Economic Modelling 78 (2019) 32–39

Contents lists available at ScienceDirect

Economic Modelling

journal homepage: www.journals.elsevier.com/economic-modelling

Forecasting stock returns: Do less powerful predictors help?

Yaojie Zhang [a], Qing Zeng [b], Feng Ma [a,*], Benshan Shi [a]

[a] School of Economics and Management, Southwest Jiaotong University, Chengdu, China
[b] College of Economics, Sichuan Agricultural University, Chengdu, China

ARTICLE INFO

JEL classification:
C53
G11
G17

Keywords:
Stock return predictability
Multivariate regression model
Complementary information
Combination forecasts
Monte Carlo simulation

ABSTRACT

This paper proposes a simple but efficient way to improve the predictability of stock returns. Instead of torturously constructing new powerful predictors, we readily select existing predictors that have low correlations and thus provide complementary information. Our forecasting strategy is to use the selected predictors based on a multivariate regression model. In our forecasting strategy, less powerful predictors are also useful for forecasting stock returns if they could provide complementary information. The empirical results show that our forecasting strategy outperforms not only the univariate regression models that use each predictor's information separately but also combination approaches that use all predictors jointly. We also document that our strategy extracts significantly more useful information from the complementary predictors than the competing models. In addition, from an asset allocation perspective, a mean-variance investor realizes substantial economic gains. Furthermore, the evidence based on Monte Carlo simulations supports the feasibility of our forecasting strategy.

1. Introduction

Predicting stock returns is a subject of great interest to both practitioners and academics in finance; however, it is notoriously difficult. The seminal paper of Welch and Goyal (2008) documents that most prevailing macroeconomic predictors fail to generate consistently superior out-of-sample forecasts of aggregate stock returns relative to the simple benchmark forecast of historical average.

To beat the historical average benchmark and improve stock return predictability, a growing number of studies have constructed many new and powerful predictors or factors, which include the variance risk premium (Bollerslev et al., 2009, 2014), technical indicators (Lin, 2018; Neely et al., 2014), the short interest index (Rapach et al., 2016), investor sentiment (Huang et al., 2015; Li et al., 2017; Ni et al., 2015; Yang et al., 2013), financial news (Narayan and Bannigidadmath, 2017; Narayan et al., 2017b), stock return autocorrelations (Xue and Zhang, 2017), news-implied volatility (Manela and Moreira, 2017), credit quality (Narayan et al., 2017a), and manager sentiment (Jiang et al., 2017), among others.

However, it is well known that constructing new and powerful predictors for stock returns is a daunting task. For this consideration, our paper's main purpose is to propose a simple but efficient method to achieve superior out-of-sample forecasting performance without constructing new predictors. Specifically, we seek existing predictors that can provide complementary information to predict stock returns and then place them into a multivariate predictive regression model to generate stock return forecasts. Our motivation is straightforward. When two or more predictors contain complementary information that is useful for predicting future stock returns, we will lose necessary information using a univariate regression model based on an individual predictor. Most of the related literature on stock return predictability prefers to use a univariate regression model for each predictor separately. This is because useful predictors tend to have high correlations, thus providing little complementary information and resulting in a multicollinearity issue. Hence, the key to our forecasting strategy is seeking useful predictors that have low correlations in order to obtain complementary information.

In the empirical analysis, we find that the short interest index (SII) and aligned investor sentiment (AIS) capture complementary information for predicting stock returns, as they show a low correlation. Compared with AIS, SII is a less powerful predictor. Nonetheless, SII can complement AIS and further enhance out-of-sample forecasting performance. Specifically, the bivariate regression model based on AIS and SII generates substantially larger out-of-sample R-squares and economic values than the competing models, including the benchmark model of the historical average, the univariate predictive regression model based on

* Corresponding author. School of Economics and Management, Southwest Jiaotong University, No. 111, North 1st Section, 2nd Ring Road, Chengdu, 610031, China.

E-mail addresses: yaojie_zhang@126.com (Y. Zhang), zengqing2017@126.com (Q. Zeng), mafeng2016@swjtu.edu.cn (F. Ma), bsshi@swjtu.cn (B. Shi).

https://doi.org/10.1016/j.econmod.2018.09.014
Received 4 February 2018; Received in revised form 7 July 2018; Accepted 9 September 2018
Available online 11 September 2018

individual AIS or SII predictors, and the prevailing combination approaches used by Rapach et al. (2010). In particular, combination approaches are very similar to our forecast strategy, as both use all the predictors jointly. Although combination forecasts are more accurate than individual forecasts, our forecasting strategy further surpasses combination approaches when using predictors with low correlations.

Our empirical findings are found to be robust to various settings, including business cycles, different forecasting windows, and alternative investment environments. In addition, we observe a robust result for more predictors including AIS, SII, and two technical indicators from Neely et al. (2014), further supporting the effectiveness of our forecasting strategy.

Furthermore, the results of forecast encompassing tests show that the forecasts based on the bivariate predictive regression model contain significantly more powerful predictive information for forecasting future stock returns than either of the competing forecasts. This supports the motivation of our forecasting strategy. That is, our strategy can extract more useful information contained in individual predictors.

Another main contribution of this paper is that we provide evidence from Monte Carlo simulations. More importantly, the simulation evidence provides reasons why combination approaches work well in the setting of Rapach et al. (2010) and why our method outperforms the competing models in this study. The first simulation case suggests that combination approaches yield the best out-of-sample forecasting performance in the presence of multicollinearity. This finding is consistent with the empirical findings of Rapach et al. (2010). Combination approaches can reduce forecast variance, whereas the multivariate regression model fails to extract useful information due to the information overlap. In sharp contrast, the second simulation case shows that our forecasting strategy (i.e., the multivariate regression model) outperforms combination approaches when the predictors used have low correlations. This is because the multivariate regression model can extract more useful information than combination approaches when the predictors' information is complementary. Based on the second case, we further add a less powerful predictor with a lower correlation relative to the existing predictors in the third simulation case. The less powerful predictor further helps to improve the forecasting performance of the multivariate regression model, while we observe a decline in forecasting performance for combination approaches. This evidence documents the feasibility of our forecasting strategy. Of course, the simulation evidence also reduces data mining concerns for our empirical analysis.

Finally, we provide some practical implications that are related to this study. In the real world, investors are unlikely to depend on a single predictor to forecast future stock returns. Moreover, with the development of technology, we can collect increasing amounts of data and information. According to our paper's findings, combination approaches are good choices in this data-rich world. However, when investors have useful predictors that can provide complementary information, we recommend using a multivariate regression model to forecast stock returns and guide asset allocation. In doing so, investors can earn more money and assume a lower risk.

