2 TERM SPREAD PREDICTION USING LASSO IN MACHINE LEARNING FRAMEWORKS [1]

Daeyun KANG[2]

Doojin RYU[2,*]

Alexander WEBB[3]

Abstract

This study predicts the term spread using various machine learning models. Given that numerous macroeconomic variables can be used for term spread prediction, 116 variables are considered, and key variables are selected and extracted using LASSO. The core of the research lies in comparing two methodologies for predicting the term spread. The first method involves directly forecasting the spread itself, while the second method predicts long-term and short-term yields separately and then generates the spread from those predictions. The results indicate that the approach of directly predicting the term spread is statistically significantly superior. Our analysis of various forecasting models reveals that Long Short-Term Memory (LSTM), which can effectively capture nonlinear characteristics, demonstrates particularly strong performance in financial time series forecasting. These findings provide an effective approach to predicting the term spread and may serve as an important foundation for future research.

Keywords: Forecasting; LASSO; Long Short-Term Memory; Machine learning; Term spread

JEL classification: C45, E43, G17

1. Introduction

The term spread represents the difference between long-term and short-term bond yields, commonly calculated as the difference between the 10-year and 2-year U.S. Treasury yields, also known as the "10-year minus 2-year spread." The term spread reflects the business cycle, with the premium on long-term bonds being lower (higher) and short-term bond yields being higher (lower) during economic expansions (recessions) (Ang, Piazzesi, and Wei, 2006; Chun, Cho, and Ryu, 2023). Many studies employ the term spread and yield curve for macroeconomic analyses, often focusing on economic growth (Ang and Piazzesi, 2003; Kim and Ryu, 2020; Hamilton and Wu, 2012; Wu and Xia, 2016).

Harvey (1988) suggests that the real yield curve provides higher explanatory power for consumption growth compared to lagged consumption growth and stock returns, indicating the

[1] *Acknowledgement : We appreciate Lucian Liviu Albu and Corina Saman (Editors-in-chief).*

[2] *Department of Economics, Sungkyunkwan University, Seoul, Korea*

[*] *Corresponding author, E-mail: sharpjin@skku.edu*

[3] *Faculty of Business and Law, University of Wollongong, Australia*

possibility of a linear relationship between the two. Taylor (1993) proposes the Taylor Rule, suggesting that the Federal Reserve should lower nominal yields by 1 percentage point for every 2 percentage point decrease in GDP growth. Haubrich and Dombrosky (1996) confirm the significant predictive power of the term spread between the yield on the 10-year U.S. Treasury note and the 3-month U.S. Treasury bill yield for economic growth. Gürkaynak, Sack, and Wright (2010) introduce a method to smooth the yield curve and extract inflation expectations. Wang, Chang, Mikhaylov, and Linyu (2024) examine the interconnections among the U.S. Treasury yield spread, the U.S. dollar, and gold prices, analyzing how market stress and resilience change across various timeframes and quantiles. Other studies on credit ratings, exchange rate fluctuations, and monetary policy underscore the importance of the term spread as a key variable in macroeconomic analysis (Bernanke and Reinhart, 2004; Chen and Tsang, 2013; Riaz, Shehzad, and Umar, 2021; Swanson and Williams, 2014; Yu and Ryu, 2019). Conversely, predicting the term spread and yield curve can involve numerous variables (Favero, Niu, and Sala, 2012; Ludvigson and Ng, 2009; Mönch, 2008). Cochrane and Piazzesi (2005) demonstrate that a single factor derived from initial forward rates can predict one- to five-year bond excess returns with an R^2 of up to 0.44, which is countercyclical, forecasts stock returns, and is unaffected by typical movements in term structure models. Diebold, Rudebusch, and Aruoba (2006) show that real activity, inflation, and monetary policy significantly influence yield curve movements. Jotikasthira, Le, and Lundblad (2015) find that the global inflation index and major currency pairs explain most yield curve movements. Coroneo, Giannone, and Modugno (2016) find that key macroeconomic factors like economic growth and real interest rates have substantial predictive power over the yield curve. In sum, forecasting the term spread is useful for understanding macroeconomic variables like business cycles and economic growth, and a variety of macroeconomic factors can be employed in its prediction.

It is often challenging to precisely determine which model is suitable for a particular study when fitting models. Consequently, probability models generated using machine learning methodologies, which are more flexible than traditional model fitting and hypothesis testing methods, tend to exhibit higher predictive accuracy (Kelly and Xiu, 2023). With their fitting ability and applicability, machine-learning-based approaches have started to be widely adopted in the field of finance and economics (Bang and Ryu, 2023; Kim, Cho, and Ryu, 2021a, 2021b, 2022, 2024; Kim, Park, and Ryu, 2024; Park and Ryu 2021; Ryu, Hong, and Jo, 2024). This trend has led to the frequent use of Artificial Neural Networks (ANN) in financial time series forecasting research, with recent growing interest in Long Short-Term Memory (LSTM) networks (Kim, Cho, and Ryu, 2023; Park and Ryu, 2021). Conducting a study on KOSPI stock price prediction using ANN, Kim and Han (2000) report that models incorporating neural networks outperform those without them. Göçken, Özçalıcı, Boru, and Dosdoğru (2016) also emphasize that models combined with ANN demonstrate superior performance from both statistical and financial perspectives. Kim, Ryu, and Webb (2024) forecast the oil futures market dynamics using XGBoost (XGB) and Random Forests (RF). Zhang, Chu, and Shen (2021) demonstrate that LSTM is superior to other neural network models in handling complex, non-linear, and non-stationary financial time series data. Ghosh, Neufeld, and Sahoo (2022) show that by using RF and LSTM to select the top 10 stocks with expected high daily returns and short the bottom 10, it is possible to achieve substantial returns even after accounting for transaction costs. Chen and Ge (2019) utilize LSTM to predict movements in Hong Kong stocks and propose trading strategies based on these predictions. Nelson, Pereira, and De Oliveira (2017) also compare LSTM with ANN and conclude that LSTM is superior. Basak, Kar, Saha, Khaidem, and Dey (2019) explore stock trend forecasting using RF and XGB, suggesting the potential for developing trading strategies and portfolio management. Their findings indicate that these algorithms improve return prediction accuracy compared to existing models, minimizing investment risk. Yıldırım, Toroslu, and Fiore (2021) demonstrate that using LSTM to forecast various currency pair movements in the foreign exchange market can create profit opportunities. As such, LSTM is commonly used in financial time series forecasting research compared to tree-based models.

