Dai Zhifeng (Orcid ID: 0000-0002-7912-1280)

Forecasting stock return volatility: The role of shrinkage approaches in a data-rich environment

Zhifeng Dai[1], Tingyu Li[1], Mi Yang[2]

[1]College of Mathematics and Statistics, Hunan Provincial Key Laboratory of Mathematical Modeling and Analysis in Engineering Changsha University of Science and Technology, Hunan, 410114 China.

[2]School of economics and management, Changsha University of Science and Technology, Hunan, 410114 China.

Abstract

This paper employs the prevailing shrinkage approaches, the lasso, adaptive lasso, elastic net and ridge regression to predict stock return volatility with a large set of variables. The out-of-sample results reveal that shrinkage approaches exhibit superior performance relative to the benchmark of the autoregressive model and a series of competing models in terms of the out-of-sample R-square and the model confidence set. By using shrinkage methods to allocate portfolio, a mean-variance investor can obtain significant economic gains. Overall, our findings confirm that shrinkage approaches can effectively improve stock return volatility forecasting in a data-rich environment.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/FOR.2841

Keywords: Shrinkage regressions; Stock return volatility; Out-of-sample forecasting; Model confidence set; Asset allocation.

1 Introduction

Forecasting stock return volatility plays an essential role in many financial fields, such as portfolio allocation, risk management and asset pricing. Therefore, a growing number of literature focus on how to predict stock return volatility, for example, Engle and Rangel (2008), Engle et al. (2013), Degiannakis et al. (2014), Huang et al., 2015, Nonejad (2017) and Wang et al. (2018).

The seminal work by Schwert (1989) believes that sources of financial volatility are closely related to real and nominal macroeconomic volatility, economic activity and financial leverage. Understanding the relationship between a series of variables and stock market return volatility is an important empirical problem in finance. From a longer-term value-at-risk perspective, it is important to understand how future stock market volatility will respond to changing conditions. However, according to Nonejad (2017), due to parameter instability and model uncertainty, a regression model that includes lags of realized volatility and an individual macroeconomic variable often fails to generate more accurate forecasts than the benchmark of autoregressive model. In view of this, finding a suitable forecasting method is the main challenge, which can explore the powerful factors related to the stock return volatility in a data-rich environment, and obtain more accurate forecasts of stock market return volatility.

In recent years, shrinkage methods such as least absolute shrinkage and selection operator (lasso) in Tibshirani (1996) has found to perform well in many applications. Li et al. (2015) and Li and Tsiakas (2017) study the predictability of exchange rates and stock returns, respectively. Zhang et al. (2019a) and Zhang et al. (2019b) investigate the predictability of oil prices and oil prices volatility, respectively. In addition, Audrino and Knaus (2016) and Audrino et al. (2019) combine heterogeneous autoregressive (HAR) with lasso and adaptive lasso to predict several individual stocks return volatility, respectively. Inspired by the above research, this paper attempts to use shrinkage methods such as lasso, elastic net and ridge regression (Zou and Hastie, 2005) to predict stock market return volatility with a large set of variables. Furthermore, we make a comparison between the forecasting strategies with respect to their out-of-sample forecasting performance in the statistical and economic senses.

In this paper, the prevailing shrinkage regression methods employ the common volatility prediction model, which takes the logarithm of the realized volatility as the dependent variable, and uses the lag term and a series of variables as the autoregressive term and predictors, respectively. Paye (2012), Christiansen et al. (2012) and Nonejad (2017) provide macroeconomic variables that can be considered as predictors in forecasting volatility. Oil has a great impact on economic conditions reflected in cash flow and discount rates. The main link between oil prices and stocks is that oil prices have a potential impact on stocks by affecting corporate cash flow and profitability. Therefore, crude oil volatility might be relevant in predicting future levels of stock market volatility. Baker and Wurgler (2007) and Huang et al. (2015) introduce investor and market sentiment indicators to predict stock returns accurately, and they also point out that investor and market sentiment indicators can also help predict financial volatility. In addition, following Neely et al. (2014), we will construct a wide range of technical indicators based on stock price, and trading volume, which are used to forecast stock return volatility. The macroeconomic variables and technical indicators mentioned above form a data-rich environment. However, the prediction performance of adding a series of variables to the stock market return volatility prediction model is unclear.

In this paper, we use four popular shrinkage regression models (lasso, adaptive lasso, elastic net and ridge) to forecast stock return volatility by using the data of the SPY range from January 1995 to December 2018. The SPDR S&P 500 Trust ETF, also known as the SPY ETF, aims to track the S&P 500 index. We consider market and investor sentiment indicators (UMS), macroeconomic fundamentals (MF), technical indicators (TI) and oil volatility indicators (OV) as predictors. There are seven UMS indicators, including the uncertainty index (UI), the economic policy uncertainty index (EPU) constructed by Baker et al. (2016), market volatility (EWV), Close end fund discount rate (CEFD), Number of IPOs (NIPO), Dividend premium (PDND), and Sentiment index (SENT). Macroeconomic fundamentals including 13 macroeconomic variables are proposed by Welch and Goyal (2008). Technical indicators are constructed based on prices and volumes, and oil volatility indicators mainly consider crude oil volatility and Brent oil volatility in the paper. We have several notable findings in our empirical results.

First, the out-of-sample results suggest that in a standard linear regression framework, most variables don't seem to predict stock return volatility well, which is consistent with the previous literature. It is well known that the univariate regression model only incorporates finite information, leading to relatively weak predictive ability. Out-of-sample results suggest that the elastic net not only outperforms the univariate regression model, but also is superior to the ridge regression and lasso in the shrinkage methods. And the elastic net has the largest MCS p-value of 1, which is the best in the overall MCS test. In general, out-of-sample results indicate that predictive performance of elastic net, adaptive lasso, lasso and ridge models outperform the univariate regression models. In addition, based on the random volatility model added to the original set of model, the contraction method is found to still has strong predictive performance in predicting the volatility of stock returns. The realized volatility used is a measure of past volatility, not future uncertainty. We replaced it with other implied volatility, such as CBOE VIX. The MCS test shows that the lasso provides the smallest loss function and has a strong predictive performance. Therefore, the superior forecasting performance of shrinkage methods in data rich environment is robust to the selection of model sets and explained variables.