The remainder of the paper is organized as follows. Section 2 provides the methodologies, including forecasting models and evaluation methods. Section 3 describes our data. Section 4 reports the results of out-of-sample forecasting performance and conducts a series of robustness checks. Section 5 provides the evidence from Monte Carlo simulations. Finally, Section 6 concludes.

2. Methodology

2.1. Forecasting models

First of all, we run the following univariate predictive regression to predict stock returns.

$$r_{t+1} = \alpha_i + \beta_i x_{i,t} + \varepsilon_{i,t+1}, \tag{1}$$

where r_{t+1} is the S&P 500 log excess return for month t+1, $x_{i,t}$ is the ith predictor available at t, and $\varepsilon_{i,}$ is an error term whose mean is equal to zero.

Combination forecasts are computed as weighted averages of the N individual forecasts based on Eq. (1). Statistically, combination forecasts are given by

$$\widehat{r}_{c,t+1} = \sum_{i=1}^{N} \omega_{i,t} \widehat{r}_{i,t+1}, \tag{2}$$

where $\widehat{r}_{c,t+1}$ is the combination forecast at month t+1, $\widehat{r}_{i,t+1}$ is the ith individual forecast, and $\omega_{i,t}$ represents the combining weight of the ith individual forecast calculated at month t.

Following Rapach et al. (2010) and Zhu and Zhu (2013), we consider five popular combination approaches: mean, median, trimmed mean, DMSPE (1), and DMSPE (0.9). The mean combination forecast takes the mean of the N individual forecasts, $\{\widehat{r}_{i,t+1}\}_{i=1}^{N}$. The median combination forecast takes the median of $\{\widehat{r}_{i,t+1}\}_{i=1}^{N}$. The trimmed mean combination forecast discards the smallest and largest individual forecasts in $\{\widehat{r}_{i,t+1}\}_{i=1}^{N}$ and sets $\omega_{i,t} = 1/(N-2)$ for the remainder of the individual forecasts. In the discount mean squared prediction error (DMSPE) combining method, the combining weights of model i at month t are expressed as

$$\omega_{i,t} = \varphi_{i,t}^{-1} \Big/ \sum_{\ell=1}^{N} \varphi_{\ell,t}^{-1}, \tag{3}$$

where

$$\varphi_{i,t} = \sum_{s=m+1}^{t} \theta^{t-s} \left(r_s - \widehat{r}_{i,s} \right)^2, \tag{4}$$

m is the length of the initial training sample period and θ is a discount factor. Following Rapach et al. (2010), Zhu and Zhu (2013), and Zhang et al. (2018), we consider two values of θ, namely, 1 and 0.9. Consequently, two DMSPE methods, DMSPE (1) and DMSPE (0.9), are used in this study. Similar to our forecasting strategy below, combination approaches are very popular for forecasting financial assets with many predictors. Furthermore, a host of related literature documents that combination forecasts significantly outperform individual forecasts (see, e.g., Ma et al., 2018a; Rapach et al., 2010; Zhu and Zhu, 2013). Given this, we treat combination approaches as important competing models in this study.

In contrast to combination approaches, this study employs a multivariate predictive regression model,

$$r_{t+1} = \alpha + \sum_{i=1}^{N} \beta_i x_{i,t} + \varepsilon_{t+1}, \tag{5}$$

which can be regarded as a kitchen sink model to predict future stock returns. The motivation for using a multivariate predictive regression model is that it can extract more useful information than combination approaches when predictors can provide complementary information. In other words, our main goal is to seek new or existing predictors with low correlations. In the following sections, we provide not only empirical evidence but also simulation evidence to support the feasibility of our method. Note that Eq. (5) turns into a bivariate regression model when we use two predictors in the empirical analysis.

2.2. Out-of-sample evaluation

Following the convention in return forecasting (see, e.g., Campbell and Thompson, 2008; Huang et al., 2015; Jiang et al., 2017; Neely et al., 2014; Pettenuzzo et al., 2014; Rapach et al., 2016; Rapach et al., 2010; Zhu, 2013), we use the out-of-sample R^2 statistic to evaluate the out-of-sample predictive performance of a forecasting model relative to the benchmark of historical average. The out-of-sample R^2 statistic is defined as

$$R_{OS}^2 = 1 - \frac{\sum_{k=1}^{q}\left(r_{m+k} - \hat{r}_{m+k}\right)^2}{\sum_{k=1}^{q}\left(r_{m+k} - \bar{r}_{m+k}\right)^2}, \tag{6}$$

where r_{m+k}, $\bar{r}_{m+k}$, and $\hat{r}_{m+k}$ are the actual return, historical average, and return forecast, respectively, at month $m+k$, and m and q denote the lengths of the initial estimation period and forecast evaluation period, respectively.

The R_{OS}^2 statistic measures the reduction in mean squared forecast error (MSFE) for the return forecast relative to the historical average benchmark. To further ascertain whether a forecasting model yields a statistically significant improvement in MSFE, the Clark and West (2007) statistic is employed. More specifically, the Clark and West (2007) statistic tests the null hypothesis that the MSFE of the historical average benchmark is smaller than or equal to the MSFE of the forecasting model of interest against the alternative hypothesis that the MSFE of the historical average benchmark is larger than the MSFE of the forecasting model of interest. Mathematically, the Clark and West (2007) statistic is computed by first defining

$$f_t = (r_t - \bar{r}_t)^2 - \left(r_t - \hat{r}_t\right)^2 + \left(\bar{r}_t - \hat{r}_t\right)^2, \tag{7}$$

where r_t, $\bar{r}_t$, and $\hat{r}_t$ are the actual stock return, the simple mean benchmark forecast of stock return, and the stock return forecast based on the forecasting model of interest, respectively. By regressing $\{f_s\}_{s=m+1}^{T}$ on a constant, we can conveniently derive the Clark and West (2007) statistic, which is just the t-statistic of the constant. Moreover, a p-value for the one-sided (upper-tail) test is conveniently derived with the standard normal distribution.

