Proceedings of ICFTBA 2026 Symposium: Strategic Management & Business Analytics: Data-Driven Performance Models
DOI: 10.54254/2754-1169/2026.LD37894

Forecasting Stock Market Synchronicity: An Out-of-Sample Comparison of Statistical and Machine Learning Models

Xujiayun Deng

College of Literature and Science, University of California Davis, Davis, USA
xjydeng@ucdavis.edu

Abstract. Stock–market synchronicity measures the extent to which individual stock returns move with the overall market and is relevant to risk assessment and portfolio allocation. This study examines whether next-quarter synchronicity can be predicted and whether machine-learning models improve on simpler benchmarks. Daily price and volume data for 25 U.S. large-cap stocks from 2010 to 2025 were used to estimate quarterly R^2 values from regressions of individual stock returns on SPY returns. A persistence benchmark, ordinary least squares (OLS), decision tree, and random forest were evaluated on an untouched 2022–2025 test set using MAE, RMSE, predictive R^2, and out-of-sample R^2. OLS achieved the lowest RMSE (0.1873), the highest predictive R^2 (0.3255), and a 10.5% reduction in squared forecast error relative to persistence, although persistence produced the lowest MAE (0.1416). Both tree-based models performed worse in point prediction. Current synchronicity and VIX conditions were important predictors, while performance varied across volatility regimes and sectors. The random forest nevertheless separated stocks with relatively high and low future synchronicity. Overall, simple linear methods were more reliable for precise forecasting, whereas machine learning offered complementary ranking insights.

Keywords: Financial Engineering, Data Science, Statistic, Stock Market, Synchronicity

1. Introduction

In the existing research on the stock market, this study can observe that the prices of individual stocks are influenced by a large number of uncontrollable factors both from the companies themselves and the market. Moreover, the synchronicity between the trends of individual stocks and the overall market trend is a very important point for investigation. Previous studies have revealed the influence of factors such as inflation, financial reports, and investor sentiment on the overall market and individual stocks. This study, however, believes that the synchronicity between individual stocks and the overall market is a very worthy research direction. This is because for financial institutions and individual investors, this synchronicity can help them establish some new understandings of investment risks and asset allocation, thereby better assisting them in profiting from the market and individual stocks.

Morck et al established R2 as an important indicator for measuring the synchronicity of stock returns [1], while Durnev et al. and Chan and Chan explored the relationship between this indicator

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2754-1169/2026.LD37894

and company-specific variables [2,3]. This is why this research chose the R2 as the measurement variable for Stock Market Synchronicity. This indicator is mainly derived from the regression analysis of the index's return rate and the return rate of individual stocks. If the R2 value is high, then the majority of the fluctuations in individual stock returns can be explained by market volatility. If the R2 value is low, it indicates that there will be more other variables that have a greater impact on the fluctuations in the return rate. In this study, R2 will be used as an indicator to measure the synchronicity of stock returns.

In the studies by Brockman et al, it was found that stock co-volatility changes along with the business cycle [4]. Meanwhile, Whaley explained VIX as an important indicator for measuring market volatility [5]. Based on the ideas provided in the above papers and referring to the reliability of machine learning in financial forecasting in Gu et al.'s paper and his study determined the specific direction and methods [6]. Also, the study done by Breiman et al, tested random forests, which integrate multiple decision trees to improve the stability of predictions [7]. At the same time, it is noted that most papers in this field focus on the prediction of returns and risks, and rarely explore the synchronicity of stock market trends and the comparison of the accuracy of machine learning with traditional methods outside the sample.

This study, in an effort to fill this gap, covered the daily stock returns of 25 large American enterprises from five industries from 2010 to 2025, calculated the quarterly market R2 based on the SPY returns, and utilized the current R2, beta, returns, volatility, and volume changes. Variables such as VIX were used to predict the R2 value for the next quarter. Within this framework, this study compared Persistence benchmark, OLS, decision tree, and random forest. The data from 2010 to 2018 was used for training, 2019 to 2021 for validation, and 2022 to 2025 for testing. The contribution of this study lies in applying machine learning to the prediction of synchrony, and comparing the advantages and disadvantages between traditional statistical models and machine learning models.