Secondly, previous studies, such as Ma et al. (2018), Wang et al. (2016) and Zhang et al. (2019b), argue that statistically significant is not sufficient to prove the superiority of a specific model in predicting stock return volatility. Market investors and decision makers are more interested in the economic value of predictive models with the purposed of making correct decisions and obtaining higher economic gains, which leads us to measure the economic value of stock return volatility predictability of shrinkage methods from an asset allocation perspective. We report the utility gains of portfolio with models of interest relative to benchmark model. A mean-variance investor allocates his assets between stock and the risk-free Treasury bill using shrinkage methods can realize positive utility gains of portfolio, while the certainty equivalent return gains in univariate regression model is negative. Shrinkage approaches obtain more accurate forecasts than univariate models. Specifically, ridge regression based on technical indicators achieves the highest economic utility benefit among all models. In addition, our results suggest that based on UMS, MF, TI and OV information using shrinkage estimation regression to calculate economic value is quite helpful for asset allocation.

This study makes two major contributions to the literature on volatility forecasting as follows. First, the out-of-sample results illustrate that the shrinkage models can generate more accurate stock return volatility forecasting based on a series of predictors. Second, we find that a mean-variance investor who uses the volatility forecasts based on the elastic net and lasso to allocate between stock index and risk-free bills can realize positive certainty equivalent return (CER) gains. In summary, the stock market return volatility forecasts generated by shrinkage regression models are statistically and economically significant for out-of-sample prediction performance.

Our paper is related to the work of Audrino and Knaus (2016) and Audrino et al. (2019) as the two studies both rely on lasso. On the contrary, we can provide at least three new contributions or major differences. First, Audrino and Knaus (2016) and Audrino et al. (2019) make an empirical analysis on several individual stocks returns volatility, while our paper focuses on the stock market return volatility. Second, Audrino and Knaus (2016) and Audrino et al. (2019) explore whether the lag structure imposed by the simple HAR-RV model can be recovered by the lasso or adaptive lasso. However, we collect all the predictors from its various extensions and investigate whether all these predictors can be efficiently used by the lasso, adaptive lasso, elastic net and ridge. In addition, this paper further does some additional work relative to Audrino and Knaus (2016), such as portfolio exercise. Therefore, our paper contributes new insights into the existing literature on stock return volatility forecasting.

The remainder of the paper is organized as follows. Section 2 presents the data sources. Section 3 describes our methodology. Section 4 provides empirical results. In Section 5 investigates asset allocation. Section 6 concludes.

2 Data

2.1 Stock return volatility

The S&P 500 serves as one of the main benchmarks of the US equity market and indicates the financial health and stability of the economy. However, the S&P 500 index is not traded from the perspective of asset allocation, so we use the data of SPDR S&P 500 Trust ETF to study the stock return volatility. The SPDR S&P 500 Trust ETF, also known as the SPY ETF, is one of the most popular funds that aims to track the Standard & Poor's 500 Index (see, e.g., Xu et al., 2016; Xu and Yin, 2017; Corbet and Twomey, 2014, Latunde et al., 2020). And the Granger causality results of Xu and Lin (2017) show that there is a strong interaction and a close correlation between the volatility of S&P 500 index and SPY ETF. Therefore, following Christiansen et al. (2012), Ma et al. (2018), Gong and Lin (2018, 2020) and Liu and Gong (2020), we use the sum of SPY ETF daily return measures monthly stock return volatility as follows

$$RV_t = \sum_{i=1}^{N_t} R_{i,t}^2 \qquad (1)$$

$$R_{i,t} = \log(P_{i,t}) - \log(P_{i-1,t}) \qquad (2)$$

where N_t indicates the number of trading days in the t-th month and $R_{i,t}$ means the daily return on the SPY on the i-th trading day of the t-th month, $P_{i,t}$ denotes the daily price on the SPY on the i-th trading day of the t-th month.

Paye (2012) find that errors of realized volatility are non-Gaussian. However, the distribution of logarithms of realized volatility, $V_t = \log(RV_t)$, is approximately Gaussian according to the finding of Andersen et al. (2001).

2.2 Market and investor sentiment indicators

As is well known, market and investor sentiment indicators have an impact on stock market volatility. Sentiment indicators are widely used in equity research. Sentiment is not only a phenomenon observed by professional traders but sentiment influences professional traders.

Long et al. (1990) point out from a theoretical point of view that investor sentiment exists widely in financial markets. Following the literature, we adopt seven variables that represent market and investor sentiment to research the influence of sentiment on future stock return volatility. Liu and Zhang (2015) apply EPU successfully to forecast stock return volatility. Therefore, Jurado et al. (2015) propose the uncertainty index (UI) and Baker et al. (2016) construct the economic policy uncertainty index (EPU) to represent market uncertainty and sentiment.[1] Moreover, we use four investor sentiment proxies of Baker and Wurgler (2007) namely CEFD, NIPO, PDND and SENT. Close end fund discount rate (CEFD) is value-weighted average difference between the net asset values of closed end stock mutual fund shares and their market prices; Number of IPOs (NIPO) is monthly number of initial public offerings; Dividend premium (PDND) is difference of the value-weighted average market-to-book ratios of dividend payers and nonpayers; Sentiment index (SENT) in Baker and Wurgler (2007) is based on first principal component of five sentiment proxies. Equity Market Volatility (EMV) is created as the equity market volatility trackers that change with the volatility index and the realized stock return volatility. EMV is also a simple volatility indicator that combines price and volume. It is made on the base of the principle of isometric chart.

2.3 Macroeconomic predictors

Our macroeconomic fundamentals use monthly data from Welch and Goyal (2008). We investigate the monthly stock return volatility based on 13 representative macroeconomic predictors selected in Welch and Goyal (2008).[2]

1. Dividend-price ratio (log), DP: log of a 12-month moving sum of dividends paid on the S&P 500 Index minus the log of stock prices (S&P 500 Index).