Following Campbell and Thompson (2008), Rapach et al. (2010), Neely et al. (2014), Rapach et al. (2016), and Jiang et al. (2017), among others, we further measure the economic value of various stock return forecasts from an asset allocation perspective. More specifically, we calculate the certainty equivalent return (CER) for a mean-variance investor who allocates between stocks and risk-free bills using various forecasts of stock returns. In order to achieve the maximum CER, the investor would allocate the weight of stocks during month $t+1$ as

$$w_t = \frac{1}{\gamma}\frac{\hat{r}_{t+1}}{\hat{\sigma}_{t+1}^2}, \tag{8}$$

where γ is the investor's risk aversion coefficient, $\hat{r}_{t+1}$ denotes a return forecast, and $\hat{\sigma}_{t+1}^2$ denotes a forecast of the stock return variance. As in Campbell and Thompson (2008), Rapach et al. (2010), Neely et al. (2014), and Jiang et al. (2017), among others, we estimate the variance forecasts using a five-year moving window of past stock returns and restrict w_t to the range between 0 and 1.5 to preclude short sales and to allow no more than 50% leverage.[1]

For a portfolio constructed by Eq. (8), the investor can realize an average CER as

$$CER = \bar{R}_p - 0.5\gamma\sigma_p^2, \tag{9}$$

where $\bar{R}_p$ and σ_p^2 denote the mean and variance, respectively, of the realized portfolio returns during the out-of-sample evaluation period. The CER gain is calculated as the difference between the CER for the investor when she uses a return forecast and the CER when she uses the prevailing mean forecasts. Accordingly, the CER gain can be regarded as the portfolio management fee that a mean-variance investor would be pleased to pay to have access to the return forecasts in instead of the historical average forecasts.

3. Data

The short interest index (SII), proposed by Rapach et al. (2016), and aligned investor sentiment (AIS), constructed by Huang et al. (2015), are two suitable predictors that strongly satisfy our requirements.[2] SII and AIS not only show a relatively strong predictive ability but also have a low correlation.[3] The sample period spans from January 1973 through December 2014. The initial estimation period is 1973:01–1990:12, and therefore, our out-of-sample period is 1991:01–2014:12. Following the literature on return predictability (see, e.g., Huang et al., 2015; Jiang et al., 2017; Neely et al., 2014; Rapach et al., 2016; Rapach et al., 2010), we generate out-of-sample forecasts of stock returns using a recursive (expanding) estimation window.

Table 1 presents the descriptive statistics of excess stock returns, AIS, and SII, as well as the correlation matrix among excess stock return, AIS with a one-month lag, and SII with a one-month lag. We find that the average stock return is positive over the entire sample period. Stock return and AIS exhibit significantly non-normal distributions. In addition, the correlation coefficients between the dependent variable (i.e., stock return) and independent variables (i.e., lagged SII and lagged AIS) are relatively large and significant at the 1% level, while the correlation coefficient of the two independent variables is relatively small and significant at a lower level of 5%. According to the correlation matrix, we can infer that the individual in-sample R-square values of AIS and SII are 1.52% and 1.74%, respectively, which are rather large in the literature on stock return predictability (see, e.g., Neely et al., 2014; Welch and Goyal, 2008).

Table 1
Descriptive statistics and correlation matrix.

Variables	Mean	Std. Dev.	Skewness	Kurtosis	Jarque-Bera
Panel A: Descriptive statistics					
Stock return	0.004	0.045	−0.718	2.707	192.580***
AIS	−0.056	0.890	1.357	1.682	211.054***
SII	−0.039	0.226	0.165	0.202	3.027
Panel B: Correlation matrix					
Variables	Stock return	Lagged AIS	Lagged SII		
Stock return	1.000	−0.123***	−0.132***		
Lagged AIS	−0.123***	1.000	−0.101**		
Lagged SII	−0.132***	−0.101**	1.000		

This table reports descriptive statistics and correlation coefficients for stock returns, aligned investor sentiment (AIS), and the short interest index (SII). The Jarque-Bera statistic tests for the null hypothesis of normal distribution. *** and ** indicate significance at the 1% and 5% levels, respectively.

[1] Alternatively, Rapach et al. (2016) uses a ten-year moving window to estimate the return volatility. The economic value results are qualitatively similar for a ten-year moving window. To save space, we do not report these results, but they are available upon request.

[2] We thank Guofu Zhou for sharing the raw data, which are available on his webpage at http://apps.olin.wustl.edu/faculty/zhou/.

[3] In this paper, we do not provide a standard rule to choose our desired predictors or provide a reasonable threshold value to define "low correlation". This is because we can accept a higher correlation when having stronger predictors. Furthermore, it is unlikely to derive an explicit function that can accurately describe the relationship between predictors' predictive ability and correlation. However, according to our empirical evidence below, we can provide a general guidance that when the predictors' in-sample R-squares (which can be viewed as a measure of predictive ability) are greater than 2%, the correlation value of 0.1 can be regarded as a low correlation.

4. Empirical results

4.1. Out-of-sample forecasting performance

Table 2 presents both the statistical and economic out-of-sample performance. We can see that SII is a less powerful predictor relative to AIS, as SII yields smaller values of R^2_{OS}, CER gain, and Sharpe ratio than AIS. More importantly, the bivariate regression model with SII and AIS exhibits a considerable improvement in out-of-sample performance relative to their univariate regression models.[4] Intuitively, the economic sources of the return predictability of the bivariate regression model result predominantly from the complementary information between SII and AIS. That is, SII and AIS can separately capture different information content and thus provide additional useful information.

Rapach et al. (2010) show that combination forecasts also incorporate meaningful information from all of the potential predictors. In this setting, however, combination approaches yield worse out-of-sample performance than the bivariate regression model, suggesting that combination forecasts may discard certain portions of useful information from SII and AIS. Finally, it should be noted that median forecasts are identical to mean forecasts and that trimmed mean forecasts do not exist when there are only two individual forecasts in our empirical case. The corresponding results are thus not reported.

4.2. Forecast encompassing tests

Forecast encompassing tests are widely used to assess the relative information content in the stock return forecasts (see, e.g., Huang et al., 2015; Jiang et al., 2017; Neely et al., 2014; Rapach et al., 2016; Rapach et al., 2010). To further compare the information content of the forecasts based on the bivariate predictive regression model to that of the forecasts based on the competing models, we conduct forecast encompassing tests. More specifically, we first generate an optimal combination forecast of actual return r_t as a convex combination of the forecast $\widehat{r}_{j,t+1}$ based on one of the competing models and the forecast $\widehat{r}_{b,t+1}$ based on the bivariate predictive regression model,

$$r_{t+1} = (1-\delta)\widehat{r}_{j,t+1} + \delta\widehat{r}_{b,t+1}, \tag{10}$$

where j indexes the competing models and b indexes the bivariate predictive regression model. If δ (1 - δ) = 0, the model j (b) forecast encompasses the model b (j) forecast, because model b (j) does not provide any useful information for forming the optimal combination forecast beyond the information already contained in model j (b). Alternatively, if δ (1 - δ) > 0, the model j (b) forecast does not encompass the model b (j) forecast. That is, model b (j) does provide information that is useful for forecasting stock returns beyond the information already contained in model j (b). Harvey et al. (1998) propose a statistic for testing the null hypothesis that the model j (b) forecast encompasses the model b (j) forecast (H_0: δ (1 - δ) = 0) against the alternative hypothesis that the model j (b) forecast does not encompass the model b (j) forecast (H_1: δ (1 - δ) > 0).