2. Method

2.1. Dataset preparation

The dataset includes daily financial data from two sources. First, the adjusted closing price and trading volume were obtained by using the yfinance open-source tool of Python to download data from the Yahoo financial interface. These data cover 25 American enterprises and the SPDR S&P 500 Index ETF(SPY), which acts as a market agent. These 25 companies are divided into 5 industries: technology, finance, healthcare, consumer goods, and industry. In addition, the VIX data is provided by the Federal Reserve. The sample period ranges from January 2010 to December 2025. The dataset of prices and trading volumes includes data for a total of 4,024 trading days, while the VIX dataset contains 4,051 observations. This study converts the daily prices into returns and employs the following model:

$$r_{i,d} = \alpha_{i,q} + \beta_{i,q} r_{\mathrm{SPY},d} + \varepsilon_{i,d} \tag{1}$$

Obtained R2 by this model, and then measures the proportion of SPY's explanation for the quarterly stock return volatility. In this study, R2 is an important indicator for measuring the synchronization between the market and stocks. This study excluded quarterly periods with fewer than 40 trading days from the dataset and only retained those with 60 or more trading days. The beta coefficient is composed of composite stock returns, annualized stock volatility, average percentage

DOI: 10.54254/2754-1169/2026.LD37894

change in trading volume, composite SPY returns, annualized SPY volatility, mean VIX, and VIX percentage change. In addition, the industry classification of stocks into five categories will be added as predictor variables to the model.

This study ultimately retained data for 25 stocks over 64 quarters, including 1,600 stock-quarter observations. The prediction dataset utilized 9 numerical predictor variables and 1 categorical predictor variable to forecast the value of R2 for the next quarter. A total of 1575 target variables were generated, which were created by shifting the R2 value of each stock one quarter forward. This study does not define the first trading volume change observation for each stock and instead uses the median to replace the missing values in the prediction pipeline. In the linear prediction model, this study standardized the data, but the tree model did not. The OLS inference model utilizes all the observed data.

The training set of this study consists of 875 observed values from Q2 2010 to Q4 2018. The validation set comprises 300 observed values from Q1 2019 to Q4 2021, while the test set consists of 400 observed values from Q1 2022 to Q4 2025. The figure below 1 presents the trend of the quarterly average synchronization rate. To preserve the temporal ordering of the data and prevent future information from entering model estimation, the sample was divided chronologically into training, validation, and test periods, following established principles for out-of-sample forecast evaluation [8].

2.2. Prediction methods

2.2.1. Persistence benchmark

The assumption of the persistence benchmark is as follows:

$$\widehat{R^2}_{i,q+1} = R^2_{i,q} \tag{2}$$

In this study, to predict the Q2 of 2022, Persistence Benchmark directly used the values from Q1 of 2022. Then, this process was repeated for all 400 stock-quarter observations in the test set, and the overall MAE and RMSE were calculated.

2.2.2. Ordinary least squares

This study employed OLS regression to correlate the R^2 value in the next quarter with nine numerical predictor variables and industry indicators. OLS mainly estimated the direction and degree of the correlation.In order to obtain more reliable statistical inference in the presence of potential heteroskedasticity and influential observations, HC3 heteroskedasticity-consistent standard errors and 95% confidence intervals were reported [9].

2.2.3. Decision tree

This study mainly employed the CART regression tree to capture nonlinear relationships and interactions. The CART framework recursively partitions the predictor space and can therefore capture nonlinear relationships and interactions without imposing a predetermined functional form [10]. After verifying the RMSE, based on the minimum leaf size for different depths and different observation values, the maximum depth of 3 was finally selected, with at least 30 observations for

DOI: 10.54254/2754-1169/2026.LD37894

each leaf. By imposing these restrictions, this study result has shown the possibility to reduce the risks of overfitting and excessive noise.

2.2.4. Random forest

This study employed the average prediction results of 400 random regression trees for modeling, and selected a maximum depth of 4, a minimum leaf node number of 10, and 70% of the features. And the random seed was fixed at 42. Through 30 repetitions of the test set, the prediction contributions of each feature were evaluated.

3. Results and discussion

In this study, all the four models mentioned above will be compared in terms of MAE, RMSE, R^2, out of sample R^2, and Persistence in Figure 1, it can be observed that the synchronicity of the stock market in the sample shows significant uncertainty and does not exhibit a clear stable trend. The average value of R^2 is 0.369, but there are also extreme values. For instance, the synchronicity reached 0.732 in Q3 of 2011, 0.768 in Q1 of 2020, while it was only 0.117 in Q2 of 2024. It can be seen from this that merely relying on descriptive models is not sufficient to determine the relationship.