2. Dividend yield (log), DY: log of a 12-month moving sum of dividends minus the log of lagged stock prices.

3. Earnings-price ratio (log), EP: log of a 12-month moving sum of earnings on the S&P

[1] UI and EPU are available at https://www.sydneyludvigson.com/data-and-appendixes/ and http://www.policyunc-ertainty.com/, respectively.

[2] The data can be downloaded from the websites: http://www.hec.unil.ch/agoyal/ (Welch and Goyal, 2008).

500 Index minus the log of stock prices.

4. Dividend-payout ratio (log), DE: log of a 12-month moving sum of dividends minus the log of a 12-month moving sum of earnings.

5. Book-to-market ratio, BM: book-to-market value ratio for the Dow Jones Industrial Average.

6. Net equity expansion, NTIS: ratio of a 12-month moving sum of net equity issues by NYSE-listed stocks to the total end-of-year market capitalization of New York Stock Exchange (NYSE) stocks.

7. Treasury bill rate, TBL: interest rate on a three-month Treasury bill (secondary market).

8. Long-term yield, LTY: long-term government bond yield.

9. Long-term return, LTR: return on long-term government bonds.

10. Term spread, TMS: long-term yield minus the Treasury bill rate.

11. Default yield spread, DFY: difference between Moody's BAA- and AAA-rated corporate bond yields.

12. Default return spread, DFR: long-term corporate bond return minus the long-term government bond return.

13. Inflation, INFL: calculated from the CPI for all urban consumers;

2.4 Technical indicators

Economic fundamentals have received much more attention than technical indicators for forecasting stock volatility. Technical indicators based on past stock prices or volume patterns try to distinguish future trends. Various studies have researched the profitability of many technical rules, such as the momentum strategy (e.g., Fuertes et al., 2010; Moskowitz et al., 2012, Dai and Shi, 2021) and the moving average and channel strategy (e.g., Szakmary et al., 2010) in stock futures trading.

We employ 18 technical indicators based on three popular trend-following strategies. The first strategy is based on moving-average (MA). It produces a buying signal when the price exceeds its most recent low by more than a given percentage and produces a selling signal when it moves in the opposite direction.

$$S_{i,t} = \begin{cases} 1 & if \ \ MA_{s,t} \geq MA_{j,t} \\ 0 & if \ \ MA_{s,t} < MA_{j,t} \end{cases} \tag{3}$$

$$MA_{j,t} = \left(\frac{1}{j} \right) \sum_{i=0}^{j-1} P_{t-i} \quad for \ \ j = s, l; \tag{4}$$

where P_t is the value of SPY ETF, and $s(l)$ is the length of the short (long) MA $(s < l)$. In fact, short MA will detect recent price movement, because the MA rules imply stock price trends. We research the monthly MA rules with $s = 1, 2, 3$ and $l = 9, 12$.

The second rule is based on momentum that generates the following signal:

$$S_{i,t} = \begin{cases} 1 & if \ \ P_t \geq P_{t-m} \\ 0 & if \ \ P_t < P_{t-m} \end{cases} \tag{5}$$

In the rule, a sell signal indicates that value is lower than its level of m periods ago, which implies “negative” momentum and a low stock return. We compute the monthly signals for $m = 1, 2, 3, 6, 9, 12$.

The third technical indicators use volume data to analysis market trends. We first define

$$OBV_t = \sum_{k=1}^{t} VOL_k D_k \tag{6}$$

where VOL_k is a measure of the trading volume during period k. If $P_k - P_{k-1} \geq 0$, then D_k takes 1 and if it is on the contrary that takes a value of -1. The following defines OBV as

$$S_{i,t} = \begin{cases} 1 & if \ \ MA_{s,t}^{OBV} \geq MA_{i,t}^{OBV} \\ 0 & if \ \ MA_{s,t}^{OBV} < MA_{i,t}^{OBV} \end{cases} \tag{7}$$

$$MA_{j,t}^{OBV} = \left(\frac{1}{j} \right) \sum_{i=0}^{j-1} OBV_{t-i} \quad for \ \ j = s, l \tag{8}$$

A relatively high volume in the recent period will form a buy signal, which is a positive market trend and the monthly signals of the trading volume is calculated for $s = 1, 2, 3$ and $l = 9, 12$.

Considering the lag rules of technical indicators, the technical indicators data in formulas (3), (5) and (7) are obtained by using the data that daily SPY and monthly trading volume data start from January 1995. Technical indicators include MA, momentum and volume rules proposed by Sullivan et al. (1999), which represent the trend of the stock market. In addition, Neely et al. (2014) show that technical indicators better detect the typical decline in the equity risk premium near business-cycle peaks, whereas macroeconomic variables more readily pick up the typical rise in the equity risk premium near cyclical troughs. There are some studies on the predictability of macroeconomic variables for stock market volatility, for example, Christiansen et al. (2012), Paye (2012), Engle et al. (2013) and Nonejad. (2017). In this paper, we try to investigate the predictability of technical variables for the stock market return volatility.

2.5 Oil volatility

We obtain two oil prices from the website of the Energy Information Administration, West Tex as Intermediate (WTI) crude oil and Brent oil.[3] We can obtain more useful information from WTI oil price and it is the benchmark of oil price, however, Brent oil price reflects more international oil demand and supply information. Feng et al., (2018) show oil volatility does exhibit statistically and economically significant predictability for stock market volatility. The sample period of two oil prices monthly data is from January 1995 to December 2018.

Insert Table 1 about here

Table 1 reports descriptive statistics for the monthly data from January 1995 through December 2018, including stock return volatility, market and investor sentiment indicators, macroeconomic variables from the Amit Goyals data set, technical indicators based on price and trading volume, and oil volatility.