Table 3 reports the p-values of forecast encompassing tests. In either case, the combination weights δ of the forecasts based on the bivariate regression of AIS and SII are significantly larger than zero, suggesting that the forecasts based on the bivariate regression of AIS and SII contain useful information for forecasting stock returns beyond the relevant information already contained in those competing forecasts. In contrast, we cannot reject the null hypothesis that (1 - δ) = 0, indicating that the forecasts based on the bivariate regression of AIS and SII encompass the competing forecasts. In other words, the information of the forecasts based on the competing models has already been contained in the forecasts based on the bivariate regression of AIS and SII. In conclusion, the results of the forecast encompassing tests show that the forecasts based on the bivariate predictive regression model exhibit significantly more powerful predictive information for forecasting future stock returns than either of the competing forecasts. This finding implies that the bivariate predictive regression model is the most efficient way to capture the complementary information between AIS and SII.

4.3. Robustness checks

4.3.1. Forecasting performance over the business cycles

From an economic point of view, although the overall R^2_{OS} is interesting, it is also important to analyze the return predictability during business cycles. Following related studies (see, e.g., Huang et al., 2015; Jiang et al., 2017; Ma et al., 2018b; Neely et al., 2014; Rapach et al., 2010; Wang et al., 2018), we compute the R^2_{OS} statistic separately for expansions ($R^2_{OS,EXP}$) and recessions ($R^2_{OS,REC}$),

$$R^2_{OS,c} = 1 - \frac{\sum_{k=1}^{q} I^c_{m+k}\left(r_{m+k} - \widehat{r}_{m+k}\right)^2}{\sum_{k=1}^{q} I^c_{m+k}\left(r_{m+k} - \overline{r}_{m+k}\right)^2} \text{for } c = EXP, REC, \tag{11}$$

Table 2
Statistical and economic out-of-sample performance.

Forecasting models	MSFE	R^2_{OS} (%)	CER gains (%)	Sharpe ratios
AIS	17.25	2.43***	5.12	0.21
SII	17.37	1.76***	5.37	0.19
AIS and SII	16.90	4.43***	8.50	0.27
Mean	17.19	2.76***	6.69	0.24
DMSPE (1)	17.19	2.78***	6.70	0.24
DMSPE (0.9)	17.20	2.73***	6.67	0.24

This table reports the out-of-sample performance from both statistical and economic perspectives. The forecasting models include two univariate regression models based on aligned investor sentiment (AIS) and the short interest index (SII), a bivariate regression model of both AIS and SII, and three combination approaches of mean combination, DMSPE(1), and DMSPE(0.9). Statistical performance includes mean squared forecast error (MSFE) and out-of-sample R-square (R^2_{OS}). Statistical significance for the R^2_{OS} statistic is derived by using the Clark and West (2007) test. *** indicates significance at the 1% level. With respect to economic performance, the annualized certainty equivalent return (CER) gains and Sharpe ratios are calculated based on a mean-variance investor with relative risk aversion coefficient of three who allocates between stocks and risk-free bills monthly. The initial estimation period is 1973:01–1990:12, while the out-of-sample period is 1991:01–2014:12.

Table 3
Forecast encompassing test results.

Forecasting models	δ	1 - δ
AIS	0.002	0.233
SII	0.008	0.817
Mean	0.012	0.510
DMSPE (1)	0.013	0.503
DMSPE (0.9)	0.012	0.525

The table reports p-values for the Harvey et al. (1998) MHLN statistic for testing the null hypothesis that δ (1 - δ) = 0 against the alternative hypothesis that δ (1 - δ) > 0. δ (1 - δ) represents the combination weight of the bivariate regression forecasts based on SSI and AIS (the competing forecasts). The forecasting models are two univariate regression models based on aligned investor sentiment (AIS) and the short interest index (SII) and three combination approaches of mean combination, DMSPE(1), and DMSPE(0.9).

[4] If we use the univariate regression models to replace the historical average benchmark in Eq. (6), the R^2_{OS}s of the bivariate regression model relative to the AIS and SII univariate regression models are 2.05% and 2.71%, respectively. They are both significant at the 1% level. Furthermore, the R^2_{OS}s of the bivariate regression model relative to the combination approaches are also significantly positive.

where I_{m+k}^{EXP} (I_{m+k}^{REC}) is an indicator that takes a value of one when month $m+k$ is in an NBER expansion (recession) period and zero otherwise. Similarly, we can compute the CER gains and Sharpe ratios separately for expansions and recessions.

Table 4 reports the out-of-sample forecasting performance over business cycles. Two observations follow the table immediately. First, consistent with the related literature on return predictability (see, e.g., Huang et al., 2015; Jiang et al., 2017; Neely et al., 2014; Rapach et al., 2010), the return predictability is concentrated over recessions for all the forecasting models. Second and more importantly, the bivariate regression model of AIS and SII exhibits more forecasting gains than the competing models over recessions. The results reported in Table 4 also suggest that the previous results are robust to alternative business cycles.

4.3.2. Alternative forecasting windows

Rossi and Inoue (2012) emphasize that the arbitrary choices of various window sizes may result in quite different out-of-sample predictive performances in practical applications. Therefore, the choice of forecasting window sizes plays an important role in out-of-sample evaluation. Thus, we additionally consider another two forecasting windows, where the length of the initial estimation windows is 15 years and 20 years. As a result, all three forecasting windows considered in this paper have a desirable trade-off between an initial estimation period that has enough in-sample observations to precisely estimate parameters and an out-of-sample period that is relatively long for forecast evaluation.

Table 5 reports the out-of-sample forecasting performance for alternative forecasting windows. We observe a robust result that the bivariate regression model based on AIS and SII shows the best out-of-sample performance from both statistical and economic perspectives. In other words, the bivariate regression model can effectively capture the complementary information between SII and AIS under alternative forecasting windows.

4.3.3. Transaction cost and alternative risk aversion choices

Our main analysis of the economic value of stock return forecasts is based on the assumption of the mean-variance investor with a risk aversion coefficient of three who allocates between stocks and risk-free bills without any transaction cost. As a robustness check, we additionally consider not only other reasonable values of risk aversion coefficients but also a proportional transaction cost of 50 basis points per transaction.

