Forecasting equity risk premium: is the predictive ability of technical indicators persistent?

Krisztián Szemán

2019

1 Introduction

Forecasting stock returns is an extremely difficult task that has been examined by numerous papers. Most of these studies focus on fundamental macroeconomic variables, and their ability to predict the excess return over the risk-free rate (equity risk premium) of a stock index. As Goyal and Welch [5] shows, some fundamental variables such as the treasury bill rate or dividend yields can explain in-sample future stock returns to an extent, but their out-of-sample performances are inferior to that of the historical average forecast, which is generally considered a stringent benchmark.

Neely[2] compares the performance of technical indicators and fundamental variables based on their ability to forecast the risk premium, and finds that technical indicators are predict more effectively during recessions, while models based on fundamental variables perform better during economic expansions. This implies that the predictive ability of a model based on a variable of either type is persistent over time, so assuming it performed well in the recent past, it is more likely do so in the current period. This phenomenon is called momentum of predictability (MoP) by Wang[4]. A MoP based strategy switches between the forecast of the model of interest and the historical average forecast based on the performance of the two separate models in the past k months. Wang shows that there is significant relationship between past and current forecasting ability for certain fundamental variables, such as the dividend-price ratio, the treasury bill rate of the long-term yield, which leads to improved out-of-sample predicitive ability of individuals models using MoP-based time varying forecasts.

This paper examines a naturally arising issue: is there an analogous, similarly valuable momentum of predictability for technical indicators? Using data from January 1950 to December 2017, and a total of 14 variables (6 moving average-based, 2 momentum-based, and 6 volume-based rules), I perform a test of momentum predictability in the same framework as Wang[4], and find that there is no significant dependence between past and current predictability for the indicators. Even though out of the 70 computed p-values, 7 are lower than 5%, employing a False Discovery Rate (developed by Benjamini and Hochberg) test yields that none of the findings can actually be considered true positives. (In contrast, Wang finds p-value lower than 1% for five different variables on multiple lookback periods.)

Furthermore, I also compute the monthly average signal for a higher number of moving average-based (354 different strategies, following Okunev's paper), momentum-based (36), and volume-based (354) rules. The three average signals, and a principal component extracted from the 744 variables taken together are also tested for momentum of predictability. The results are similar to the preceding section: the variables do not exhibit significant momentum of predictability.

2 Data and descriptive statistics

This section describes the variables that are used to carry out the analysis. For the equity risk premium, the log of the value-weighted equity return including dividends minus the log of the lagged risk-free rate is used (available on Amit Goyal's homepage).

The 14 technical indicators that are used are the same that Neely[2] uses, and the methodology of computing the technical indicators can also be found in that paper. Nevertheless, it is included in this section for the sake of easier understanding.

Out of the 14 variables, the first 6 are moving-average based rules, which generate a buy or sell signal ($S_{i,t} = 1$ or $S_{i,t} = 0$) at the end of period t by comparing two moving averages:

$$S_{i,t} = \begin{cases} 1 & if \quad MA_{s,t} \geq MA_{l,t} \\ 0 & \text{if} \quad \text{MA}_{s,t} < \text{MA}_{l,t}, \end{cases}$$

where

$$\text{MA}_{j,t} = \left(\frac{1}{\text{j}}\right) \sum_{i=0}^{j-1} P_{t-i},$$

P_t is the level of the stock index, s and l are the lengths of the short and long MA ($s < l$). I analyze monthly MA rules with $s = 1, 2, 3$ and $l = 9, 12$, and denote the indicators with MA lengths s and l by MA(s, l).

Secondly, 2 momentum based strategies are calculated. A momentum rule strategy guves a buy signal if the current stock price is higher than its level m periods ago:

$$S_{i,t} = \begin{cases} 1 & if \quad P_t \geq P_{t-m} \\ 0 & \text{if} \quad \text{P}_t < \text{P}_{t-m.} \end{cases}$$

I employ monthly signals for m = 9,12 and denote them MOM(m) for m.

Thirdly, 6 volume based strategies are considered, whicTherefore,h utilize trading volume as well as past price changes to identify market trends. Initially, the on-balance volume is defined as follows:

$$\mathrm{OBV}_t = \sum_{\mathrm{k=1}}^{\mathrm{t}} \mathrm{VOL}_k \mathrm{D}_k,$$

where VOL_k is the trading volume during period k, and D_k is a binary variable which takes the value 1 if $P_k \geq P_{k-1}$ and -1 otherwise. From the values ofOBV_t, the trading signal is computed in the following way:

$$S_{i,t} = \begin{cases} 1 & if \quad MA_{s,t}^{OBV} \geq MA_{l,t}^{OBV} \\ 0 & \text{if} \quad \mathrm{MA}_{s,t}^{\mathrm{OBV}} < \mathrm{MA}_{l,t}^{\mathrm{OBV}}, \end{cases}$$

where

$$\mathrm{MA}_{j,t}^{\mathrm{OBV}} = \left(\frac{1}{\mathrm{j}}\right) \sum_{i=0}^{j-1} \mathrm{OBV}_{t-i}.$$

I calculate montly volume-based rules for $s = 1, 2, 3$ and $l = 9, 12$.