3 Forecasting methods

3.1 Predictive regressions

The model that we use to predict one-month stock volatility model is autoregressive model (AR), and its standard benchmark is as follows:

[3] Data for the second set of variables can be obtained from the Energy Information Administration (EIA) (http://www.eia.gov/).

$$V_{t+1} = \omega + \sum_{i=0}^{p-1} \alpha_i V_{t-i} + \varepsilon_{t+1} \tag{9}$$

where V_{t+1} is t+1-th month stock volatility, V_{t-i} is i+1-th lag of V_{t+1}. ω and α_i are constant coefficient and lag coefficient respectively, the ε_{t+1} is a disturbance term with zero mean. Following Paye (2012), we should set the lag order p as 6 when our data is monthly stock return volatility. The reason for using this length is that it can get auto correlation information of stock fluctuations.

In order to explore the influence of various factors on stock volatility, we added the influence of other individual predictor to the equation (10), and the model obtained is as follows:

$$V_{t+1} = \omega + \sum_{i=0}^{p-1} \alpha_i V_{t-i} + \beta x_{i,t} + \varepsilon_{t+1} \tag{10}$$

where $x_{i,t}$ represents the *i*-th indicator in the *t*-th month that affect stock volatility. The influence of other factors on the stock market volatility is reflected in the parameter β.

We use least squares estimation (OLS) to calculate the coefficient estimates in (10). In a standard regression framework, the null hypothesis of unpredictability, $\beta = 0$, can be tested using the standard t-statistic. Considering that there may be serial correlation between the data, we employ heteroscedasticity and autocorrelation consistent t-statistic to test the null hypothesis.

To make full use of past information, we perform predictive regression to generate out-of-sample stock volatility forecasts by using a recursive window. We separate into in-sample part and out-of-sample part from the T total sample, including stock volatility series and other indicators. There are M observations in-sample and $T-M$ observations out-of-sample.

3.2 Shrinkage approach

A popular economic and financial forecasting methodology is called Least absolute shrinkage and selection operator (LASSO) in the out-of-sample forecasting method.

Following Elliott et al. (2013), Li et al. (2015), Li and Tsiakas (2017) and Zhang et al. (2019a), we mainly draw lessons from the most popular and simple three of their research methods, respectively, the original lasso of Tibshirani (1996), the ridge and the elastic net of Zou and Hastie (2005).

We apply shrinkage estimation because it can minimize the out-of-sample mean squared error (MSE) through shrinking the regression coefficients towards zero. The traditional OLS estimators are unbiased but might have a greater out-of-sample MSE, however, a shrinkage estimator is biased but have lower variance than OLS. Compared with OLS, the shrinkage estimator can produce more accurate predictions and smaller out-of-sample mean square errors, which is helpful for stock volatility forecasts.

We quote the following example to explain the role of the shrinkage estimator more clearly. OLS estimation will produce unbiased estimates with high variance, and the benchmark model in which regression coefficients are set as zero will generate bias estimates but zero variance. Therefore, the shrinkage estimator is a combination of the above two situations and is a good bias trade-off. Statistically, the lasso forecasts of stock return volatility are expressed as:

$$V_{t+1} = \hat{\beta}_0 + \sum_{i=0}^{p-1} \alpha_i V_{t-i} + \sum_{i=1}^{N} \hat{\beta}_i x_{i,t} \tag{11}$$

where

$$\hat{\beta} = \arg\min_{\beta} \left(\frac{1}{2(t-1)} \sum_{l=1}^{t-1} \left(V_{t+1} - \beta_0 - \sum_{i=0}^{p-1} \alpha_i V_{t-i} - \sum_{i=1}^{N} \beta_i x_{i,t} \right)^2 + \lambda \sum_{i=1}^{N} |\beta_i| \right) \tag{12}$$

V_{t+1} is the actual log stock return volatility at month *t*+1, $x_{i,t}$ refer to the *i*-th indicator of each variable collection type at *t*-th month. We regress the available data before month t to obtain the shrinkage estimator $\hat{\beta}$ in the lasso, and coming true non-negative regularization of β by limiting the penalty function λ. Therefore, the ordinary least squares estimator can also be expressed as equation (11).

$$\hat{\beta} = \arg\min_{\beta} \frac{1}{2} \sum_{l=1}^{t-1} (V_{l+1} - \beta_0 - \sum_{i=0}^{p-1} \alpha_i V_{t-i} - \sum_{i=1}^{N} \beta_i x_{i,t})^2 \tag{13}$$

$$\hat{\beta} = \arg\min_{\beta} \left(\frac{1}{2(t-1)} \sum_{l=1}^{t-1} \left(V_{t+1} - \beta_0 - \sum_{i=0}^{p-1} \alpha_i V_{t-i} - \sum_{i=1}^{N} \beta_i x_{i,t} \right)^2 + \lambda \sum_{i=1}^{N} \left((1-\alpha)\beta_i^2 + \alpha |\beta_i| \right) \right) \quad (14)$$

where α and λ are positive constant between 0 and 1, which are based on the principle of minimizing the mean square error of the estimated forecasts. Specifically,
in both circumstances $\alpha = 1$ and $\alpha = 0$, the elastic net is transformed into lasso regression and ridge regression respectively.

When employing the lasso, elastic net and ridge regression to forecast stock return volatility at month t+1, we should use the data up to t period to estimate the optimal coefficients α and λ. Following Li and Tsiakas (2017) in a recent study, we find that the commonly used method for calculating the lasso parameters of α and λ is cross-validation. Specifically, when we use the lasso, elastic net and ridge regression to generate stock return volatility forecast at t+1-th period, we adopt cross-validation.

