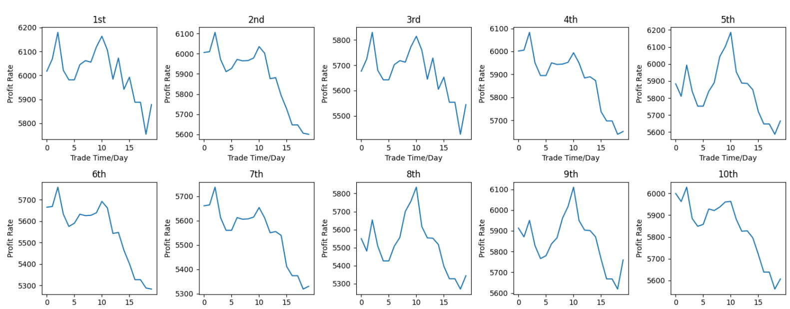

Figure 1. Yield Curve of 10 Candidate Portfolios

3.2 Empirical analysis of cross-border ETF and stock index futures trading strategies

This paper empirically analyzes the trading data for the period from November 9 to November 17, 2023, which includes stock index futures data from the Chinese mainland market on the RoyalFlush iFinD platform, ETF trading data from the Hong Kong and U.S. markets, as well as the selected cross-border ETFs and stock index futures portfolios as mentioned before. By simulating the algorithm proposed in this paper in a high-frequency trading environment, the data of 206,966 times in 7 trading days are subdivided and preprocessed in time steps to ensure the consistency and comparability of the data.

The simulation trading sets the initial investment at 1.5 million yuan, of which 1 million yuan in futures and 500,000 yuan in spot, with a margin ratio of 10% for futures, and the transaction fee rate set at 0.005% for futures and 0.01% for stocks, respectively, and considering a price slippage of 0.01%. Through simulated trading, the selected 10 cross-border ETF trading portfolios are calculated and analyzed for a number of evaluation indexes, such as unit time step strategy return, unit time relative profit index, total cross-border profit, maximum retracement, yield curve, and trading data prediction, aiming at verifying the application effect of this paper's algorithms and potential profitability in the actual trading environment, as shown in Table 2 and Figure 2. This empirical analysis not only demonstrates the arbitrage potential of cross-border ETF and stock index futures portfolios, but also provides a practical reference for future trading strategies.

Table 2. Weighting ratios, portfolio risk and portfolio return for 10 candidate portfolios.

Ranking	Transaction Number	Investment Weight	Portfolio Risk	Portfolio Returns	Mean Returns Per Unit Time Step Strategy	Relative Profit Index Per Time Step
1	IM2311.CFE, 09151.HK	0.101	0.097	CNY ¥1,074,792.12	84.56702	61.5021
2	IM2311.CFE, 09173.HK	0.089	0.107	CNY ¥1,016,364.85	81.23501	61.50148
3	IC2311.CFE, 09151.HK	0.158	0.098	CNY ¥997,335.58	75.34784	55.89536
4	IM2311.CFE, 09031.HK	0.099	0.100	CNY ¥984,742.14	79.14683	61.50123
5	IM2311.CFE, 09812.HK	0.073	0.119	CNY ¥951,633.98	76.50042	61.4957
6	IC2311.CFE, 09173.HK	0.079	0.108	CNY ¥942,565.92	72.32218	55.89481
7	IC2311.CFE, 09031.HK	0.107	0.101	CNY ¥912,887.15	70.42639	55.89457
8	IC2311.CFE, 09812.HK	0.093	0.120	CNY ¥881,839.49	68.02418	55.88955
9	IM2311.CFE, 09801.HK	0.107	0.106	CNY ¥808,703.72	65.36713	61.49888
10	IM2311.CFE, 09839.HK	0.094	0.104	CNY ¥783,098.68	63.63108	61.50145
Total profit value of the total portfolio of cross-border transactions					CNY ¥13,519,150.29	

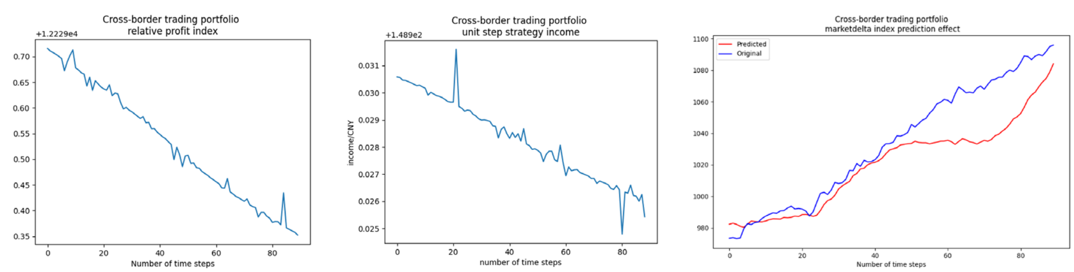

Figure 2. Predictive Effects of Cross-Border Trading Relative Profit Index (Left Panel), Cross-Border Trading Unit Time Strategy Returns (Center Panel), and δ_{mkt} Index (Right Panel).