The market data of the S&P500 stock index is used from January 1950 to December 2017. Due to the lack of earlier data for trading volume, we have observations for all technical indicators starting from only December 1950.

It is important to note that due to the long-term upward trend in the index level, technical indicators more often generate buy signal than not. More precisely, every indicators generates a buy signal between 67 % and 73% of the time. We can observe that the indicators are highly correlated, with minimum correlation between two indicators being 0.44, and the minimum correlation between two indicators of the same class (moving average, momentum, volume) being 0.64.

3 Methodology of testing predictability

This section describes the way of testing the out-of-sample predictive performance of a forecasting model compared to the benchmark historical average model. The methodology is also detailed by Wang[4].

Strategy	MA(1,9)	MA(1,12)	MA(2,9)	MA(2,12)	MA(3,9)
R^2_{OOS}	0.24	0.66	0.23	0.73	0.18
Strategy	MA(3,12)	MOM(9)	MOM(12)	VOL(1,9)	VOL(1,12)
R^2_{OOS}	0.01	0.14	0.16	0.13	0.49
Strategy	VOL(2,9)	VOL(2,12)	VOL(3,9)	VOL(3,12)	
R^2_{OOS}	0.34	0.40	0.09	0.57	

Table 1: $R^2_{OOS}$is positive for all variables, which implies stock return predictability. The values in the table are expressed in percentages.

The starting point is the following univariate linear regression:

$$r_{t+1} = \alpha + \beta x_t + \varepsilon_{t+1},$$

where r_t is the aforementioned equity risk premium, x_t is a vector of predictive variables (consisting of zeros and ones in this case), and ε_t is an error-term assumed to be independent and identically distributed. Following the framework of Wang, I employ a recursive, expanding estimation window to generate out-of-sample forecasts. In particular, I divide the whole sample of T observations of r_t into an in-sample part containing the first M (in this study, $M = 180$) observations, and an out-of-sample part containing the remaining $N = T - M$ observations. The out-of sample forecast for some period $K+1$, where $K \geq M$, $K < T$ is then given by

$$\hat{r}_{K+1} = \hat{\alpha}_K + \hat{\beta}_K x_K,$$

where $\hat{\alpha}_K$and $\hat{\beta}_K$ are OLS-estimates of α and β, respectively. These estimates are obtained by regressing $\{r_t\}_{t=2}^{K}$ on a constant and $\{x_t\}_{t=1}^{K-1}$. Throughout this process, a series of N out-of-sample forecasts are generated. We compare these predictions to the benchmark model by using out-of-sample R^2. This is the defined as the percentage reduction in mean squared forecast error of the model of interest relative to the benchmark model, thus theoretically, it can obtain a negative value. The R^2_{OOS} results are reported in Table 1. We can observe that the R^2_{OOS} is positive for all of the indicators, suggesting the predictability of stock returns, although a test of the significance of the difference between the models of interest and the benchmark model is not carried out.

4 Testing the momentum of predictability

Wang[4] defines momentum of predictability as the property whereby predictability present during a past period is also found in the current period. In other words, we can say that momentum of predictability is present in a model, if it is more likely to outperform the benchmark model in the current period, given that it did so in the recent past.

Current and past predictability is defined based on the squared error of the predictions. Current predictability is given by:

$$cp = \left((r_t - \hat{r}_t)^2 < (r_t - \bar{r}_t)^2 \right),$$

a logical function, which thus takes the value 1,if the condition in the parenthesis is satisfied and -1 otherwise. Similarly,

$$pp\left(k\right) = \left(\sum_{j=t-k}^{t-1} (r_t - \hat{r}_t)^2 < \sum_{j=t-k}^{t-1} (r_t - \dot{r}_t)^2 \right),$$

where k is the lookback period. $pp\left(k\right)$ takes the value 1, when the model of interest produced a smaller sum of squared error than the benchmark model throughout the past k months., and -1 otherwise. I considered lookback periods of $k = 1, 3, 6, 9, 12$. I test the momentum of predictability under the null hypothesis that $pp\left(k\right)$ and cp_t are serially independent. For this purpose, I use the Pesaran-Timmermann statistic for directional accuracy testing.