This study aims to predict the term spread using various macroeconomic indicators. First, we use LASSO, a powerful regression technique that combines variable selection and regularization, to extract the variables that influence the term spread and long-term and short-term bond yields among the 116 macroeconomic indicators in a high-dimensional data context (Bang and Ryu, 2024). Next, we investigate which approach is superior: directly forecasting the term spread or separately forecasting the 10-year long-term government bond yield and the 2-year short-term government bond yield before calculating the spread. As these two methods are compared on a one-to-one basis within the same model, we use the Giacomini-White statistic (GW statistic) to test the statistical significance of the difference between the two sets of predictions. Lastly, we compare tree-based models and the LSTM model to identify the best-performing model. The Model Confidence Set (MCS) is utilized to facilitate comparisons among multiple models.

The remaining sections of this study are structured as follows. Section 2 examines the sample data, and Section 3 details the methodology. Section 4 presents the prediction results of each model along with the findings. Section 5 concludes the study.

2. Sample Data

This study aims to forecast the term spread of U.S. Treasury bonds. Since various macroeconomic variables, including the term spread, have significant interactions with each other, previous studies have used numerous macroeconomic variables to make predictions when forecasting macroeconomic outcomes (Bokun, Jackson, Kliesen, and Owyang, 2023; Kim and Ryu, 2020, 2021; Kim, Ryu, and Yu, 2021; McCracken and Ng, 2016; Zhang, Wahab, and Wang, 2023). To ensure data consistency, we use the Federal Reserve Economic Data (FRED) and download the data via API (https://fred.stlouisfed.org/). The dataset consists of monthly data from January 2000 to January 2024. The macroeconomic variables used span across comprehensive economic indicators, covering output, income, labor market, housing, money, credit, interest rates, exchange rates, prices, and stock markets. The methods for variable selection and data stabilization are based on the research by Medeiros, Vasconcelos, Veiga, and Zilberman (2021). Details on the economic indicators and data stabilization methods are presented in Table 1.

Table 1.

Data description

Panel A. Output and income		
Series ID	Title	Tcode
RPI	Real Personal Income	5
W875RX1	Real personal income excluding current transfer receipts	5
INDPRO	Industrial Production: Total Index	5
IPFPNSS	Industrial Production: Final Products and Nonindustrial Supplies	5
IPFINAL	Industrial Production: Final Products	5
IPCONGD	Industrial Production: Consumer Goods	5
IPDCONGD	Industrial Production: Durable Consumer Goods: Durable Consumer Goods	5
IPNCONGD	Industrial Production: Nondurable Consumer Goods	5

IPBUSEQ	Industrial Production: Equipment: Business Equipment	5
IPMAT	Industrial Production: Materials	5
IPDMAT	Industrial Production: Durable Goods Materials: Durable Goods Materials	5
IPNMAT	Industrial Production: Nondurable Goods Materials	5
IPMANSICS	Industrial Production: Manufacturing (SIC)	5
IPB51222s	Industrial Production: Nondurable Energy Consumer Goods: Residential Utilities	5
IPFUELS	Industrial Production: Nondurable Energy Consumer Goods: Fuels	5
CUMFNS	Capacity Utilization: Manufacturing (SIC)	2

Panel B. Labor market		
Series ID	Title	Tcode
CLF16OV	Civilian Labor Force Level	5
CE16OV	Employment Level	5
UNRATE	Unemployment Rate	2
UEMPMEAN	Average Weeks Unemployed	2
UEMPLT5	Number Unemployed for Less Than 5 Weeks	5
UEMP5TO14	Number Unemployed for 5-14 Weeks	5
UEMP15OV	Number Unemployed for 15 Weeks & over	5
UEMP15T26	Number Unemployed for 15-26 Weeks	5
UEMP27OV	Number Unemployed for 27 Weeks & over	5
PAYEMS	All Employees, Total Nonfarm	5
USGOOD	All Employees, Goods-Producing	5
CES1021000001	All Employees, Mining, Quarrying, and Oil and Gas Extraction	5
USCONS	All Employees, Construction	5
MANEMP	All Employees, Manufacturing	5
DMANEMP	All Employees, Durable Goods	5
NDMANEMP	All Employees, Nondurable Goods	5
SRVPRD	All Employees, Service-Providing	5
USTPU	All Employees, Trade, Transportation, and Utilities	5
USWTRADE	All Employees, Wholesale Trade	5
USTRADE	All Employees, Retail Trade	5
USFIRE	All Employees, Financial Activities	5
USGOVT	All Employees, Government	5
CES0600000007	Average Weekly Hours of Production and Nonsupervisory Employees, Goods-Producing	2
AWOTMAN	Average Weekly Overtime Hours of Production and Nonsupervisory Employees, Manufacturing	2
AWHMAN	Average Weekly Hours of Production and Nonsupervisory Employees, Manufacturing	2