Table 4
Out-of-sample forecasting performance over business cycles.

Forecasting models	Recession periods			Expansion periods		
	R^2_{OS} (%)	CER gains (%)	Sharpe ratios	R^2_{OS} (%)	CER gains (%)	Sharpe ratios
AIS	2.03**	3.04	0.25	3.54*	22.93	−0.12
SII	1.17**	2.90	0.22	3.43*	26.53	−0.04
AIS and SII	2.91***	4.99	0.27	8.68**	38.86	0.37
Mean	2.40***	3.79	0.25	3.79**	31.73	0.01
DMSPE (1)	2.39***	3.79	0.25	3.87**	31.84	0.02
DMSPE (0.9)	2.34***	3.74	0.25	3.83**	31.95	0.02

This table reports the out-of-sample performance from both statistical and economic perspectives over business cycles. The forecasting models include two univariate regression models based on aligned investor sentiment (AIS) and the short interest index (SII), a bivariate regression model of both AIS and SII, and three combination approaches of mean combination, DMSPE(1), and DMSPE(0.9). Statistical performance includes mean squared forecast error (MSFE) and out-of-sample R-square (R^2_{OS}). Statistical significance for the R^2_{OS} statistic is derived by using the Clark and West (2007) test. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. With respect to economic performance, the annualized certainty equivalent return (CER) gains and Sharpe ratios are calculated based on a mean-variance investor with relative risk aversion coefficient of three who allocates monthly between stocks and risk-free bills. The initial estimation period is 1973:01–1990:12, while the out-of-sample period is 1991:01–2014:12.

Table 5
Statistical and economic out-of-sample performance for alternative forecasting windows.

Forecasting models	MSFE	R^2_{OS} (%)	CER gains (%)	Sharpe ratios
Panel A: The length of the initial estimation period is 15 years				
AIS	17.13	2.25***	4.72	0.20
SII	17.32	1.16**	4.51	0.18
AIS and SII	16.86	3.79***	7.78	0.26
Mean	17.12	2.29***	5.64	0.22
DMSPE (1)	17.12	2.31***	5.64	0.22
DMSPE (0.9)	17.13	2.26***	5.62	0.21
Panel B: The length of the initial estimation period is 20 years				
AIS	17.74	2.36***	5.28	0.20
SII	17.79	2.06***	5.94	0.20
AIS and SII	17.34	4.56***	9.00	0.27
Mean	17.64	2.90***	7.25	0.24
DMSPE (1)	17.64	2.89***	7.22	0.24
DMSPE (0.9)	17.65	2.85***	7.19	0.24

This table reports the out-of-sample performance from both statistical and economic perspectives for alternative forecasting windows. The forecasting models include two univariate regression models based on aligned investor sentiment (AIS) and the short interest index (SII), a bivariate regression model of both AIS and SII, and three combination approaches of mean combination, DMSPE(1), and DMSPE(0.9). Statistical performance includes the mean squared forecast error (MSFE) and out-of-sample R-square (R^2_{OS}). Statistical significance for the R^2_{OS} statistic is derived by using the Clark and West (2007) test. *** and ** indicate significance at the 1% and 5% levels, respectively. With respect to economic performance, the annualized certainty equivalent return (CER) gains and Sharpe ratios are calculated based on a mean-variance investor with relative risk aversion coefficient of three who allocates between stocks and risk-free bills monthly.

Table 6 provides the economic values for alternative investment environments. In short, we obtain a robust result that the bivariate regression model based on AIS and SII consistently yields the largest economic gains (that is, the largest CER gain and Sharpe ratio) for alternative asset allocation exercises, including various risk aversion coefficients and transaction costs. That is, the results of economic significance are consistent when the real world has various investors and investment environments.

4.3.4. Alternative predictors

The last but most important robustness check is that we further consider more predictors to support our strategy's superior predictive ability. Specifically, in the face of a limited number of useful predictors that could yield a positive R_{OS}^2, we, fortunately, find another two suitable predictors from the technical indicators of Neely et al. (2014). The first technical indicator is based on the moving-average (MA) rule and thus termed MA in this study. MA takes the value of one if the moving average of stock prices over the recent two months is larger than the one over the recent twelve months and zero otherwise. The second technical indicator is based on the trading volume and thus termed VOL in this study. VOL takes the value of one if the moving average of "on-balance" volume over the recent three months is larger than the one over the recent twelve months and zero otherwise.[5]

Table 7 reports the out-of-sample performance for the four predictors of AIS, SII, MA, and VOL. Several important and interesting findings emerge. First, in terms of R_{OS}^2, MA and VOL are less powerful than AIS but more powerful than SII. However, compared with the bivariate regression model based on AIS and SII, the bivariate regression model based on AIS and MA (VOL) shows a relatively small increase in R_{OS}^2. This is because although MA and VOL exhibit powerful individual predictive ability relative to SII, they also have relatively high correlation coefficients (around 0.2) with AIS. Therefore, MA and VOL provide less

[5] See Neely et al. (2014) for more details about the definitions of the technical indicators.

Table 6
Robustness checks of economic out-of-sample performance.

Forecasting models	No transaction cost		Transaction cost is 50 bps	
	CER gains (%)	Sharpe ratios	CER gains (%)	Sharpe ratios
Panel A: Risk aversion coefficient is 1				
AIS	6.03	0.23	5.83	0.23
SII	2.90	0.18	2.72	0.17
AIS and SII	7.61	0.28	7.14	0.27
Mean	5.96	0.23	5.66	0.22
DMSPE (1)	6.04	0.23	5.73	0.22
DMSPE (0.9)	5.99	0.23	5.68	0.22
Panel B: Risk aversion coefficient is 3				
AIS	5.12	0.21	4.88	0.20
SII	5.37	0.19	5.14	0.18
AIS and SII	8.50	0.27	8.20	0.26
Mean	6.69	0.24	6.40	0.23
DMSPE (1)	6.70	0.24	6.41	0.23
DMSPE (0.9)	6.67	0.24	6.35	0.22
Panel C: Risk aversion coefficient is 5				
AIS	3.08	0.18	2.88	0.17
SII	3.96	0.19	3.66	0.18
AIS and SII	6.59	0.26	6.22	0.25
Mean	4.74	0.23	4.49	0.22
DMSPE (1)	4.74	0.23	4.49	0.22
DMSPE (0.9)	4.71	0.23	4.45	0.21
Panel D: Risk aversion coefficient is 7				
AIS	2.12	0.18	1.97	0.16
SII	3.06	0.20	2.78	0.19
AIS and SII	5.05	0.26	4.67	0.25
Mean	3.42	0.23	3.22	0.22
DMSPE (1)	3.42	0.23	3.22	0.22
DMSPE (0.9)	3.41	0.23	3.20	0.21

This table reports the economic out-of-sample performance for alternative transaction costs and alternative risk aversion choices. The forecasting models include two univariate regression models based on aligned investor sentiment (AIS) and the short interest index (SII), a bivariate regression model of both AIS and SII, and three combination approaches of mean combination, DMSPE(1), and DMSPE(0.9). The annualized certainty equivalent return (CER) gains and Sharpe ratios are calculated based on a mean-variance investor with alternative risk aversion coefficients who allocates between stocks and risk-free bills monthly. The initial estimation period is 1973:01–1990:12, while the out-of-sample period is 1991:01–2014:12.

complementary information beyond that already contained in AIS. This is consistent with our suggestion.