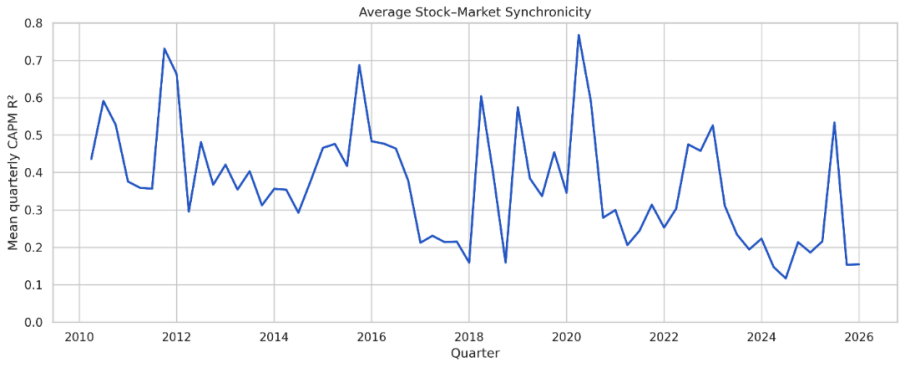


Figure 1. Average stock-market synchronicity (picture credit: original)

Table 1. Model performance on the 2022–2025 test set

Model	MAE	RMSE	Predictive R^2	OOS R^2 vs. persistence
OLS	0.1558	0.1873	0.3255	0.1050
Persistence	0.1416	0.1980	0.2464	0.0000
Random Forest	0.1836	0.2127	0.1302	-0.1541
Decision Tree	0.1944	0.2320	-0.0345	-0.3727

As a result, that OLS has the lowest RMSE and the highest R^2. In Table 1, the out-of-sample R $R^2$is 0.105, which indicates that the squared error of the prediction has decreased by 10.5%. However, the MAE of the Persistence Benchmark is the smallest, only 0.1416, while OLS is 0.1558. In summary, OLS performs better in controlling large errors.

What's more, in the Table1 can see that the performance of Random Forest and Decision Tree is inferior to that of OLS and Persistence Benchmark. Based on the quarter wise dependency relationships among stocks, also in this research, the quartered bootstrap method estimates that the

DOI: 10.54254/2754-1169/2026.LD37894

random forest RMSE difference is 0.0254, and 0.35 of the bootstrap samples show negative differences. Therefore, the high RMSE of the random forest is not caused by a few observations or the test quarter.

Table 2. Regression results with HC3 standard errors

Predictor	Coefficient	Robust SE	t statistic	p value	CI 2.5%	CI 97.5%
Intercept	0.150213	0.031032	4.840546	0.000001	0.089391	0.211035
C(Sector) [Energy Industrials]	0.128792	0.021707	5.933109	0.000000	0.086247	0.171338
C(Sector) [Financials]	0.159084	0.026564	5.988629	0.000000	0.107019	0.211149
C(Sector) [Healthcare]	0.048820	0.019235	2.538132	0.011145	0.011121	0.086519
C(Sector) [Technology]	0.051766	0.021559	2.401153	0.016344	0.009512	0.094021
market r2	0.402997	0.075500	5.337728	0.000000	0.255021	0.550974
market beta	-0.039574	0.045346	-0.872723	0.382814	-0.128450	0.049302
stock q return	0.182746	0.076233	2.397213	0.016520	0.033333	0.332159
stock volatility	0.071950	0.219548	0.327717	0.743126	-0.358357	0.502256
volume change	0.023266	0.030547	0.761648	0.446270	-0.036605	0.083137
spy q return	-0.444453	0.201686	-2.203689	0.027546	-0.839751	-0.049156
spy volatility	-3.061293	0.542474	-5.643204	0.000000	-4.124523	-1.998063
Vix mean	0.026230	0.003764	6.969265	0.000000	0.018853	0.033606
Vix change	0.158592	0.025268	6.276455	0.000000	0.109068	0.208116