CES0600000008	Average Hourly Earnings of Production and Nonsupervisory Employees, Goods-Producing	5
CES2000000008	Average Hourly Earnings of Production and Nonsupervisory Employees, Construction	5
CES3000000008	Average Hourly Earnings of Production and Nonsupervisory Employees, Manufacturing	5

Panel C. Housing		
Series ID	Title	Tcode
HOUST	New Privately-Owned Housing Units Started: Total Units	4
HOUSTNE	New Privately-Owned Housing Units Started: Total Units in the Northeast Census Region	4
HOUSTMW	New Privately-Owned Housing Units Started: Total Units in the Midwest Census Region	4
HOUSTS	New Privately-Owned Housing Units Started: Total Units in the South Census Region	4
HOUSTW	New Privately-Owned Housing Units Started: Total Units in the West Census Region	4
PERMIT	New Privately-Owned Housing Units Authorized in Permit-Issuing Places: Total Units	4
PERMITNE	New Privately-Owned Housing Units Authorized in Permit-Issuing Places: Total Units in the Northeast Census Region	4
PERMITMW	New Privately-Owned Housing Units Authorized in Permit-Issuing Places: Total Units in the Midwest Census Region	4
PERMITS	New Privately-Owned Housing Units Authorized in Permit-Issuing Places: Total Units in the South Census Region	4
PERMITW	New Privately-Owned Housing Units Authorized in Permit-Issuing Places: Total Units in the West Census Region	4

Panel D. Money and credit		
Series ID	Title	Tcode
M1SL	M1	5
M2SL	M2	5
M2REAL	Real M2 Money Stock	5
TOTRESNS	Reserves of Depository Institutions: Total	5
NONBORRES	Reserves of Depository Institutions: Nonborrowed	5
BUSLOANS	Commercial and Industrial Loans, All Commercial Banks	5
REALLN	Real Estate Loans, All Commercial Banks	5
NONREVSL	Nonrevolving Consumer Credit Owned and Securitized	5
DTCOLNVHFNM	Consumer Motor Vehicle Loans Owned by Finance Companies, Level	5

DTCTHFNM	Total Consumer Loans and Leases Owned and Securitized by Finance Companies, Level	5
INVEST	Securities in Bank Credit, All Commercial Banks	5
Panel E. Interest rates and exchange rates		
Series ID	Title	Tcode
FEDFUNDS	Federal Funds Effective Rate	2
AAA	Moody's Seasoned Aaa Corporate Bond Yield	2
BAA	Moody's Seasoned Baa Corporate Bond Yield	2
AAAFFM	Moody's Seasoned Aaa Corporate Bond Minus Federal Funds Rate	2
BAAFFM	Moody's Seasoned Baa Corporate Bond Minus Federal Funds Rate	2
DFF	Federal Funds Effective Rate	2
DGS1	Market Yield on U.S. Treasury Securities at 1-Year Constant Maturity	2
DGS10	Market Yield on U.S. Treasury Securities at 10-Year Constant Maturity	2
DGS2	Market Yield on U.S. Treasury Securities at 2-Year Constant Maturity	2
DGS20	Market Yield on U.S. Treasury Securities at 20-Year Constant Maturity	2
DGS3	Market Yield on U.S. Treasury Securities at 3-Year Constant Maturity	2
DGS30	Market Yield on U.S. Treasury Securities at 30-Year Constant Maturity	2
DGS5	Market Yield on U.S. Treasury Securities at 5-Year Constant Maturity	2
DGS7	Market Yield on U.S. Treasury Securities at 7-Year Constant Maturity	2
DPRIME	Bank Prime Loan Rate	2
DTB1YR	1-Year Treasury Bill Secondary Market Rate, Discount Basis	2
DTB3	3-Month Treasury Bill Secondary Market Rate, Discount Basis	2
DTB6	6-Month Treasury Bill Secondary Market Rate, Discount Basis	2
T10Y2Y	10-Year Treasury Constant Maturity Minus 2-Year Treasury Constant Maturity	2
T10Y3M	10-Year Treasury Constant Maturity Minus 3-Month Treasury Constant Maturity	2
T10YFF	10-Year Treasury Constant Maturity Minus Federal Funds Rate	2
T1YFF	1-Year Treasury Constant Maturity Minus Federal Funds Rate	2
T5YFF	5-Year Treasury Constant Maturity Minus Federal Funds Rate	2
EXUSEU	U.S. Dollars to Euro Spot Exchange Rate	2
EXCAUS	Canadian Dollars to U.S. Dollar Spot Exchange Rate	2
EXJPUS	Japanese Yen to U.S. Dollar Spot Exchange Rate	2
EXCHUS	Chinese Yuan Renminbi to U.S. Dollar Spot Exchange Rate	2
EXSZUS	Swiss Francs to U.S. Dollar Spot Exchange Rate	2