Second, we find that the correlation between MA and VOL is as high as 0.674, suggesting that the complementary information between MA and VOL is probably scarce. As expected, the predictive regressions that incorporate either of MA and VOL (including only MA, only VOL, AIS and MA, AIS and VOL, SII and MA, SII and VOL) will yield smaller R^2_{OS}s when the corresponding regressions further include MA or VOL. For example, the regression model based on AIS and VOL generates the R^2_{OS} of 3.58%, while the regression model based on AIS, MA, and VOL yields a smaller R^2_{OS} of 3.40%. This evidence suggests that the predictors with high correlation cannot provide sufficient complementary information and are thus not suitable for our forecasting strategy. In addition, we find similar results for economic values of CER gains and Sharpe ratios.

Third, in this robustness check, we document the feasibility of our forecasting strategy not only for bivariate regression models but also for multivariate regression models. For example, the multivariate regression based on AIS, SII, and VOL always outperforms the regressions based on any of the predictors' subsets. As mentioned above, we observe a decline in R^2_{OS}, CER gain, and Sharpe ratio when further adding MA into the regression of AIS, SII, and VOL. This is because MA has high correlations with the three predictors, especially with VOL. Given this, it is important to note that we can further consider more predictors such as the 14 widely used economic variables of Welch and Goyal (2008) and the technical indicators of Neely et al. (2014), but we will not use these popular predictors due to their negative predictive ability or high correlations. More specifically, in our data sample, all of the economic variables generate negative R^2_{OS}s and thus cannot provide useful information by themselves. Although the technical indicators can yield positive R^2_{OS}s, they cannot provide complementary information due to high correlations. This is consistent with our case of MA and VOL.

Finally, our strategy using multivariate regressions based on the predictors that can provide complementary information always outperforms the combination approaches. For brevity, we only report the out-of-sample performance of the combination approaches based on all the four individual forecasts. The results are similar when we use the subset of the four predictors. In addition, the reader may wonder whether too few predictors are used in the combination approaches. For example, Rapach et al. (2010) use 14 economic variables to generate combination forecasts and obtain forecasting gains. The success of combination approaches is mainly due to the reduction of forecast volatility (Rapach et al., 2010). Therefore, when we add more predictors such as the 14 economic variables into our used predictor set, the combination forecasts become more stable. However, the stable forecasts are more close to the relatively inaccurate forecasts generated by the useless economic variables and thus are far away from the relatively accurate forecasts generated by the powerful predictors such as AIS and SII. Hence, the reduction of forecast volatility is useless in our case. In other words, our forecasting strategy can still beat the combination approaches when we further include the economic variables.

Table 7
Statistical and economic out-of-sample performance for alternative predictors.

Forecasting models	MSFE	R^2_{OS} (%)	CER gains (%)	Sharpe ratios
AIS	17.25	2.43***	5.12	0.21
SII	17.37	1.76***	5.37	0.19
MA	17.36	1.84**	6.14	0.22
VOL	17.30	2.14**	6.21	0.23
AIS, SII	16.90	4.43***	8.50	0.27
AIS, MA	17.09	3.37***	7.35	0.26
AIS, VOL	17.05	3.58***	7.00	0.25
SII, MA	17.20	2.74***	5.76	0.21
SII, VOL	17.14	3.07***	7.02	0.23
MA, VOL	17.31	2.09**	5.98	0.22
AIS, SII, MA	16.87	4.60***	7.57	0.25
AIS, SII, VOL	16.83	4.81***	8.65	0.27
AIS, MA, VOL	17.08	3.40***	7.23	0.25
SII, MA, VOL	17.17	2.92***	6.24	0.22
AIS, SII, MA, VOL	16.88	4.56***	7.52	0.25
Mean	17.21	2.69***	6.35	0.23
Median	17.17	2.89***	6.50	0.23
Trimmed mean	17.17	2.89***	6.50	0.23
DMSPE (1)	17.21	2.69***	6.36	0.23
DMSPE (0.9)	17.21	2.66***	6.34	0.23

This table reports the statistical and economic out-of-sample performance for alternative predictors. The used predictors include the aligned investor sentiment (AIS), the short interest index (SII) and two technical indicators of MA and VOL. Five combination approaches of mean combination, median combination, trimmed mean combination, DMSPE(1), and DMSPE(0.9) are based on the individual forecasts of the four used predictors. Statistical performance includes mean squared forecast error (MSFE) and out-of-sample R-square (R^2_{OS}). Statistical significance for the R^2_{OS} statistic is derived by using the Clark and West (2007) test. *** and ** indicate significance at the 1% and 5% levels, respectively. With respect to economic performance, the annualized certainty equivalent return (CER) gains and Sharpe ratios are calculated based on a mean-variance investor with relative risk aversion coefficient of three who allocates between stocks and risk-free bills monthly. The initial estimation period is 1973:01–1990:12, while the out-of-sample period is 1991:01–2014:12.

5. Simulation evidence

Why do combination approaches outperform the multivariate regression model (i.e., kitchen sink model) in the setting of Rapach et al. (2010)? Why do we observe an opposite pattern in this study? To answer these two questions, we conduct Monte Carlo simulations, which also reduce data mining concerns.

We simply simulate the stock return as $r_{t+1} = \mu + \sigma\varepsilon_t$ and the ith predictor as $x_{i,t} = \varepsilon_{i,t}$. The innovations ($\varepsilon$) of the stock return and predictors follow a standard normal distribution $N(0,1)$, and we define their covariance (i.e., correlation) matrix as Σ. According to the stock return data of the S&P 500 index from December 1950 to December 2016, we set the average return as $\mu = 0.407$ (in percent) and the return volatility as $\sigma = 4.378$. In addition, there are 5 predictors and 600 observations in this Monte Carlo simulation. That is, $N = 5$, $T = 12 \times 50$. The first 20 years (12×20 observations) are the initial estimation period, and the remaining 30 years are therefore the out-of-sample evaluation period.