3.3 Adaptive Lasso

In some cases, the lasso is inconsistent in the choice of variables. However, the adaptive lasso proposed by Zou (2006) and Nardi and Rinaldo (2011) leads to the condition that the variable selection is consistent, which is to assign different weights to the penalty coefficient. The same penalty coefficient is unfair in a sense, which will make the coefficients of different variables suffer the same penalty. Therefore, we think about adaptive lasso, namely meaning different coefficients are assigned different weights.

$$\hat{\beta} = \arg\min_{\beta} \left(\frac{1}{2(t-1)} \sum_{l=1}^{t-1} \left(V_{t+1} - \beta_0 - \sum_{i=0}^{p-1} \alpha_i V_{t-i} - \sum_{i=1}^{N} \beta_i x_{i,t} \right)^2 + \lambda \sum_{i=1}^{N} \omega_i |\beta_i| \right) \quad (15)$$

where ω_i is the weight of the i-th coefficient. Considering the correlation between weights and data and making the weighted lasso have oracle properties, the new method obtained is adaptive lasso. We define the parameters of the adaptive lasso, choose a positive γ, and the weight vector $\hat{\omega} = 1/\left|\hat{\beta}\right|^{\gamma}$, then the adaptive lasso estimate $\hat{\beta}^{*(n)}$ is the following formula.

$$\hat{\beta}^{*(n)} = \arg\min_{\beta} \left(\frac{1}{2(t-1)} \sum_{l=1}^{t-1} \left(V_{t+1} - \beta_0 - \sum_{i=0}^{p-1} \alpha_i V_{t-i} - \sum_{i=1}^{N} \beta_i x_{i,t} \right)^2 + \lambda_n \sum_{i=1}^{N} \hat{\omega}_i |\beta_i| \right) \quad (16)$$

The adaptive lasso, which is essentially a penalty method, can be calculated according to the algorithm of the lasso.

3.4 Model confidence set evaluation method

The model confidence set (MCS) test originally proposed by Hansen et al. (2011) tends to choose a model that dominates competitors even in the case of many alternatives. The MCS test can also verify whether the out-of-sample forecasting performance of multiple forecasting models for stock return volatility is statistically significant. We use two popular loss functions, mean squared error (MSE) and mean absolute error (MAE). The two loss functions are defined as follows:

$$\text{MSE}=\frac{1}{q}\sum_{t=m+1}^{m+q}(RV_t - RV_t)^2 \tag{17}$$

$$\text{MAE}=\frac{1}{q}\sum_{t=m+1}^{m+q}\left|RV_t - RV_t\right| \tag{18}$$

The meaning of the p-value generated by the MCS test is not much different from the classical p-value, and a model with good predictive performance is accompanied by a larger p-value. The MCS p-value reported in this paper is a statistic calculated based on stationary bootstrap as in Zhang et al. (2019b). Following Hansen et al. (2011), Wang et al. (2016) and Liu et al. (2018) the MCS p-value is less than 0.1, the model is defined as an inferior model, and the prediction model with p-value greater than 0.1 is selected and included in the MCS set.

4 Empirical results for the model confidence set evaluation method

4.1 Out-of-sample forecasting performance for the MCS approach

In this section, we conduct an MCS test on the stock volatility forecasts generated in the period from January 2001 to December 2018, and compare the univariate model with the shrinkage methods in the data-rich environment.

Insert Table 2 about here

Panel A of Table 2 compares the forecasting performance of the prediction models of the four types of variables respectively, including the MCS p-values of the two popular loss functions. All four shrinkage methods appear in the MCS model of classification comparison of the four variables with 90% confidence level of every variable. In addition, the adaptive lasso in four shrinkage methods always provides the maximum MCS p-value of 1. This indicates that the out-of-sample forecasting performance of shrinkage methods outperform univariate model of various indicators, including investor sentiment indexes, macroeconomic variables, technical indicators, and oil volatility indicators.

To directly observe the prediction performance of the shrinkage methods in a data-rich environment, we use MCS to test all variables and the fluctuation prediction generated by the shrinkage methods based on all variables. Panel B of Table 4 reports the out-of-sample prediction performance of all prediction models in the case of rich data without classification comparison, which is the essential difference from panel A. We find an interesting phenomenon that almost all univariate prediction models cannot enter the MCS sets with 90% confidence for two loss functions in the MCS test of variable unclassified comparison, and univariate prediction model are judged as inferior models.

Elastic net is a good choice to forecast the stock return volatility in the rich data environment. The MSE and MAE values of Elastic net are the lowest values, which indicates that the Elastic net results based on the MSE and MAE loss functions have the best overall performance of the shrink model. Only the shrinkage methods always enter the MCS sets with 90% confidence level in all cases, but the shrinkage methods which produce the maximum MCS p-value in different data cases is different. Finally, comparing the results of panel A and panel B, we find that the performance of shrinkage methods is better in the richer data environment.

Insert Figure 1 about here

Figure 1 is a graphical representation of the four shrinkage estimation methods, which are used to predict the sentiment variables, economic fundamentals, technical indicators, oil volatility indicators and stock volatility based on all indicators in the sampling period. Some interested conclusions can be drawn by describing the actual and predictive changing trend of stock volatility. First, some large events always occur with stock volatility. For example, the stock volatility reached the historical highest peak in the entire sample period, which is because the financial crisis broke out in the United States in the same year. More importantly, the forecast of stock volatility is highly consistent with the actual trajectory of stock volatility, and the phenomenon shows that our models have good forecasting performance.

4.2 Alternative benchmark GRACH model

Choosing an appropriate benchmark model can avoid artificially discovering the forecasting performance of shrinkage methods. The main purpose of this article is to evaluate the effectiveness of the shrinkage methods relative to the autoregressive model in predicting the stock return volatility. Specifically, it might be interesting if we compare their forecasting approach with forecasts from stochastic volatility family of models and shrinkage methods. The generalized autoregressive conditional heteroscedasticity model in the volatility model has a large number of literature for the research of the oil market (see, e.g., Aloui and Mabrouk, 2010; Kang et al., 2009; Narayan and Narayan, 2007, Kawabata, 2020). Therefore, we use the GRACH model as an alternative benchmark model to predict the stock return volatility market. We determine the order of GRACH according to AIC and SIC criteria and select GARCH (1,1) to forecast stock return volatility. Abounoori et al. (2016) found that the forecasting performance of the GARCH model is superior.