The graphs analyzed in this paper show the trends of cross-border trading portfolios in terms of relative profitability indices and strategy returns per unit time step, revealing a downward trend in both with increasing time steps. This downward trend implies that the profitability of the trading portfolios decreases over time and the profit margins weaken, especially the decline in the unit time step strategy returns further validates the phenomenon of diminishing returns realized by the strategies over time. Meanwhile, the predicted versus actual curve of the cumulative cross-border trading portfolio index rises with increasing time step, and the blue line (actual value) and the red line (predicted value) have basically the same trend, but the increase of the actual value exceeds that of the predicted value, which suggests that the prediction model, although able to capture the market trend to a certain extent, still needs to be optimized to improve the prediction accuracy. The results of this analysis not only reflect the profitability of the cross-border trading strategy over time, but also point to the potential for improvement in the accuracy of the forecasting model.

4 CONCLUSION

This study successfully proposes and validates an innovative arbitrage strategy by thoroughly analyzing cross-market arbitrage opportunities in international financial markets, especially the trading potential between stock index futures and cross-border ETFs. The strategy is based on an accurate mathematical model that integrates key factors such as asset ratio coefficients, market volatility indices and price correlation, providing a new perspective and methodology for cross-market arbitrage. Through simulation tests and empirical analysis of relevant data from the China Financial Futures Exchange and the U.S. and Hong Kong markets, this study not only proves the effectiveness of the proposed strategy, but also demonstrates its profitability and risk control advantages in actual operation. The contribution of this study is that it provides a scientific and systematic methodology for arbitrage trading in financial markets, enriches the theoretical basis of cross-market arbitrage, and provides financial practitioners with strategic tools with practical application value, which has an important impact on promoting the efficiency of financial markets and the internationalization process.

REFERENCES

[1] He Nie, Yonghong Jiang and Baoqing Yang. **2018**. Do different time horizons in the volatility of the US stock market significantly affect the China ETF market? Applied Economics Letters, 25(12), 830-834. https://doi.org/10.1080/13504851.2017.1363853.

[2] Xiaokang Zhao, Guanghe Ran, Bing Shen and Xiaolong Li. **2021**. Do ETFs Improve the pricing efficiency of the A-share market—examining ETF holdings of individual stocks. Applied Economics, 53(18), 2090-2105. https://doi.org/10.1080/00036846.2021.1897075.

[3] Andrea Barbon and Virginia Gianinazzi. **2019**. Quantitative Easing and Equity Prices: Evidence from the ETF Program of the Bank of Japan. The Review of Financial Studies, 32(11), 4222-4265. https://doi.org/10.2139/ssrn.2925198.

[4] Kimie Harada and Tatsuyoshi Okimoto. **2021**. The BOJ's ETF purchases and its effects on Nikkei 225 stocks. International Review of Financial Analysis, 101826. https://doi.org/10.1016/J.IRFA.2021.101826.

[5] Christophe Hurlin, Grégoire Iseli, Christophe Pérignon and Stanley Yeung. **2019**. The Counterparty Risk Exposure of ETF Investors. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2462747.

[6] Jie Yan. **2007**. Research on ETF Funds and Arbitrage. Doctoral dissertation, Southwest University of Finance and Economics.

[7] Shenhao Tu, **2019**. Research on Investment Portfolio Based on Deep Reinforcement Learning and Improved Mean-Variance Model. Doctoral dissertation, University of Chinese Academy of Sciences (School of Sciences Institute for Artificial Intelligence).

[8] Haoran Zhang. **2022**. Research on the Prediction of Treasury Bond Yield Curve and Investment Portfolio Strategy Based on LSTM Method. Doctoral dissertation, Southwest University of Finance and Economics.

[9] Qijun Tan. **2021**. Research on Financial Time Series Prediction Based on Multiscale Decomposition and LSTM Network. Doctoral dissertation, South China University of Technology.