FINANCIAL TIME SERIES FORECASTING: A COMPREHENSIVE ANALYSIS OF UNIVARIATE NONLINEAR AND LINEAR MODELS' PERFORMANCE

Oleg Stepanov & Petr Parshakov,National Research University Higher School of Economics, Russia
Email:oleoste@yandex.ru

Abstract. Financial markets have always been attractive as a means of increasing one's wealth, and those who make accurate predictions take the prize. Forecasting models such as linear ones are simple to compute, however, they give rough approximations of the underlying relationships in the data, thus, producing poor forecasts. The solution to this issue could be the nonlinear models which try to fit the data and display the relationships with higher accuracy. Previous research seems to prove this statement from the statistician's point of view which might be of little use for an investor. Therefore, the focus of this paper is on the comparison of three types of models (nonlinear: ANN, STAR, and linear: AR) in terms of financial performance. Our research is based on the initial code for GAUSS and papers by Dick van Dijk. The data used is the monthly S&P 500 Index values from 1970 to 2012 provided by the Robert Shiller's website. Forecasting index changes begins at 1995 and ends in 2012 providing up-to-date results for 14 model specifications. The best model proves to be the flexible ANN, beating the linear AR in the majority of cases, leaving the underperforming heavy-parameterized STAR model behind. Thus, it is evident that the more flexible nonlinear models outperform the heavily parameterized ones as well as linear models for the S&P 500 Index. The introduced type of performance evaluation has a more comprehensible application to the financial market analysis.

Keywords: forecasting, nonlinear models, ANN, STAR, simulation, model performance, stock market, securities

JEL classification: G17

1 Introduction

Two previous decades have shown an increasing interest in the studies of nonlinear models. This is especially true for their application in the analysis of financial time series. It was noted in several research papers (Franses & van Dijk, 2000; Kanas & Yannopoulos, 2001) that they tend to outperform the conventional linear models. While this is generally valid for analyzing time series in experimental environments, there is still a significant degree of uncertainty in their performance in real time and market conditions. This paper focuses on the comparison of three types of models in terms of financial performance. In particular, these models are the autoregressive model (AR), the smooth transition autoregressive time series model (STAR (Chan & Tong, 1986) and the Artificial Neural Networks (ANN (Cheng and Titterington, 1994), with the last two belonging to the nonlinear models group.

Testing these models will demonstrate whether linear models make better real time forecasts than nonlinear models. Also we examine which of the two nonlinear models: the tightly parameterized STAR or the more flexible ANN – displays superior results. Additionally, forecast pooling (Timmermann, 2006) is considered in order to see if the results improve. Finally, we discuss the necessity of forecast restraints.

Another aspect of this paper is also to test the Efficient Market Hypothesis (EMH) according to the Granger and Timmermann definition (Granger & Timmermann, 2004). We will try to examine if the progressive methods outperform the traditional 'buy and hold' strategy (BH) as a part of the financial performance comparison.

2 Financial Performance

The aforementioned financial performance is used to evaluate these models from the standpoint of a stock market investor. Generating profit is possible if the model estimates the sign of the future change correctly in the majority of cases. Thus, it is necessary to use three measures of a model's performance: the success rate (Pesaran & Timmermann, 1992), the investor's final equity value and economic profit.

As it was outlined above, the model should correctly forecast the sign of the change which we will call the success ratio. It is vital to mention that the initial equity of an investor is 1 million dollars. We use the S&P 500 Index as our means of investment. The monthly data begins in January 1970 and ends in January 2012. The forecast period is from January 1995 through January 2012. In our experiment we exercised real time simulation of the stock market. This means using the information only up to the date of the forecast with the date not included. Later estimated forecasts were used to conduct operations on the market (open or hold a long or a short position, cash out). Each time the investor's equity was recalculated according to real market changes. Finally, the latest equity value was used as a measure of a model's performance.

3 Forecast restraints

It is known that there are various types of investors in terms of their maximum risk exposure. For this instance it is essential to assume that those who cannot tolerate high degrees of risk would avoid taking short positions while trading and, therefore, would sell their stocks if the forecast was negative. This rationale (Campbell & Thompson, 2007) is behind our forecast restraint that we implemented in order to see if it would improve the performance of different models.

4 Estimation

The estimation procedure included using the expanding sample for our dataset. The reason why we avoided using the rolling sample is that it demonstrates identical to poorer performance than the alternative (Franses & van Dijk, 2000). The next step was using one of the available models.

The order of the AR model (Franses & van Dijk, 2000) used was specified in the first case by the Akaike Information Criterion during in-sample testing, second was fixed at 6 and third at 12 since they represent whole and a half of a year. Also we decided to see whether a single intercept or a monthly intercept would improve the forecasts, therefore, simultaneously presenting us with 6 available forecasts.

The best parameters of the STAR model (van Dijk &Terasvirta, 2000) were estimated by using various criteria. Its maximum lag length was set to 18 months and the value of the 'long difference' used in the transition variable was 12. Since the best specification of the model was chosen automatically, we were left with one type of forecasts.