First, we consider the innovation correlation matrix as

$$\Sigma_1 = \begin{vmatrix} 1.00 & 0.06 & 0.06 & 0.06 & 0.06 & 0.06 \\ 0.06 & 1.00 & 0.80 & 0.80 & 0.80 & 0.80 \\ 0.06 & 0.80 & 1.00 & 0.80 & 0.80 & 0.80 \\ 0.06 & 0.80 & 0.80 & 1.00 & 0.80 & 0.80 \\ 0.06 & 0.80 & 0.80 & 0.80 & 1.00 & 0.80 \\ 0.06 & 0.80 & 0.80 & 0.80 & 0.80 & 1.00 \end{vmatrix}, \tag{12}$$

so that the correlations between stock return and all predictors are identically 0.06. That is, the in-sample R^2 of the univariate predictive regression would be as low as 0.36% (0.06×0.06). The correlations among all predictors are identically 0.80, which is so high that it may lead to the multicollinearity issue. Panel A of Table 8 presents the out-of-sample predictive performance of the first Monte Carlo experiment.[6] It is clear that individual forecasts are less accurate than the prevailing historical average. Moreover, our method, i.e., multivariate regression, yields the smallest (largest) value of R^2_{OS} (MSFE) due to the over-fitting issue. In contrast, combination forecasts generate a positive R^2_{OS}, which indicates that combination forecasts can still capture useful information from various predictors despite the presence of the multicollinearity problem. This simulation evidence is consistent with the empirical evidence of Rapach et al. (2010). Rapach et al. (2010) argue that the individual forecasts are very unstable, while combination approaches can reduce forecast variance.

Second, we consider another innovation correlation matrix as

$$\Sigma_2 = \begin{vmatrix} 1.00 & 0.30 & 0.30 & 0.30 & 0.30 & 0.30 \\ 0.30 & 1.00 & 0.10 & 0.10 & 0.10 & 0.10 \\ 0.30 & 0.10 & 1.00 & 0.10 & 0.10 & 0.10 \\ 0.30 & 0.10 & 0.10 & 1.00 & 0.10 & 0.10 \\ 0.30 & 0.10 & 0.10 & 0.10 & 1.00 & 0.10 \\ 0.30 & 0.10 & 0.10 & 0.10 & 0.10 & 1.00 \end{vmatrix}. \tag{13}$$

Compared with the first simulation case, this case is more similar to the setting of this study. In this case, the predictors have a substantially higher predictive ability with an in-sample R^2 of 9% (0.30×0.30). In addition, the information contents of these predictors are complementary to each other because they have a low correlation. Panel B of Table 8 presents the out-of-sample performance of the second Monte Carlo experiment. Consistent with the empirical results in this paper, the multivariate predictive regression model outperforms all the competing models. This indicates that when the correlation among the predictors is low, the multivariate regression model can capture more useful information than combination approaches.

Third, we add an additional predictor to the second case, and the corresponding innovation correlation matrix is expressed as

[6] In each Monte Carlo experiment, we follow Kelly and Pruitt (2015) and conduct at least 5000 simulations and report the median values. The mean values are qualitatively similar.

Table 8
Out-of-sample performance of Monte Carlo simulations.

Forecasting models	MSFE	R^2_{OS} (%)
Panel A: First simulation case		
x_1	19.13	−0.05
x_2	19.13	−0.05
x_3	19.13	−0.03
x_4	19.12	−0.03
x_5	19.12	−0.04
Multivariate regression, $x_{1\text{-}5}$	19.31	−0.95
Mean	19.10	0.08
Median	19.10	0.07
Trimmed mean	19.10	0.08
DMSPE (1)	19.10	0.07
DMSPE (0.9)	19.10	0.08
Panel B: Second simulation case		
x_1	17.48	8.69***
x_2	17.49	8.60***
x_3	17.50	8.69***
x_4	17.50	8.63***
x_5	17.48	8.70***
Multivariate regression, $x_{1\text{-}5}$	13.17	31.32***
Mean	16.22	15.39***
Median	16.40	14.50***
Trimmed mean	16.29	15.01***
DMSPE (1)	16.22	15.39***
DMSPE (0.9)	16.23	15.37***
Panel C: Third simulation case		
x_1	16.89	8.60***
x_2	16.88	8.62***
x_3	16.88	8.64***
x_4	16.89	8.62***
x_5	16.88	8.65***
x_6	17.81	3.62***
Multivariate regression, $x_{1\text{-}6}$	12.34	33.25***
Mean	15.84	14.29***
Median	15.99	13.44***
Trimmed mean	15.93	13.84***
DMSPE (1)	15.83	14.36***
DMSPE (0.9)	15.83	14.35***

This table reports the out-of-sample performance based on Monte Carlo simulations. The forecasting models include several univariate regression models based on individual predictors, a multivariate regression model based on all used predictors, and five combination approaches of mean combination, median combination, trimmed mean combination, DMSPE(1), and DMSPE(0.9). MSFE represents the mean squared forecast error. Statistical significance for the out-of-sample R-square (R^2_{OS}) statistic is derived by using the Clark and West (2007) test. *** indicates significance at the 1% level.

$$\Sigma_3 = \begin{vmatrix} 1.00 & 0.30 & 0.30 & 0.30 & 0.30 & 0.30 & 0.20 \\ 0.30 & 1.00 & 0.10 & 0.10 & 0.10 & 0.10 & 0.05 \\ 0.30 & 0.10 & 1.00 & 0.10 & 0.10 & 0.10 & 0.05 \\ 0.30 & 0.10 & 0.10 & 1.00 & 0.10 & 0.10 & 0.05 \\ 0.30 & 0.10 & 0.10 & 0.10 & 1.00 & 0.10 & 0.05 \\ 0.30 & 0.10 & 0.10 & 0.10 & 0.10 & 1.00 & 0.05 \\ 0.20 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 1.00 \end{vmatrix}. \tag{14}$$

We can expect that the new predictor has an in-sample R^2 of 4% (that is, its predictive power is less than half of the existing predictors). Nonetheless, the correlation between the new predictor and the previous predictors is very low. The new predictor is thus likely to provide extra information for predicting future returns. Panel C of Table 8 presents the out-of-sample performance of the third Monte Carlo experiment. A remarkable result is that compared with the second case, there is a sizeable increase in R^2_{OS} for the multivariate regression model. This indicates that the new predictor can make a contribution to further enhancing out-of-sample performance, despite its limited individual predictive ability. Thus, we exploit a new way to forecast stock returns. Specifically, except for new powerful predictors, we can construct new predictors or seek existing predictors that have relatively low correlations with other predictors to improve the forecasting performance, despite their own relatively weak predictive ability. In contrast, we observe a slight decrease in R^2_{OS} for combination forecasts in this case, suggesting that the combination approaches fail to capture the information of the new predictor.