Panel F. Prices and stock

Series ID	Title	Tcode
WPSFD49207	PPI: Final Demand: Finished Goods	5
WPSFD49502	PPI: Final Demand: Personal Consumption Goods	5
WPSID61	PPI: Intermediate Demand by Commodity Type: Processed Goods for Intermediate Demand	5
WPSID62	PPI: Intermediate Demand by Commodity Type: Unprocessed Goods for Intermediate Demand	5
CPIAUCSL	CPI for All Urban Consumers: All Items in U.S. City Average	5
CPIAPPSL	CPI for All Urban Consumers: Apparel in U.S. City Average	5
CPITRNSL	CPI for All Urban Consumers: Transportation in U.S. City Average	5
CPIMEDSL	CPI for All Urban Consumers: Medical Care in U.S. City Average	5
CUSR0000SAC	CPI for All Urban Consumers: Commodities in U.S. City Average	5
CUUR0000SAD	CPI for All Urban Consumers: Durables in U.S. City Average	5
CUSR0000SAS	CPI for All Urban Consumers: Services in U.S. City Average	5
CPIULFSL	CPI for All Urban Consumers: All Items Less Food in U.S. City Average	5
CUUR0000SA0L2	CPI for All Urban Consumers: All Items Less Shelter in U.S. City Average	5
CUSR0000SA0L5	CPI for All Urban Consumers: All Items Less Medical Care in U.S. City Average	5
PCEPI	PCE: Chain-type Price Index	5
DDURRG3M086SBEA	PCE: Durable goods (chain-type price index)	5
DNDGRG3M086SBEA	PCE: Nondurable goods (chain-type price index)	5
DSERRG3M086SBEA	PCE: Services (chain-type price index)	5
DCOILWTICO	Crude Oil Prices: West Texas Intermediate (WTI) - Cushing, Oklahoma	5
DCOILBRENTEU	Crude Oil Prices: Brent - Europe	5
NASDAQ100	NASDAQ 100 Index	5
NASDAQCOM	NASDAQ Composite Index	5
VIXCLS	CBOE Volatility Index (VIX)	1

Note: The "Series ID" refers to the identifier code provided by FRED for each data series, and the "Title" describes each data series. The "Tcode" represents the methodology used for data stabilization, where 1 indicates the original variable, 2 is for the first difference, 3 is for the second difference, 4 denotes a log transformation, and 5 indicates a log difference.

3. Method

RF is an ensemble learning method that generates multiple decision trees and performs the final prediction through majority voting. The key hyperparameters of RF are as follows: the number of trees is set to 500, and the node size is set to 5, requiring each leaf node to contain at least 5 samples. This helps control model complexity and prevents overfitting. RF uses bootstrap sampling to randomly select samples from the original dataset, allowing the same sample to be selected multiple times.

XGB is a boosting technique introduced by Chen and Guestrin (2016) that sequentially trains multiple weak prediction models (typically decision trees), enhancing prediction performance by correcting the errors of previous models. The main hyperparameters of XGB include a learning rate of 0.05 to control the pace of learning, and the number of iterations is set to 1000 to ensure sufficient training. The number of CPU threads is set to 1, and the column sample rate per tree level is set to 2/3 to prevent overfitting. The sample rate is set to 1, and the maximum tree depth is limited to 4 to control model complexity. The minimum child weight is set to 1/200 of the data row count, determining the minimum sample size for each split.

Light Gradient Boosting Machine (L-GBM), introduced by Ke et al. (2017), is a gradient-boosting framework designed for fast training and memory efficiency, particularly for large datasets. The key hyperparameters for L-GBM are as follows: the maximum number of leaf nodes per tree is set to 31 to control model complexity. A learning rate of 0.05 is used to adjust the update speed, and the number of iterations is set to 1,000 to allow the model to generate predictions through multiple trees.

LSTM introduced by Hochreiter and Schmidhuber (1997) is a neural network structure capable of processing both long-term and short-term dependencies, making it highly effective for time-series data and natural language processing. The key hyperparameters of LSTM include batch size set to 25 to control the amount of data fed at once, and the number of explanatory variables is specified as features. The model is trained for 100 epochs to ensure adequate learning, with a single-layer architecture of 32 units, and a final output layer with a single node. The model uses mean squared error (MSE) as the loss function and applies the Adam optimizer. After prediction, results are denormalized to their original scale.

For input variables, only those extracted through LASSO, as outlined by Tibshirani (1996), are used. A rolling window method is applied with a window size of 200. Each variable incorporates up to 4 lags, and principal component analysis (PCA) is used to extract 4 principal components, following the methodology of Medeiros, Vasconcelos, Veiga, and Zilberman (2021).

4. Results

Table 2 shows the variables selected through LASSO. T10Y2Y represents the yield spread between 10-year and 2-year Treasury securities and is linked to various macroeconomic indicators. Industrial production indicators are associated with T10Y2Y, reflecting production levels within the industrial sector. The production of nondurable energy consumer goods indicates energy demand. Labor market data also affect T10Y2Y. The unemployment rate provides insights into employment levels, while long-term unemployment reflects sustained employment conditions. In general, higher unemployment rates can lead to a smaller T10Y2Y spread. Employment figures in the goods-producing and durable goods sectors offer insights into overall labor market trends, with the important caveat that low unemployment can sometimes be misleading if the overall labor force participation rate is not taken into account. Employment in trade, transportation, and utilities is tied to industrial activity, while wholesale trade employment indicates broader economic trends.

Financial indicators significantly influence T10Y2Y. The risk premium for corporate bonds is tied to credit risk, and the federal funds rate reflects short-term borrowing costs. Short-term Treasury bill rates reflect market expectations for interest rates. Inflation pressures are evaluated through price changes in final consumer goods and intermediate goods, with crude oil price changes directly impacting inflation and economic growth expectations. Black swan events such as the Sept 11th, 2001 attacks, and the COVID-19 Pandemic can bring about sudden changes in the market and act in counterintuitive ways. These variables collectively contribute to the T10Y2Y spread's dynamics and indicate market perceptions of the broader economy.