Insert Table 3 about here

In this part, we want to know whether the result is robust to the GRACH model. Table 3 reports the forecasting performance of different contraction operators for stock return volatility in the model set including GRACH model. We find that elastic net appears in MCS, and provide the smallest loss function and get the largest p-value in both two types of model sets. Consistent with the main results of our original model set, elastic net is all models with the best forecasting performance.

4.3 Stock market implied volatility predictability

The CBOE Volatility Index, referred to as CBOE VIX, is one of the most commonly used implied volatility. The forward-looking forecast of volatility over the coming 30 days generated by the CBOE VIX also represents the market's expectation of the future price change of the S&P 500 index. The speed of price changes and the magnitude of volatility are often an expression of investor sentiment in the market. Therefore, the CBOE VIX provides a measure of market risk and is also an important reference index in the field of trading and investment.

Insert Table 4 about here

Table 4 reports the out-of-sample prediction performance of CBOE VIX for a series of explanatory variables. An interesting phenomenon is that elastic nets are no longer the shrinking method with the best predictive ability, and lasso provides the smallest loss function. Furthermore, the lasso that is one of the shrinking methods we are concerned on delivers the largest MCS p-value of 1, and the shrinking methods occupy the three maximum values of MCS p-value. Therefore, our MCS results are robust to CBOE VIX of out-of-sample evaluation periods.

5 Portfolio exercise

For investors, the economic value of predicting stock volatility is greater than the statistical significance of volatility predictability. According to Fleming et al. (2003) and Dai and Kang (2021), we take the view of maximizing investor utility, and use the volatility forecast as the key factor for portfolio optimization to calculate economic benefits. The reliability of the volatility forecast is reflected by the result of the investment portfolio performance. We refer to a great deal of literature on stock returns and stock volatility forecast (see, e.g., Neely et al., 2014; Wang et al., 2016; Ma et al., 2018; Chen and wang., 2021), from which we realize that a rational investor reasonably divides his/her wealth between stocks and risk-free assets. The total utility obtained by optimal allocation strategy for the mean-variance investor as follows:

$$\mathrm{L}_t(r_t) = E_t(\omega_t R_t + R_{t,f}) - \frac{1}{2}\gamma \,\mathrm{var}_t(\omega_t R_t + R_{t,f}) \tag{19}$$

where ω_t is the proportion of stocks in the total assets of investors, R_t and $R_{t,f}$ are the return in the stock market and the profits of risk-free investment, respectively. γ represents the highest risk coefficient for investors to accept asset investment. $E_t(\cdot)$ and var_t are, respectively, conditional expectation and variance of investors' asset allocation portfolio.

In order to maximize the total utility return of investors' assets, we go through the essential steps of determining ω_t.

$$\omega_t^* = \frac{1}{\gamma}\left(\frac{\hat{r}_{t+1}}{\hat{\sigma}_{t+1}^2}\right) \tag{20}$$

where $\hat{r}_{t+1}$ is the historical average of stock return at t+1-th month, and $\hat{\sigma}_{t+1}^2$ is the five-year rolling-window variance of stock return. The benchmark model is consistent with the literature of stock return forecasts, so the historical average of stock return in the t-th period is unchanged. It is the stock return volatility that determines the weight of the stock in the portfolio.

It is widely known that the value of risk aversion is closely related to the optimal weight. In this paper, a lower γ value means that the higher the cost of investment in the stock market. Therefore, γ is taken as 3, which not only ensures that the risk of investors is not high, but also obtain excellent investment return. According to the literature (e.g., Rapach et al., 2010; Neely et al., 2014), we set the weight of stocks in the portfolio range from 0 to 1.5, which is to prevent the effect of short selling on the asset return and to take precautions against exceeding 50% leverage.

Portfolio returns with the meaning of explanatory parameters are expressed as:

$$P_{t+1} = \omega_t^* R_{t+1} + R_{t+1,f} \tag{21}$$

The calculated portfolio is evaluated using the universally acknowledged certainty equivalent return:

$$CER_p = \hat{\mu}_p - \frac{\gamma}{2}\hat{\sigma}_p^2 \tag{22}$$

Insert Table 5 about here

Taking the autoregressive benchmark model as the standard, Table 4 forms the certainly equivalent return with market and investor sentiment indicators, macroeconomic fundamentals variables, technical indicators and oil volatility variables. When using recursive window in out-of-sample period, the portfolio for univariate regression based on four types of indicators utility gains are not more than the result of the shrinking method. For lasso, adaptive lasso, elastic net and ridge regression, the utility return increases to 6.826 after adding a type of variables that emotional indicators, macroeconomic variables, technical indicators and oil volatility variables into the autoregressive model. From a perspective of portfolio utility, Lasso, adaptive lasso, elastic net and ridge regression are more notable in improving the stock volatility forecasting model. The portfolio utility returns of macroeconomic fundamental variables outperform the other three varieties of variables in lasso, adaptive lasso, elastic net and ridge regression respectively for whole out-of-sample period. Generally, the empirical results of four shrinkage methods based on a series of variables indicate that plays a great role in the allocation of investors' assets compared with the univariate model.

6 Conclusion

In this paper, we advocate a set of predictive methods to forecast stock volatility with various indicators. There are following several findings. First, most of market and investor sentiment indicators, macroeconomic fundamentals, technical indicators and oil volatility predictors fail to predict stock return volatility in univariate regression. Elastic net has the best prediction ability of stock return volatility in MCS test, and the MCS p-value of most of them reaches 1. Second, and more importantly, we provide evidence that elastic net, adaptive lasso, lasso and ridge models have statistically significant predictive performance among the popular shrinkage methods. Hence, shrinking methods contribute to the improvement of stock return volatility. In the case of CBOE VIX as the explained variable, the performance of lasso is better than other models in the MCS test with rich data environment. Third, we find that the shrinkage strategies obtain better results than single predictors in portfolio. Therefore, lasso, adaptive lasso, ridge and elastic net are not only statistically significant, but also have greatly economic value. In summary, shrinkage methods always bring better benefits in a data-rich environment for stock volatility forecasting.