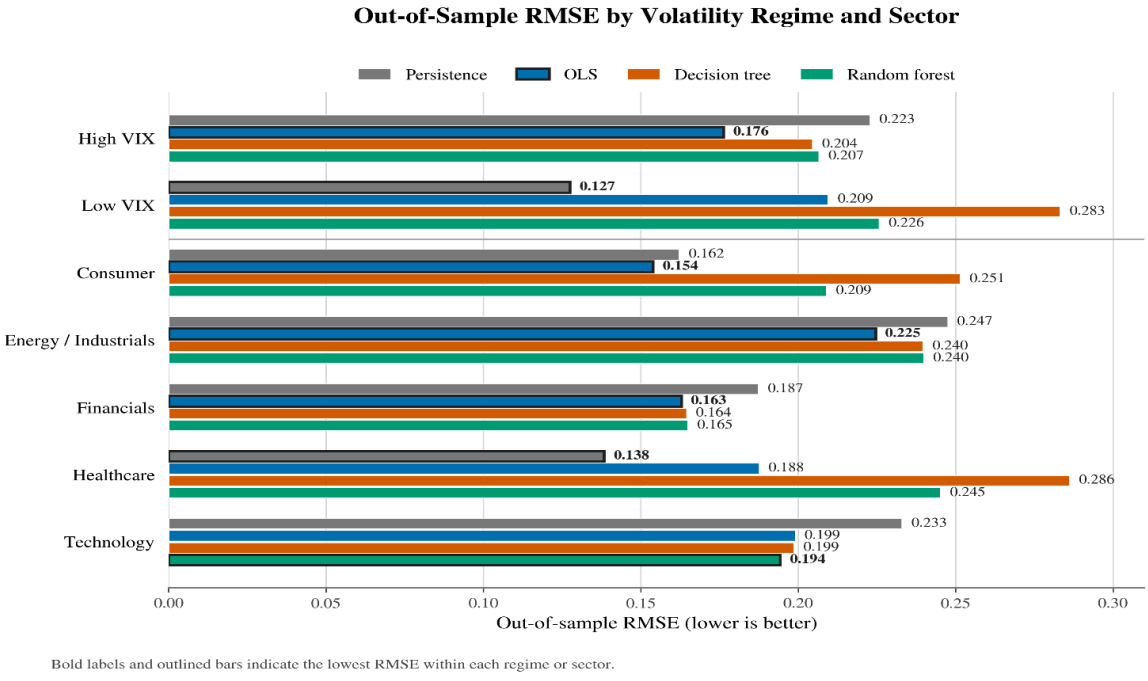


Figure 2. Out of sample RMSE by volatility regime and sector (picture credit: original)

In the OLS regression analysis shown in Table 2, it was found that the average VIX and VIX changes were positively correlated with the synchronicity of the next quarter, with coefficients of 0.026 and 0.159 respectively, and both were highly significant. Stock returns were also positively

DOI: 10.54254/2754-1169/2026.LD37894

correlated with future synchronicity, while SPY returns and SPY volatility were negatively correlated. Compared to consumer stocks, financial stocks and industrial stocks had the largest positive industry coefficients. Moreover, stock volatility and trading volume changes were not statistically significant.

In Figure 2, after defining the high VIX quarter using the median of 16.07, the RSME of OLS was 0.1763 while that of the Persistence Benchmark was 0.228. However, in the low VIX quarter, the RMSE of the Persistence Benchmark was 0.1275, which was significantly lower than the 2.095 of OLS. Furthermore, in the OLS model, the RMSE was the lowest for consumption, industrial, and financial stocks. In the random forest model, stocks from the technology sector outperformed other models, while the Persistence Benchmark performed the best in the healthcare sector. All the results of the decision tree model are basically the worst.

Table 3. High- versus low-predicted group comparison

Portfolio group	Observations	Mean predicted R^2	Mean realized R^2	Median realized R^2
Predicted High 5	80	0.461	0.447	0.429
Predicted Low 5	80	0.308	0.106	0.051

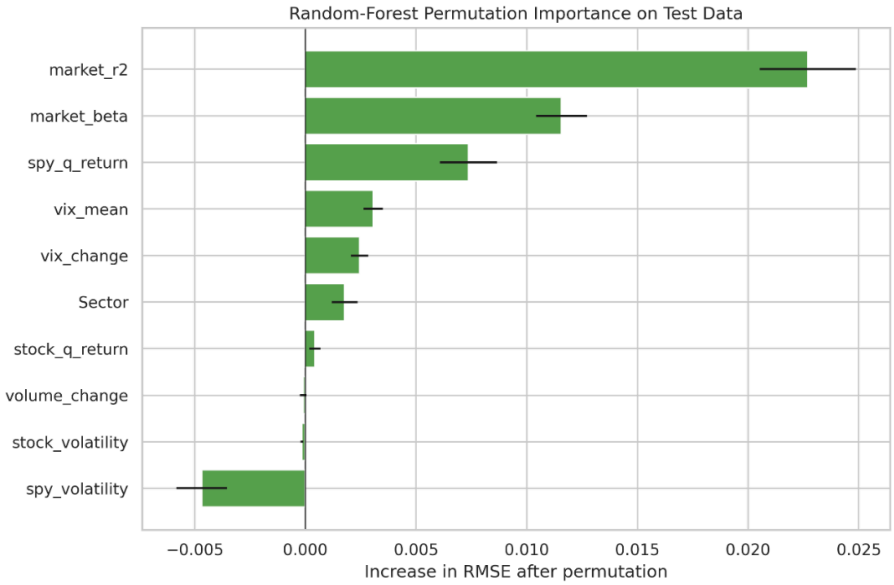


Figure 3. Random-forest feature importance (picture credit: original)