Initial tests for the specification of ANN (Franses & van Dijk, 2000) proved that the best lag lengths would be 1, 4 and 7 months. The network itself demonstrated acceptable results with 1 and 2 hidden units in the first hidden layer. As a result, we were able to retrieve 6 different forecasts for each period.

Since we had up to 6 available forecasts for each period in addition with previous forecasting history, it was possible to perform back-testing (Mamaysky, Spiegel & Zhang 2007), that is, comparing each type of a forecast to previous real changes and evaluating the current performance of a specification of a model. Then the most accurate specification and its forecast for the next month were chosen. Additionally, we implemented moving windows of 1, 5 and 15 years for the history of forecasts.

5 Bootstrap and forecast pooling

One of the most significant moments of our research as stated above is comparing three types of models. In this block bootstrapping (Bühlmann, 2002) will aid us by testing the robustness of acquired statistics for our models. If the models' performance is stable, then they could be considered reliable for usage in a broad scope of markets. The data generating process can be described by random mixing blocks of elements of the original data set. Blocks of data were used for saving the initial relationships in the data and acquiring a similar dataset.

Forecast pooling implies uniting all available forecasts into a single set and implementing back-testing to choose the best forecast among all models.

6 Results

6.1 1 year back-testing window

	No Constraint					
	Success Rate[1]	Real Equity[2]	Real BH Equity[3]	Real SR[4]	EP BH[5]	EP[6]
ANN	0,5280529	2 405 823,00	2 746 171,44	0,588235	3,17334	3,09082549
AR	0,5289854	2 437 187,22	2 746 171,44	0,590244	3,17334	3,10000532
STAR	0,5081936	792 940,22	2 746 171,44	0,502439	3,17334	-0,28571424
	Forecast Constraint					
ANN	0,5316543	2 446 236,92	2 746 171,44	0,59803922	3,17334	3,12040122
AR	0,5267141	2 348 382,37	2 746 171,44	0,5902439	3,17334	3,10320297
STAR	0,5081936	2 276 826,07	2 746 171,44	0,50243902	3,17334	3,08254644

Table 1: Model performance for 1 year back-testing window

The best model proves to be the AR, beating the ANN by only 0,2% in the success rate of forecasts for actual S&P 500 Index, making 59% overall. This results in $2,43 million dollars for the AR with the ANN behind by $2,41 million. The bootstrapped success rate goes down for both to 52,8%, so the models are not robust in their performance.

Since implementing the 'buy and hold' strategy generates $2,75 million, both models show inferior performance. Economic profit demonstrates comparable results for all models proving its close relation to investor's equity value.

The STAR model shows only a half of correct sign forecasts making the investor's assets depreciate to $800 thousand. Bootstrapped results demonstrate 50,8% success rate which is slightly better than the real one.

The implementation of forecast restraint significantly increases the ANN's performance increasing the investor's equity to $2,45 million. The same goes for the STAR model increasing the results to $2,28. Despite the growth in economic profit, AR shows a drop in equity to $2,35 million.

Forecast restraints demonstrated that the ANN model had a certain percentage of wrong negative sign forecasts which were negated by the implementation of restraints. We can also note that the AR model had a significant portion of correct negative forecasts allowing the investor to make the right decisions about taking short positions while trading.

It is also worth noting that restraints improve economic profit for all models since they reduce the standard deviation of investor's equity value which is associated with ignoring short positions.

1 Success Rate – percentage of correct sign forecasts of returns based on bootstrapping
2 Real Equity – investor's equity value based on the initial dataset forecasts
3 Real BH Equity – investor's equity value based on the initial dataset for the Buy and Hold strategy
4 Real SR (Success Rate) – percentage of correct sign forecasts of returns based on the initial dataset
5 EP BH – Economic Profit for the Buy and Hold strategy
6 EP – Economic Profit gained from a model

6.2 5 year back-testing window

	No Constraint					
	Success Rate	**Real Equity**	**Real BH Equity**	**Real SR**	**EP BH**	**EP**
ANN	0,5246711	2 609 368,24	2 746 171,44	0,671569	3,17334	3,13050201
AR	0,5356098	2 518 807,77	2 746 171,44	0,604878	3,17334	3,11077975
STAR	0,5081936	792 940,22	2 746 171,44	0,502439	3,17334	-0,2857142
	Forecast Constraint					
ANN	0,5323956	2 393 733,12	2 746 171,44	0,6372549	3,17334	3,10902961
AR	0,5363317	2 355 490,51	2 746 171,44	0,60487805	3,17334	3,10894114
STAR	0,5081936	2 276 826,07	2 746 171,44	0,50243902	3,17334	3,08254644

Table 2: Model performance for 5 year back-testing window

This case demonstrates that the ANN improved its performance. The investor's equity grew to $2,61 million with the success rate to 67%. AR made positive changes as well but they were slightly less in scale. Its equity rose to $2,52 million and the success rate only to 60% meaning only a 1% increase compared with the 1 year window. The bootstrapped success rate for both is approximately 53%.