DGS2, representing the yield on 2-year U.S. Treasury securities, is impacted by its historical yields, new privately-owned housing permits, the Baa corporate bond yield, primary and secondary market yields for similar maturities, VIX, and the exchange rate (EXSZUS). Among these factors, the most significant influence on DGS2 comes from the consumer price index for all urban consumers: services in the U.S. city average. Similarly, DGS10 is highly correlated with itself and the 5-year constant maturity treasury yield, and is also influenced by commercial and industrial loans and the VIX, which is implied by options price dynamics and gauges investors' fears and sentiments (Chen, Han, Ryu, and Tang, 2022; Song, Ryu, and Webb, 2018).

Table 2.

Selected variables

T10Y2Y		DGS2		DGS10	
Variable	Coef.	Variable	Coef.	Variable	Coef.
T10Y2Y	0.203	DGS2	0.066	DGS10	-0.003
IPFPNSS	-0.660	PERMIT	0.001	BUSLOANS	0.001
IPFUELS	-0.008	PERMITMW	0.015	DGS5	0.133
UNRATE	0.018	BAA	0.004	VIXCLS	0.005
UEMP27OV	0.066	DGS1	0.101		
USGOOD	-0.034	DTB1YR	0.001		
DMANEMP	-0.014	DTB6	0.002		
USTPU	-0.296	EXSZUS	0.109		
USWTRADE	-0.082	CUSR0000SAS	0.572		
M2SL	-0.195	NASDAQ100	0.017		
FEDFUNDS	-0.007	VIXCLS	0.000		
BAAFFM	0.000				
DFF	0.000				
DTB3	-0.010				
DTB6	-0.028				
WPSFD49502	0.020				
WPSID62	0.000				
DCOILWTICO	0.010				

Note: "Variable" refers to the selected variables, while "Coef." indicates the average coefficient for each respective variable.

Table 3 presents the forecasting results. For the RF model, the Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) for Spread1 are 0.123 and 0.097, respectively, while for Spread2, the RMSE and MAE are 0.108 and 0.084. Although Spread2 seems to perform better based on these metrics, the Giacomini-White statistic, as proposed by Giacomini and White (2006), is -1.17, indicating that Spread1 has a smaller prediction error. However, the statistical significance is 0.24, suggesting that the difference in approach between Spread1 and Spread2 is not statistically significant.

In the XGB, Spread1 has an RMSE and MAE of 0.132 and 0.105, respectively, while Spread2's RMSE and MAE are 0.154 and 0.123. For L-GBM, Spread1's RMSE and MAE are 0.132 and 0.106, while Spread2's are 0.175 and 0.137. In the LSTM model, the RMSE and MAE for Spread1 are 0.129 and 0.098, and for Spread2, they are 0.124 and 0.097. For these models, Spread1 generally outperforms Spread2 based on the forecast errors, and the GW statistic confirms that Spread1 has significantly smaller prediction errors. Therefore, it can be concluded that predicting the term spread directly is more effective than predicting the long-term and short-term yields separately when forecasting term spreads.

Table 3.

Forecasting results

Model	Spread1		Spread2		GW_Statistic	GW_Pvalue
	RMSE	MAE	RMSE	MAE		
RF	0.123	0.097	0.108	0.084	-1.17	0.24
XGB	0.132	0.105	0.154	0.123	-4.49	0.00
L-GBM	0.132	0.106	0.175	0.137	-3.83	0.00
LSTM	0.129	0.098	0.124	0.097	-4.06	0.00

Note: "Model" refers to the model used for prediction, where "Spread1" represents the case where the term spread itself is predicted, and "Spread2" represents the scenario where the short- and long-term yields are predicted separately, and the spread is then calculated from the results. "RMSE" stands for RMSE, which measures the square root of the average squared differences between predicted and actual values. "MAE" stands for MAE, which calculates the average absolute differences between the predicted and actual values. "GW_Statistic" refers to the test statistic used to assess the difference in predictive performance between two models, and "GW_Pvalue" evaluates whether this difference is statistically significant.

Table 4 shows the superior set of models about Spread2. According to the MCS introduced by Hansen, Lunde, and Nason (2011), this statistical methodology is used to evaluate the relative superiority of multiple models. MCS compares alternative models and selects a set of reliable models based on a specified confidence level. In the results, RF and L-GBM are eliminated, leaving LSTM and XGB in the superior set. Among these, LSTM ranks first, showing a performance value of -0.065 and a Loss of -0.0057, demonstrating outstanding performance and a high-reliability score of 1.00 in MCS. In contrast, XGB ranks second, with a performance value of 0.065, an MCS score of 0.95, and a loss value of -0.0051, indicating slightly lower performance than LSTM but still strong overall. In summary, LSTM demonstrates better predictive performance and reliability compared to XGB. Figure 1 presents the prediction results for each model.

Table 4.