Since I found the aforementioned test to be rather inadequately documented, I shortly include its methodology. The test examines whether the sign of an n-element vector x does a significantly better job than a random forecast in oredicting the sign of an n-element vector y, First, we define the vectors

$$d_x = \begin{cases} 1 & where \quad x > 0 \\ 0 & \text{where} \quad \text{x} \leq 0 \end{cases},$$

$$d_y = \begin{cases} 1 & where \quad y > 0 \\ 0 & \text{where} \quad \text{y} \leq 0 \end{cases},$$

$$d_{xy} = \begin{cases} 1 & where \quad xy > 0 \\ 0 & \text{where} \quad \text{xy} \leq 0, \end{cases}$$

where xy is the vector which is the elementwise product of x and y. For each of these three vectors, let p_x, p_y, and p_{xy}be the proportion of ones in d_x, d_y and d_{xy} respectively, and

$$q_i = \frac{p_i\left(1-p_i\right)}{n},$$

for $i = x, y, xy$. Now

$$p_* = p_x p_y + \left(1-p_x\right)\left(1-p_y\right),$$

$$p_{xy}^{VAR} = \frac{p_*\left(1-p_*\right)}{n},$$

$$p_*^{VAR} = \left(2p_y-1\right)^2 q_x + \left(2p_x-1\right)^2 q_y + 4q_x q_y,$$

where p^* can be interpreted as the expected proportion of correct guesses of sign change assuming serial independence. The test statistic is given by the formula

$$PT = \frac{p_{xy} - p_*}{\sqrt{p_{xy}^{VAR} - p_*^{VAR}}},$$

which follows a standard normal distribution under the null hypothesis. The test results with the p-values are reported in Table 2.

I find that 8 of the employed 14 variables has no significant momentum of predictability for any lookback period. The moving average based indicators MA(1,9), MA(1,12), MA(2,12), the momentum-based MOM(9) and the volume-based VOL(1,9) and VOL(2,9) show somewhat significant momentum of predictability (on a significance level of 5%) with one of the longer lookback periods ($k = 6, 9$ or 12), which. Moreover, none of the 70 p-values of Table3 is less than 1 percent, which suggests that there is no momentum of predictability for these 14 technical indicators.

To more formally conclude that there is no signficant momentum of predictability present, I used False Discovery Rate controlling (by Benjamini Hochberg[1]), more precisely the Benjamini–Hochberg–Yekutieli procedure. In this test, we rearrange the computed n p-values in ascending order: $P_{(1)} \leq P_{(2)} \leq \ldots \leq P_{(n)}$, and look for the largest k that satisfies $P_{(k)} \leq \frac{k}{n \cdot c(n)} q^*$, where $q^*$is the control parameter, the expected proportion of false discoveries, and $c\left(n\right) = 1$, since the tests in this case are positively correlated. With $q^* = 5\%$, the test yields that none of the discoveries are true positive.

k	MA(1,9)	MA(1,12)	MA(2,9)	MA(2,12)	MA(3,9)
1	0.540	0.669	0.753	0.947	0.654
3	0.508	0.364	0.489	0.565	0.428
6	0.066	0.137	0.036	0.314	0.080
9	0.278	0.287	0.119	0.458	0.121
12	0.013	0.018	0.148	0.366	0.068

k	MA(3,12)	MOM(9)	MOM(12)	VOL(1,9)	VOL(1,12)
1	0.861	0.736	0.575	0.515	0.419
3	0.330	0.291	0.329	0.145	0.363
6	0.104	0.044	0.096	0.036	0.228
9	0.158	0.389	0.341	0.389	0.194
12	0.260	0.137	0.839	0.029	0.092

k	VOL(2,9)	VOL(2,12)	VOL(3,9)	VOL(3,12)
1	0.452	0.812	0.389	0.392
3	0.115	0.373	0.141	0.347
6	0.185	0.338	0.209	0.121
9	0.021	0.454	0.148	0.311
12	0.105	0.650	0.246	0.456

Table 2: Results of testing the MoP. The p-values are much higher than in Wang's similar table, indicating there is no significant MoP for technical indicators

5 MoP of different groups of technical indicators

In the previous section, I tested 14 technical indicators, and concluded that they do not exhibit any significant momentum of predictability. Some papers, on the other hand, use a higher number of technical variables to examine their performance (e. g. Okunev[3], using 354 moving average rules on the foreign exchange market)-

Since different classes (moving-average, momentum, volume-based) of technical strategies tend to give more homogenous signals, I also analyse the performance of the average MA-based, momentum-based, and volume based strategies.