Acknowledgment: This research is supported by the National Natural Science Foundation of China granted: 71771030, 72131011, National Social Science Foundation of China granted: 21BGL111, Natural Science Foundation of Hunan Province granted: 2021JJ30025, and Fund of Hunan Provincial Education Department granted: 20B035.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. Abounoori, E., Elmi, Z. (Mila), & Nademi, Y. (2016). Forecasting Tehran stock exchange volatility; Markov switching GARCH approach. Physica A-Statistical Mechanics and Its Applications, 445, 264–282.
2. Aloui, C., Mabrouk, S. (2010). Value-at-risk estimations of energy commodities via long-memory, asymmetry and fat-tailed GARCH models. Energy Policy, 38(5), 2326–2339.
3. Andersen, T. G., Bollerslev, T., Diebold, F. X., Ebens, H. (2001). The distribution of realized stock return volatility. Journal of Financial Economics, 61(1), 43-76.
4. Audrino, F., Knaus, S. D. (2016). Lassoing the HAR model: A model selection perspective on realized volatility dynamics. Econometric Reviews, 35(8-10), 1485-1521.
5. Audrino, F., Huang, C., Okhrin, O. (2018). Flexible HAR model for realized volatility. Studies in Nonlinear Dynamics & Econometrics, 20170080.
6. Baker, M., Wurgler, J. (2007). Investor sentiment in the stock market. Journal of Economic Perspectives, 21(2), 129-152.
7. Baker, S. R., Bloom, N., Davis, S. J. (2016). Measuring Economic Policy Uncertainty. Quarterly Journal of Economics, 131(4), 1593-1636.
8. Christiansen, C., Schmeling, M., Schrimpf, A. (2012). A comprehensive look at financial

volatility prediction by economic variables. Journal of Applied Econometrics, 27(6), 956-977.

9. Chen, T., & Wang, S. (2021). Incomplete information model of credit default of micro and small enterprises. International Journal of Finance & Economics, 1-19
10. Corbet, S., Twomey, C. (2014). Have Exchange Traded Funds Influenced Commodity Market Volatility. International Journal of Economics and Financial Issues, 4(2), 323-335.
11. Dai, Z., Shi, R. (2021). Forecasting commodity prices with technical indicators: A partial least squares approach. International Review of Economics and Finance.
12. Dai, Z., Kang, J. (2021). Forecasting equity risk premium: A new method based on wavelet de-noising, 1-13. doi:10.1002/ijfe. 2400.
13. Degiannakis, S., Filis, G., Kizys, R. (2014). The effects of oil price shocks on stock market volatility: evidence from European data. The Energy Journal, 35(1), 35-56.
14. Elliott, G., Gargano, A., Timmermann, A. (2013). Complete subset regressions. Journal of Econometrics, 177(2), 357-373.
15. Engle, R. F., Rangel, J. G. (2008). The spline-garch model for low-frequency volatility and its global macroeconomic causes. Review of Financial Studies, 21(3), 1187-1222.
16. Engle, R. F., Ghysels, E., Sohn, B. (2013). Stock Market Volatility and Macroeconomic Fundamentals. The Review of Economics and Statistics, 95(3), 776-797.
17. Feng, J., Wang, Y., & Yin, L. (2018). Oil volatility risk and stock market volatility predictability: Evidence from G7 countries. Energy Economics, 68, 240-254.
18. Fleming, J., Kirby, C., Ostdiek, B. (2003). The economic value of volatility timing using “realized” volatility. Journal of Financial Economics, 67(3), 473-509.
19. Fuertes, A. M., Miffre, J., Rallis, G. (2010). Tactical Allocation in Commodity Futures Markets: Combining Momentum and Term Structure Signals. Journal of Banking and Finance, 34(10), 2530-2548.
20. Gong, X., Lin, B. (2018). Structural breaks and volatility forecasting in the copper futures market. Journal of Futures Markets, 38(3): 290-339.
21. Gong, X., Lin, B. (2020). Predicting the volatility of crude oil futures: The roles of leverage effects and structural changes. International Journal of Finance & Economics,

https://doi.org/10.1002/ijfe.2171.

22. Hansen, P. R., Lunde, A., Nason, J. M. (2011). The Model Confidence Set. Econometrica, 79(2), 453–497.
23. Huang, D., Jiang, F., Tu, J., Zhou, G. (2015). Investor sentiment aligned: A powerful predictor of stock returns. Review of Financial Studies, 28(3), 791-837.
24. Jiang, F., Lee, J. A., Martin, X., Zhou, G. (2019). Manager Sentiment and Stock Returns. Journal of Financial Economics, 132(1), 126-149.
25. Jurado, K., Ludvigson, S. C., Ng, S. (2015). Measuring uncertainty. American Economic Review, 105(3), 1177-1216.
26. Kang, S. H., Kang, S. M., Yoon, S.M. (2009). Forecasting volatility of crude oil markets. Energy Economics, 31(1), 119-125.
27. Kawabata, T. Private governance schemes for green bond standard: influence on public authorities' policy making. Green Finance, 2020, 2(1): 35-54.
28. Li, J., Tsiakas, I., Wang, W. (2015). Predicting Exchange Rates Out of Sample: Can Economic Fundamentals Beat the Random Walk? Journal of Financial Econometrics, 13(2), 293-341.
29. Li, J., Tsiakas, I. (2017). Equity premium prediction: The role of economic and statistical constraints. Journal of Financial Markets, 36, 56-75.
30. Liu, T., Gong, X. (2020). Analyzing time-varying volatility spillovers between the crude oil markets using a new method. Energy Economics, 87, 104711.
31. Latunde, T., Shina L., Damilola, A. Analysis of capital asset pricing model on Deutsche bank energy commodity. Green Finance, 2020, 2(1): 20-34
32. Liu, L., Zhang, T. (2015). Economic policy uncertainty and stock market volatility. Finance Research Letters, 15, 99-105.
33. Liu, J., Ma, F., Yang, K., Zhang, Y. (2018). Forecasting the oil futures price volatility: Large jumps and small jumps. Energy Economics, 72, 321–330.
34. Long, J. B. D., Shleifer, A., Summers, L. H., Waldmann, R. J. (1990). Noise Trader Risk in Financial Markets. Journal of Political Economy, 98(4), 703-738.
35. Ma, F., Liu, J., Wahab, M. I. M., Zhang, Y. (2018). Forecasting the aggregate oil price volatility in a data-rich environment. Economic Modelling, 72, 320-332.