Despite the multidirectional changes caused by restraints for the 1 year window, this time they demonstrate straightly negative effects on the investor's equity dropping them to the same level or lower.

As a result, it is evident that increasing the back-testing window improves real market forecasts.

6.3 15 year back-testing window

	No Constraint					
	Success Rate	**Real Equity**	**Real BH Equity**	**Real SR**	**EP BH**	**EP**
ANN	0,5220225	2 624 810,35	2 746 171,44	0,663415	3,17334	3,1367672
AR	0,5366907	2 496 626,84	2 746 171,44	0,595122	3,17334	3,1075997
STAR	0,5081936	792 940,22	2 746 171,44	0,502439	3,17334	-0,2857142
	Forecast Constraint					
ANN	0,5220225	2 527 457,39	2 746 171,44	0,66341463	3,17334	3,12409951
AR	0,5398088	2 381 165,36	2 746 171,44	0,59512195	3,17334	3,10533208
STAR	0,5081936	2 276 826,07	2 746 171,44	0,50243902	3,17334	3,08254644

Table 3: Model performance for 15 year back-testing window

Same results could be noticed for this window size. However, it should be outlined that while there are slight improvements for equity, it could be observed that there is a moderate decrease in the success rate by 0,5% for the AR and ANN. Such behaviour of the results should provide incentives for a slight reduction of the window size if the bootstrapped statistics are of more value than the investor's equity.

It should also be said that the STAR model does not provide us with multiple types of forecasts to choose from each period and, therefore, leaving us with the same results for 3 back-testing windows, making it the worst of the 3 models.

6.4 Forecast pooling

	No Constraint					
	Success Rate	**Real Equity**	**Real BH Equity**	**Real SR**	**EP BH**	**EP**
1	0,515282	2 354 228,65	2 746 171,44	0,643902	3,17334	3,08573331
5	0,515282	2 354 228,65	2 746 171,44	0,643902	3,17334	3,08573331
15	0,515282	2 354 228,65	2 746 171,44	0,643902	3,17334	3,08573331
	Forecast Constraint					
1	0,515282	2 362 451,97	2 746 171,44	0,64390244	3,17334	3,11078799
5	0,515282	2 362 451,97	2 746 171,44	0,64390244	3,17334	3,11078799
15	0,515282	2 362 451,97	2 746 171,44	0,64390244	3,17334	3,11078799

Table 4: Model performance for forecast pooling

The proposed simple method of uniting forecasts and choosing the best one proved to demonstrate poor performance. Despite the success rate being 64,4% which is second after ANN, the equity value is only $2,35 million being lower than the single models' performance. Bootstrapped results showed average results among all models making the success rate 51,5%.

The implementation of forecast restraints barely makes any difference only slightly increasing the equity value.

7 Summary

As a result it is possible to state that certain nonlinear models such as Artificial Neural Networks either outperform or show the same results as linear univariate models such as the Autoregressive model at different back-testing window sizes for the S&P 500 Index. The other nonlinear STAR model demonstrated the poorest results (as well as the simple forecast pooling procedure) among the three discussed models.

Forecast restraints showed their usefulness as a tool of analyzing the number of wrong negative sign forecasts in case of moderate model performance, and a bad choice for high performance models. However, they always managed to increase the economic profit for all models.

Since none of the models managed to get the same equity value as the 'Buy and Hold' strategy, the Efficient Market Hypothesis still has its power. This conclusion is not surprising since all the models are well-known to the public and the information that is used is open for everyone.

8 Acknowledgements

We are grateful to Philip Hans Franses and Dick van Dijk for providing the initial code for our research done in GAUSS.

9 References

Bühlmann P. (2002). Bootstraps for Time Series, *Statistical Science* Vol. 17, No. 1, 52-72

Campbell J., Thompson S. (2007). Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average? *The Review of Financial Studies 21(4)*: 1509-1531.

Chan K.S., and Tong H.(1986).On estimating thresholds in autoregressive models, *Journal of Time Series Analysis* 7, 179–90

Cheng B., and D.M. Titterington (1994).Neural networks: a review from a statistical perspective, *Statistical Science* 9, 2–54

Franses P.H., and van Dijk D. (2000). *Nonlinear Time Series Models in Empirical Finance.* New York: Cambridge University Press.

Granger C., Timmermann A. (2004).Efficient Market Hypothesis and Forecasting. *International Journal of Forecasting.* Volume 20, Issue 1, 15–27

Kanas A., and Yannopoulos A.(2001).Comparing linear and nonlinear forecasts for stock returns. *International Review of Economics & Finance.* Volume 10, Issue 4, 383–398.

Mamaysky H., Spiegel M., Zhang H. (2007), Improved Forecasting of Mutual Fund Alphas and Betas*, Review of Finance* 11 (3): 359-400

Pesaran M.H., and Timmermann A.(1992).A simple nonparametric test of predictive performance. *Journal of Business & Economic Statistics* 10, 461–5.

Timmermann A. (2006). Forecast Combinations, *Handbook of Economic Forecasting* Vol. 1: 135-196