First, following Okunev[3], I compute monthly signals for 354 moving-average based strategies with $s = 1, 2, ..., 12$ and $l = 2, 3, ..., 36$ with $s < l$, meaning

$$\mathrm{MA}_t^{\mathrm{AVG}} = \frac{\sum_{s=1}^{12} \sum_{l=s+1}^{36} I\left(\mathrm{MA}_{s,t} \geq \mathrm{MA}_{l,t}\right)}{354},$$

k	PC^{TECH}	MA^{AVG}	MOM^{AVG}	VOL^{AVG}
1	0.327	0.389	0.789	0.158
3	0.203	0.092	0.168	0.100
6	0.130	0.099	0.271	0.010
9	0.022	0.107	0.121	0.074
12	0.268	0.331	0.380	0.176

Table 3: p-values for serial dependency test current and past performance of PC^{TECH},MA^{AVG}, MOM^{AVG},and VOL^{AVG}strategies

where I is the indicator function which takes the value 1 is the condition in the parenthesis is satisfied and 0 otherwise.

Secondly, the average signal given by momentum based strategies is computed using MOM(m) for $m = 1, 2, ..., 36$:

$$\mathrm{MOM}_t^{\mathrm{AVG}} = \sum_{m=1}^{36} I\left(P_t \geq P_{t-m}\right).$$

Lastly, the average of the volume-based strategies is computed for the same short and long moving-averages as for simple MA-based strategies, meaning

$$\mathrm{VOL}_t^{\mathrm{AVG}} = \frac{\sum_{s=1}^{12} \sum_{l=s+1}^{36} I\left(\mathrm{MA}_{s,t}^{\mathrm{OBV}} \geq \mathrm{MA}_{l,t}^{\mathrm{OBV}}\right)}{354}$$

Furthermore, since all of them trying to extract the trend from price movements, technical indicators are highly correlated. Neely found that in the principal component analysis of the the14 technical indicators that paper employed, the first component explains over 70% of the variance. Considering this, I also analyse the momentum of predictability of the first principal component extracted from the PCA for the previously constructed 744 strategies. This component, further denoted by PC^{TECH}still manages to explain more than 55% of the variance in spite of the magnitudes larger number of variables.

After constructing these variables, I perform the same PT-test to test the momentum of predictability, the results of which are reported in Table 3.

For the $PC^{TECH}$and VOL^{AVG}strategies, there is significant MoP on an 5% level for one lookback period, namely 9 months and 6 months, respectively. Considering the MA^{AVG} and MOM^{AVG}strategies, there is no significant MoP for any lookback period. The use of the FDR-test with $q^* = 5\%$ again yields that all of our discoveries can be claimed false positives.

6 Reasons for the non-existence of MoP

All things considered, it can be concluded that technical indicators do not exhibit the same momentum of predictability as some fundamental variables do. This can be a bit counter-intuitive at first, as Neely[2] showed that forecasts based on technical indicators tend to be stronger in economic recessions, and weaker in expansions, while fundamental variables tend to display an opposite behaviour, which could imply that the forecasting ability of the technical indicators is just as persistent as the same property of fundamentals.

On the other hand, only some of the fundamental variables analysed by Wang[4] have significant momentum of predictability, and some do not admit this property. Focusing on technical variables whose correlation structure is much stronger than that of the fundamentals, it is unlikely to conclude a similar result: if some of them do not exhibit MoP, then it is highly likely that none of them does. Considering this, the findings about the non-presence of the MoP is not that surprising.

7 Conclusion

In this paper, I examined the persistence of forecasting ability, also known as momentum of predictability for technical indicators in predicting the equity risk premium. Namely, I tested the MoP for 14 technical indicators used by Neely[2], then 3 technical rules based on averaging out 354 moving-average, 36 momentum-based and 354 volume-based indicators, respectively, and a principal component which catches the average of the signals given by the mentioned indicators. In contrast to Wang[4], who found that the predictive ability of certain fundamental variables is persisent, I found that the same is not true for technical indicators, as there is no significant dependence between predictive ability of models based on the technical indicators in the recent past and in the present.

References

[1] Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)*, pages 289–300, 1995.

[2] Christopher J Neely, David E Rapach, Jun Tu, and Guofu Zhou. Forecasting the equity risk premium: the role of technical indicators. *Management Science*, 60(7):1772–1791, 2014.

[3] John Okunev and Derek White. Do momentum-based strategies still work in foreign currency markets? *Journal of Financial and Quantitative Analysis*, 38(2):425–447, 2003.

[4] Yudong Wang, Li Liu, Feng Ma, and Xundi Diao. Momentum of return predictability. *Journal of Empirical Finance*, 45:141–156, jan 2018.

[5] Ivo Welch and Amit Goyal. A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies*, 21(4):1455–1508, 2007.