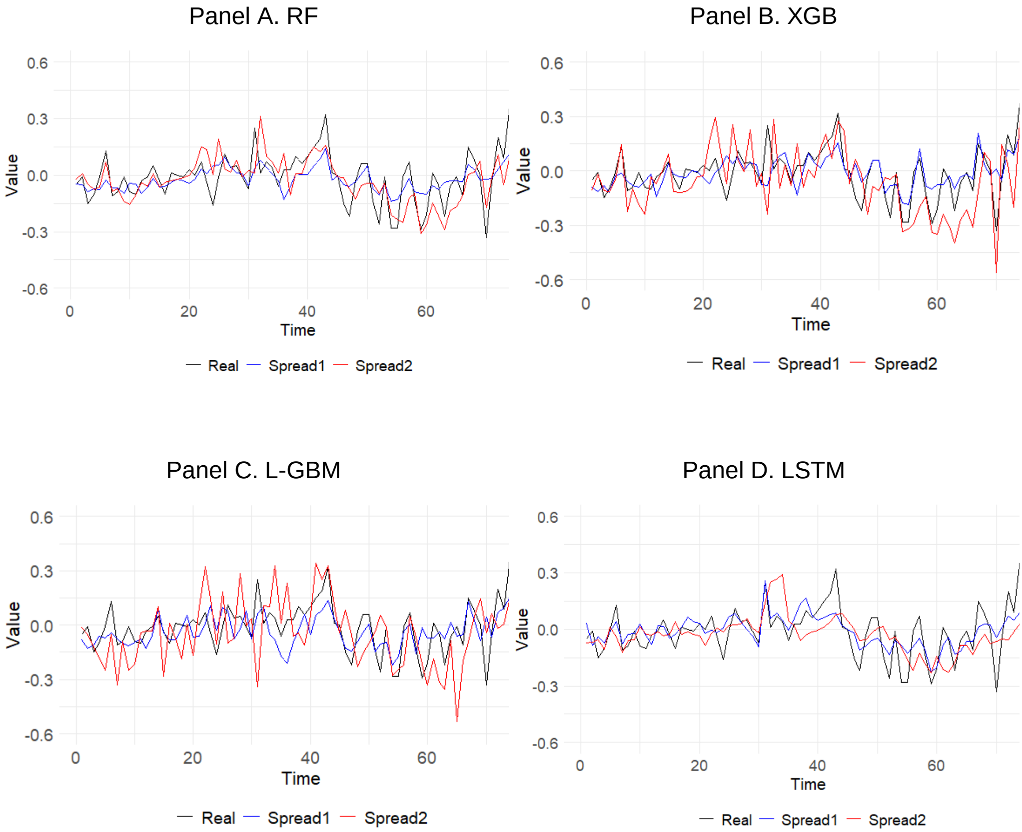
Superior set of models

Model	Rank	PV	MCS	Loss
LSTM	1	-0.065	1.00	-0.0057
XGB	2	0.065	0.95	-0.0051

Note: "Model" refers to the model used for prediction, while "Rank" indicates the ranking among the models. "PV" stands for the model performance value, and "MCS" represents the statistical confidence level associated with this performance. "Loss" refers to the model's average loss value.

Figure 1.

Model Performance



5. Conclusion

This study focuses on predicting the term spread by extracting and analyzing key macroeconomic variables that influence long-term and short-term Treasury yields, as well as the term spread itself, using the LASSO technique. Variables related to the term spread, in addition to the spread itself, primarily include those that capture economic conditions. For long-term and short-term yields, variables extracted include not only the yields themselves but also those related to similar-maturity bonds. The central focus of this research is to compare two methodologies for predicting the term spread. The first method entails forecasting the spread directly, whereas the second method involves predicting long-term and short-term yields independently and subsequently deriving the spread from those predictions. The findings suggest that directly forecasting the term spread is statistically significantly more effective, as confirmed by the GW statistics. Finally, when comparing the models RF, XGB, L-GBM, and LSTM using the MCS, it is found that RF and L-GBM are eliminated, with LSTM ranking first and XGB ranking second. This indicates that LSTM outperforms the other models in financial time series forecasting.

References

Ang, A. and Piazzesi, M., 2003. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics*, 50(4), pp.745-787. https://doi.org/10.1016/S0304-3932(03)00032-1

Ang, A., Piazzesi, M. and Wei, M., 2006. What does the yield curve tell us about GDP growth?. *Journal of Econometrics*, 131(1-2), pp.359-403. https://doi.org/10.1016/j.jeconom.2005.01.032

Bang, J. and Ryu, D., 2023. CNN-based stock price forecasting by stock chart images*. Romanian Journal of Economic Forecasting,* 26(3), pp.120-128. https://doi.org/10.1016/j.frl.2023.104315

Bang, J. and Ryu, D., 2024. ESG factors and the cross-section of expected stock returns: A LASSO-based approach. *Finance Research Letters*, 65, 105482. https://doi.org/10.1016/j.frl.2024.105482

Basak, S., Kar, S., Saha, S., Khaidem, L. and Dey, S.R., 2019. Predicting the direction of stock market prices using tree-based classifiers. *North American Journal of Economics and Finance*, 47, pp.552-567. https://doi.org/10.1016/j.najef.2018.06.013

Bernanke, B.S. and Reinhart, V.R., 2004. Conducting monetary policy at very low short-term interest rates. *American Economic Review*, 94(2), pp.85-90. https://doi.org/10.1257/0002828041302118

Bokun, K.O., Jackson, L.E., Kliesen, K.L. and Owyang, M.T., 2023. FRED-SD: A real-time database for state-level data with forecasting applications. *International Journal of Forecasting*, 39(1), pp.279-297. https://doi.org/10.1016/j.ijforecast.2021.11.008

Chen, J., Han, Q., Ryu, D. and Tang, J., 2022. Does the world smile together? A network analysis of global index option implied volatilities. *Journal of International Financial Markets, Institutions and Money*, 77, 101497. https://doi.org/10.1016/j.intfin.2021.101497