6. Conclusions

To explore stock return predictability, this paper proposes a simple but efficient approach in which we seek to utilize existing predictors with relatively limited predictive power and relatively low correlations with the other used predictors. Adding such predictors into a multivariate predictive regression model can provide complementary information that is useful to forecast stock returns. Not only the empirical results but also the simulation evidence verifies that our tractable method outperforms the benchmark model of the historical average, the univariate predictive regression model based on individual predictors, and the prevailing combination approaches used by Rapach et al. (2010) under the required conditions. The forecast encompassing tests show that our simple method can extract more complementary information that is useful for predicting stock returns with respect to the AIS and SII predictors. Furthermore, our empirical results are robust to various settings.

Acknowledgements

The authors are grateful to the editor and four anonymous referees for their insightful comments that help to improve the quality of this paper. This work is supported by Doctoral Innovation Fund Program of Southwest Jiaotong University [D-CX201724] and Service Science and Innovation Key Laboratory of Sichuan Province [KL1704]. Feng Ma acknowledges the support from the Natural Science Foundation of China [71671145, 71701170], the Humanities and Social Science Fund of the Ministry of Education [17YJC790105, 17XJCZH002], and Fundamental research funds for the Central Universities [682017WCX01, 2682018WXTD05].

References

Bollerslev, T., Marrone, J., Xu, L., Zhou, H., 2014. Stock return predictability and variance risk premia: statistical inference and international evidence. J. Financ. Quant. Anal. 49, 633–661.
Bollerslev, T., Tauchen, G., Zhou, H., 2009. Expected stock returns and variance risk premia. Rev. Financ. Stud. 22, 4463–4492.
Campbell, J.Y., Thompson, S.B., 2008. Predicting excess stock returns out of sample: can anything beat the historical average? Rev. Financ. Stud. 21, 1509–1531.
Clark, T.E., West, K.D., 2007. Approximately normal tests for equal predictive accuracy in nested models. J. Econom. 138, 291–311.
Harvey, D.S., Leybourne, S.J., Newbold, P., 1998. Tests for forecast encompassing. J. Bus. Econ. Stat. 16, 254–259.
Huang, D., Jiang, F., Tu, J., Zhou, G., 2015. Investor sentiment aligned: a powerful predictor of stock returns. Rev. Financ. Stud. 28, 791–837.
Jiang, F., Lee, J.A., Martin, X., Zhou, G., 2017. Manager sentiment and stock returns. J. Finan. Econ. forthcoming.
Kelly, B., Pruitt, S., 2015. The three-pass regression filter: a new approach to forecasting using many predictors. J. Econom. 186, 294–316.
Li, X., Shen, D., Xue, M., Zhang, W., 2017. Daily happiness and stock returns: the case of Chinese company listed in the United States. Econ. Modell. 64, 496–501.
Lin, Q., 2018. Technical analysis and stock return predictability: an aligned approach. J. Financ. Market. 38, 103–123.
Ma, F., Li, Y., Liu, L., Zhang, Y., 2018a. Are low-frequency data really uninformative? A forecasting combination perspective. N. Am. J. Econ. Finance 44, 92–108.
Ma, F., Liu, J., Wahab, M.I.M., Zhang, Y., 2018b. Forecasting the aggregate oil price volatility in a data-rich environment. Econ. Modell. 72, 320–332.
Manela, A., Moreira, A., 2017. News implied volatility and disaster concerns. J. Financ. Econ. 123, 137–162.
Narayan, P.K., Bannigidadmath, D., 2017. Does financial news predict stock returns? New evidence from Islamic and non-Islamic stocks. Pac. Basin Finance J. 42, 24–45.
Narayan, P.K., Narayan, S., Phan, D.H.B., Thuraisamy, K.S., Tran, V.T., 2017a. Credit quality implied momentum profits for Islamic stocks. Pac. Basin Finance J. 42, 11–23.
Narayan, P.K., Phan, D.H.B., Narayan, S., Bannigidadmath, D., 2017b. Is there a financial news risk premium in Islamic stocks? Pac. Basin Finance J. 42, 158–170.
Neely, C.J., Rapach, D.E., Tu, J., Zhou, G., 2014. Forecasting the equity risk premium: the role of technical indicators. Manag. Sci. 60, 1772–1791.
Ni, Z.-X., Wang, D.-Z., Xue, W.-J., 2015. Investor sentiment and its nonlinear effect on stock returns—new evidence from the Chinese stock market based on panel quantile regression model. Econ. Modell. 50, 266–274.
Pettenuzzo, D., Timmermann, A., Valkanov, R., 2014. Forecasting stock returns under economic constraints. J. Financ. Econ. 114, 517–553.
Rapach, D.E., Ringgenberg, M.C., Zhou, G., 2016. Short interest and aggregate stock returns. J. Financ. Econ. 121, 46–65.
Rapach, D.E., Strauss, J.K., Zhou, G., 2010. Out-of-sample equity premium prediction: combination forecasts and links to the real economy. Rev. Financ. Stud. 23, 821–862.
Rossi, B., Inoue, A., 2012. Out-of-sample forecast tests robust to the choice of window size. J. Bus. Econ. Stat. 30, 432–453.
Wang, Y., Liu, L., Ma, F., Diao, X., 2018. Momentum of return predictability. J. Empir. Finance 45, 141–156.
Welch, I., Goyal, A., 2008. A comprehensive look at the empirical performance of equity premium prediction. Rev. Financ. Stud. 21, 1455–1508.
Xue, W.-J., Zhang, L.-W., 2017. Stock return autocorrelations and predictability in the Chinese stock market—evidence from threshold quantile autoregressive models. Econ. Modell. 60, 391–401.
Yang, C., Yan, W., Zhang, R., 2013. Sentiment approach to negative expected return in the stock market. Econ. Modell. 35, 30–34.
Zhang, Y., Ma, F., Shi, B., Huang, D., 2018. Forecasting the prices of crude oil: an iterated combination approach. Energy Econ. 70, 472–483.
Zhu, X., 2013. Perpetual learning and stock return predictability. Econ. Lett. 121, 19–22.
Zhu, X., Zhu, J., 2013. Predicting stock returns: a regime-switching combination approach and economic links. J. Bank. Finance 37, 4120–4133.