In Table 3 and Figure 3, the result of the random forest permutation importance ranking in this study indicates that R2 is the most effective predictor, followed by the beta coefficient and the SPY yield. Although its predictive ability is weaker than that of OLS, the random forest still has a good sorting ability. Among the 5 stocks with the lowest synchronization in the test, the average R2 value for the next quarter was 0.106, while for the top 5 stocks, the average was 0.447. This difference indicates that the random forest model was better at identifying relative market exposure than the precise market exposure value. The regime and sector analysis in Figure 2 further shows that model performance is heterogeneous. OLS performs best in high-volatility environments, while persistence dominates in calmer periods, and no single model is uniformly superior across sectors. Notably, both tree-based methods—especially the decision tree—tend to underperform in most regimes, reinforcing the advantage of linear structure in this setting.

DOI: 10.54254/2754-1169/2026.LD37894

Finally, in this study, the decision tree method performed significantly worse than the random forest in all indicators. Its prediction R2 was negative, and the RMSE was the highest among the four models. This clearly indicates that the single-tree structure model is unable to accurately capture the unstable linear relationships in synchronization.

4. Conclusion

This study utilized 25 specific U.S. stocks from five distinct sectors of the U.S. stock market during the period from 2010 to 2025. Based on these data, this paper examines the potential synchronicity of the stock market. The final conclusion of this study is that there is a moderate degree of persistence in the synchronicity, but the temporal changes are still significant, as shown in Figure 1. Among the four prediction methods used in this study, OLS demonstrated the lowest RMSE and the highest R on the test set, significantly outperforming the decision tree and random forest models. The continuous benchmark model maintained the lowest MAE. This study also found that during periods of high VIX, the OLS method performed the best, while during periods of low VIX, the persistent benchmark model was the most optimal. Based on the regression results, it can also be concluded that the current synchronicity and VIX are important predictors of synchronicity in the next quarter. Although the random forest performed poorly, it can be observed from Figure 2 that it has a significant advantage in sorting stocks. It might be more suitable for research on market exposure directions rather than for numerical predictions. Overall, based on all the data and evidence obtained from this experiment, the complexity of the model does not improve the accuracy of the predictions. Simple linear models still have significant competitiveness in numerical prediction compared to machine learning methods, and the combination of different types of models is the direction for future exploration.

References

[1] Morck, R., Yeung, B., & Yu, W. (2000). The information content of stock markets: Why do emerging markets have synchronous stock price movements? *Journal of Financial Economics, 58*(1–2), 215–260.
[2] Durnev, A., Morck, R., Yeung, B., & Zarowin, P. (2003). Does greater firm-specific return variation mean more or less informed stock pricing? *Journal of Accounting Research, 41*(5), 797–836. https://doi.org/10.1046/j.1475-679X.2003.00124.x
[3] Chan, K., & Chan, Y.-C. (2014). Price informativeness and stock return synchronicity: Evidence from the pricing of seasoned equity offerings. *Journal of Financial Economics, 114*(1), 36–53. https://doi.org/10.1016/j.jfineco.2014.07.002
[4] Brockman, P., Liebenberg, I., & Schutte, M. (2010). Comovement, information production, and the business cycle. *Journal of Financial Economics, 97*(1), 107–129. https://doi.org/10.1016/j.jfineco.2010.03.008
[5] Whaley, R. E. (2009). Understanding the VIX. *The Journal of Portfolio Management, 35*(3), 98–105. https://doi.org/10.3905/JPM.2009.35.3.098
[6] Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies, 33*(5), 2223–2273. https://doi.org/10.1093/rfs/hhaa009
[7] Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32.
[8] Tashman, L. J. (2000). Out-of-sample tests of forecasting accuracy: An analysis and review. *International Journal of Forecasting, 16*(4), 437–450.
[9] MacKinnon, J. G., & White, H. (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics, 29*(3), 305–325.
[10] Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Belmont: Wadsworth International Group.