

CHAPTER 5

DISCUSSION

5.1 Summary

Data preprocessing: First, the researchers merged the three data sets into a complete data set, which contains information such as time series, stock prices, turnover rates, sentiment indicators, China's GDP, and company net profit margins. Subsequently, the data was cleaned, missing values were processed, and unnecessary variables were deleted to lay the foundation for subsequent analysis.

VAR model:

- a) Stationary test: The researchers used the ADF unit root test to test the stationary nature of the variables and found that the original sequences of stock prices, net profit margins and GDP were non-stationary, while their first-order difference sequences were stationary.
- b) Determination of lag order: The lag order of the VAR model was determined to be 1 through the AIC, SC and HQ criteria.
- c) Cointegration test: The Johansen cointegration test shows that there is a long-term equilibrium relationship between the variables, and a VAR model can be established.
- d) Model stability test: The AR characteristic root test shows that the model is stable.
- e) Granger causality test: The Granger causality test shows that China's GDP is the Granger cause of stock prices, while turnover rate and net profit margin are not.

- f) Pulse response analysis: The pulse response analysis shows that China's GDP has the greatest impact on stock prices, followed by net profit margins, and turnover rate has the least impact.
- g) Variance decomposition analysis: Variance decomposition analysis shows that in the short term, stock price fluctuations are mainly affected by their own changes, while in the long term, China's GDP has the greatest impact.

ARIMA model:

- a) Data transformation: The turnover rate is converted to logarithmic form and adjusted according to its impact on China's GDP, net profit margin and stock price.
- b) Model parameter determination: The parameters (p, d, q) of the ARIMA model are determined through ACF and PACF plots.
- c) Model training and prediction: The ARIMA model is trained using historical data and future stock prices are predicted.
- d) Result visualization: The actual stock price is compared with the predicted stock price to evaluate the prediction effect of the model.

5.2 Future Work

In the future, we can further study the mechanism by which China's GDP affects stock prices, and try to use other time series models for prediction, such as the SARIMA model or the LSTM model, to improve the accuracy of the prediction. In addition, the model can also be applied to other stocks or markets for analysis and prediction, providing investors with more effective decision support.