Chen, S. and Ge, L., 2019. Exploring the attention mechanism in LSTM-based Hong Kong stock price movement prediction. *Quantitative Finance*, 19(9), pp.1507-1515. https://doi.org/10.1080/14697688.2019.1622287

Chen, T. and Guestrin, C., 2016. *Xgboost: A scalable tree boosting system*. In: Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, pp.785-794. http://dx.doi.org/10.1145/2939672.2939785

Chen, Y.C. and Tsang, K.P., 2013. What does the yield curve tell us about exchange rate predictability?. *Review of Economics and Statistics*, 95(1), pp.185-205. https://doi.org/10.1162/REST_a_00231

Chun, D., Cho, H. and Ryu, D., 2023. Discovering the drivers of stock market volatility in a data-rich world. *Journal of International Financial Markets, Institutions and Money*, 82, 101684. https://doi.org/10.1016/j.intfin.2022.101684

Cochrane, J.H. and Piazzesi, M., 2005. Bond risk premia. *American Economic Review*, 95(1), pp.138-160. https://doi.org/10.1257/0002828053828581

Coroneo, L., Giannone, D. and Modugno, M., 2016. Unspanned macroeconomic factors in the yield curve. *Journal of Business & Economic Statistics*, 34(3), pp.472-485. https://doi.org/10.1080/07350015.2015.1052456

Diebold, F.X., Rudebusch, G.D. and Aruoba, S.B., 2006. The macroeconomy and the yield curve: a dynamic latent factor approach. *Journal of Econometrics*, 131(1-2), pp.309-338. https://doi.org/10.1016/j.jeconom.2005.01.011

Favero, C.A., Niu, L., and Sala, L., 2012. Term structure forecasting: No-arbitrage restrictions versus large information set. *Journal of Forecasting*, 31, pp.124–156. https://doi.org/10.1002/for.1181

Ghosh, P., Neufeld, A. and Sahoo, J. K., 2022. Forecasting directional movements of stock prices for intraday trading using LSTM and random forests. *Finance Research Letters*, 46, 102280. https://doi.org/10.1016/j.frl.2021.102280

Giacomini, R. and White, H., 2006. Tests of conditional predictive ability. *Econometrica*, 74(6), pp.1545-1578. https://doi.org/10.1111/j.1468-0262.2006.00718.x

Göçken, M., Özçalıcı, M., Boru, A. and Dosdoğru, A.T., 2016. Integrating metaheuristics and artificial neural networks for improved stock price prediction. *Expert Systems with Applications*, 44, pp.320-331. https://doi.org/10.1016/j.eswa.2015.09.029

Gürkaynak, R.S., Sack, B. and Wright, J.H., 2010. The TIPS yield curve and inflation compensation. *American Economic Journal: Macroeconomics*, 2(1), pp.70-92. https://doi.org/10.1257/mac.2.1.70

Hamilton, J.D. and Wu, J.C., 2012. Identification and estimation of Gaussian affine term structure models. *Journal of Econometrics*, 168(2), pp.315-331. https://doi.org/10.1016/j.jeconom.2012.01.035

Hansen, P.R., Lunde, A. and Nason, J.M., 2011. The model confidence set. *Econometrica*, 79(2), pp.453-497. https://doi.org/10.3982/ECTA5771

Harvey, C.R., 1988. The real term structure and consumption growth. *Journal of Financial Economics*, 22(2), pp.305-333. https://doi.org/10.1016/0304-405X(88)90073-6

Haubrich, J.G. and Dombrosky, A.M., 1996. Predicting real growth using the yield curve. *Economic Review*, 32(1), 26-35. https://core.ac.uk/download/pdf/6229957.pdf

Hochreiter, S. and Schmidhuber, J., 1997. Long short-term memory. *Neural Computation*, 9(8), pp.1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735

Jotikasthira, C., Le, A. and Lundblad, C., 2015. Why do term structures in different currencies co-move?. *Journal of Financial Economics*, 115(1), pp.58-83. https://doi.org/10.1016/j.jfineco.2014.09.004

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q. and Liu, T. Y., 2017. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*, 30. https://www.microsoft.com/en-us/research/wp-content/uploads/2017/11/lightgbm.pdf

Kelly, B. and Xiu, D., 2023. Financial machine learning. *Foundations and Trends® in Finance*, 13(3-4), pp.205-363. http://dx.doi.org/10.1561/0500000064

Kim, A., Ryu, D. and Webb. A., 2024. Forecasting oil futures markets using machine learning and seasonal trend decomposition. *Investment Analysts Journal*, (forthcoming).

Kim, H., Cho, H. and Ryu, D., 2021. Predicting corporate defaults using machine learning with geometric-lag variables. *Investment Analysts Journal*, 50(3), pp.161-175. https://doi.org/10.1080/10293523.2021.1941554

Kim, H., Cho, H. and Ryu, D., 2022. Corporate bankruptcy prediction using machine learning methodologies with a focus on sequential data. *Computational Economics*, 59(3), pp.1231-1249. https://doi.org/10.1007/s10614-021-10126-5

Kim, H., Cho, H. and Ryu, D., 2023. Measuring corporate failure risk: Does long short-term memory perform better in all markets?. *Investment Analysts Journal*, 52(1), pp.40-52. https://doi.org/10.1080/10293523.2022.2155353

Kim, H., Cho, H. and Ryu, D., 2024. Characteristics of student loan credit recovery: evidence from a micro-level data set. *Journal of Credit Risk*, 20(1), pp.27-49. https://doi.org/10.21314/JCR.2024.001

Kim, K. and Han, I.,2000. Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. *Expert Systems with Applications*, 19(2), pp.125-132. https://doi.org/10.1016/S0957-4174(00)00027-0

Kim, K. and Ryu, D., 2020. Predictive ability of investor sentiment for the stock market. *Romanian Journal of Economic Forecasting*, 23(4), pp.33-46. https://ipe.ro/new/rjef/rjef4_20/rjef4_2020p33-46.pdf

Kim, K. and Ryu, D., 2021. Does sentiment determine investor trading behaviour?. *Applied Economics Letters*, 28(10), pp.811-816. https://doi.org/10.1080/13504851.2020.1782331

Kim, K., Ryu, D. and Yu, J., 2021. Do sentiment trades explain investor overconfidence around analyst recommendation revisions?. *Research in International Business and Finance*, 56, 101376. https://doi.org/10.1016/j.ribaf.2020.101376

Kim, T., Park, E. and Ryu, D., 2024. Determinants of housing rental prices in Seoul: Application of explainable AI. *Spatial Economic Analysis*, (forthcoming).

Ludvigson, S.C., and Ng, S., 2009. Macro Factors in Bond Risk Premia. *Review of Financial Studies*, 22, pp.5027–5067. https://doi.org/10.1093/rfs/hhp081

McCracken, M.W. and Ng, S., 2016. FRED-MD: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics*, 34(4), pp.574-589. https://doi.org/10.1080/07350015.2015.1086655

Medeiros, M.C., Vasconcelos, G.F., Veiga, Á. and Zilberman, E., 2021. Forecasting inflation in a data-rich environment: the benefits of machine learning methods. *Journal of Business & Economic Statistics*, 39(1), pp.98-119. https://doi.org/10.1080/07350015.2019.1637745

Mönch, E., 2008), Forecasting the yield curve in a data-rich environment: A no-arbitrage factor-augmented var approach. *Journal of Econometrics*, 146, pp.26–43. https://doi.org/10.1016/j.jeconom.2008.06.002

Nelson, D.M., Pereira, A.C. and De Oliveira, R.A., 2017. *Stock market's price movement prediction with LSTM neural networks*. In: 2017 International Joint Conference on Neural Networks (IJCNN), pp.1419-1426, IEEE. https://doi.org/10.1109/IJCNN.2017.7966019

Park, D. and Ryu, D., 2021. A machine learning-based early warning system for the housing and stock markets. *IEEE Access*, 9, pp.85566-85572. https://doi.org/10.1109/ACCESS.2021.3077962

Park, D. and Ryu, D., 2021. Forecasting stock market dynamics using bidirectional long short-term memory. *Romanian Journal of Economic Forecasting*, 24(2), pp.22-34. http://www.ipe.ro/rjef/rjef2_21/rjef2_2021p22-34.pdf

Riaz, Y., Shehzad, C.T. and Umar, Z., 2021. The sovereign yield curve and credit ratings in GIIPS. *International Review of Finance*, 21(3), pp.895-916. https://doi.org/10.1111/irfi.12306

Ryu, D., Hong, J. and Jo, H., 2024. Capturing locational effects: application of the K-means clustering algorithm. *Annals of Regional Science*, 73, pp.265-289. https://doi.org/10.1007/s00168-024-01263-4

Song, W., Ryu, D. and Webb, R.I., 2018. Volatility dynamics under an endogenous Markov-switching framework: A cross-market approach. *Quantitative Finance*, 18(9), pp.1559-1571. https://doi.org/10.1080/14697688.2018.1444551

Swanson, E.T. and Williams, J.C., 2014. Measuring the effect of the zero lower bound on medium- and longer-term interest rates. *American Economic Review*, 104(10), pp.3154-3185. https://doi.org/10.1257/aer.104.10.3154

Taylor, J.B., 1993. *Discretion versus policy rules in practice*. In: Carnegie-Rochester conference series on public policy, 39, 195-214. https://doi.org/10.1016/0167-2231(93)90009-L

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(1), pp.267-288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Wang, M.C., Chang, T., Mikhaylov, A. and Linyu, J., 2024. A measure of quantile-on-quantile connectedness for the US treasury yield curve spread, the US Dollar, and gold price. *North American Journal of Economics and Finance*, 74, 102232. https://doi.org/10.1016/j.najef.2024.102232

Wu, J.C. and Xia, F.D., 2016. Measuring the macroeconomic impact of monetary policy at the zero lower bound. *Journal of Money, Credit and Banking*, 48(2-3), pp.253-291. https://doi.org/10.1111/jmcb.12300

Yıldırım, D.C., Toroslu, I.H. and Fiore, U., 2021. Forecasting directional movement of Forex data using LSTM with technical and macroeconomic indicators*. Financial Innovation*, 7, pp.1-36. https://doi.org/10.1186/s40854-020-00220-2

Yu, J. and Ryu, D., 2019. Predicting banks' subordinated bond issuances. *Romanian Journal of Economic Forecasting*, 22(4), 87-99. https://ipe.ro/new/rjef/rjef4_19/rjef4_19p87-99.pdf

Zhang, Y., Chu, G. and Shen, D., 2021. The role of investor attention in predicting stock prices: The long short-term memory networks perspective. *Finance Research Letters*, 38, 101484. https://doi.org/10.1016/j.frl.2020.101484

Zhang, Y., Wahab, M.I.M. and Wang, Y., 2023. Forecasting crude oil market volatility using variable selection and common factor. *International Journal of Forecasting*, 39(1), pp.486-502. https://doi.org/10.1016/j.ijforecast.2021.12.013