36. Moskowitz, T. J., Ooi, Y. H., Pedersen, L. H. (2012). Time Series Momentum. Journal of Financial Economics, 104(2), 228-250.
37. Narayan, P. K., Narayan, S. (2007). Modelling oil price volatility. Energy Policy, 35(12), 6549-6553.
38. Nardi, Y., Rinaldo, A. (2011). Autoregressive process modeling via the lasso procedure. Journal of Multivariate Analysis, 102(3), 528-549.
39. Neely, C. J., Rapach, D. E., Tu, J., Zhou, G. (2014). Forecasting the Equity Risk Premium: The Role of Technical Indicators. Management Science, 60(7), 1772-1791.
40. Nonejad, N. (2017). Forecasting aggregate stock market volatility using financial and macroeconomic predictors: Which models forecast best, when and why? Journal of Empirical Finance, 42, 131-154.
41. Paye, B. S. (2012). Deja Vol: Predictive Regressions for Aggregate Stock Market Volatility Using Macroeconomic Variables. Journal of Financial Economics, 106(3), 527-546.
42. Rapach, D. E., Strauss, J. K., Zhou, G. (2010). Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy. Review of Financial Studies, 23(2), 821-862.
43. Schwert, G. W. (1989). Why Does Stock Market Volatility Change over Time. Journal of Finance, 44(5), 1115-1153.
44. Sullivan, R., Timmermann, A. G., White, H. L. (1999). Data-Snooping, Technical Trading Rule Performance, and the Bootstrap. Journal of Finance, 54(5), 1647-1691.
45. Szakmary, A. C., Shen, Q., Sharma, S. C. (2010). Trend-following trading strategies in commodity futures: A re-examination. Journal of Banking and Finance, 34(2), 409-426.
46. Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. Journal of the Royal Statistical Society Series b-Methodological, 58(1), 267-288.
47. Wang, Y., Ma, F., Wei, Y., Wu, C. (2016). Forecasting realized volatility in a changing world: A dynamic model averaging approach. Journal of Banking and Finance, 64(3), 136-149.
48. Wang, Y., Wei, Y., Wu, C., Yin, L. (2018). Oil and the short-term predictability of stock return volatility. Journal of Empirical Finance, 47, 90-104.

49. Welch, I., Goyal, A. (2008). A Comprehensive Look at the empirical performance of equity premium prediction. Review of Financial Studies, 21(4), 1455-1508.
50. Xu, L., Yin, X., Zhao, J. (2016). Differently Motivated ETF Trading Activities and the Volatility of the Underlying Index. Social Science Research Network.
51. Xu, L., Yin, X. (2017). Exchange Traded Funds and Stock Market Volatility. International Review of Finance, 17(4), 525-560.
52. Zhang, Y., Ma, F., Wang, Y. (2019a). Forecasting crude oil prices with a large set of predictors: Can LASSO select powerful predictors? Journal of Empirical Finance, 54, 97-117.
53. Zhang, Y., Wei, Y., Zhang, Y., Jin, D. (2019b). Forecasting oil price volatility: Forecast combination versus shrinkage method. Energy Economics, 80, 423-433.
54. Zou, H., Hastie, T. (2005). Regularization and variable selection via the elastic net. Journal of The Royal Statistical Society Series B-Statistical Methodology, 67(2), 301-320.
55. Zou, H. (2006). The adaptive lasso and its oracle properties. Journal of the American Statistical Association, 101(476), 1418-1429.

Table 1: Summary statistics

Panel A: Monthly realized variance summary statistics									
Log（RV）	Mean	Std.dev.	Ske	Kur					
	0.003	0.005	7.051	67.277					
Panel B: Predictor summary statistics									
Predictor	Mean	Std.dev.	Ske	Kur	Predictor	Mean	Std.dev.	Ske	Kur
Panel C: Uncertainty and sentiment (UMS)					Panel E: Technical (TI) indicators				
UI	0.639	0.086	2.379	7.495	MA(1,9)	0.757	0.430	-1.196	-0.569
EPU	113.394	42.793	1.181	1.422	MA(2,9)	0.783	0.413	-1.376	-0.107
SENT	-6.577	8.664	-1.298	6.192	MA(3,9)	0.769	0.422	-1.273	-0.378
PDND	24.938	20.008	1.150	0.992	MA(1,12)	0.777	0.417	-1.334	-0.220
NIPO	6.996	4.260	-0.397	-0.611	MA(2,12)	0.774	0.419	-1.314	-0.275
CEFD	20.158	7.725	2.472	9.183	MA(3,12)	0.777	0.417	-1.334	-0.220
EMV	0.108	0.625	1.742	4.419	MOM(1)	0.644	0.480	-0.601	-1.639
Panel D: Macroeconomic (MF) fundamentals					MOM(2)	0.671	0.471	-0.726	-1.473
DP	-3.934	0.256	0.080	-0.201	MOM(3)	0.691	0.463	-0.829	-1.313
DY	-3.928	0.257	0.048	-0.177	MOM(6)	0.745	0.437	-1.123	-0.739
EP	-3.125	0.351	-2.245	7.413	MOM(9)	0.780	0.415	-1.355	-0.165
DE	-0.809	0.413	2.788	10.406	MOM(12)	0.789	0.408	-1.419	0.013
BM	0.282	0.077	-0.062	-0.413	OBV(1,9)	0.484	0.500	0.065	-1.996
NTIS	0.005	0.021	-0.492	0.040	OBV(2,9)	0.487	0.501	0.053	-1.997
TBL	2.569	2.110	0.142	-1.542	OBV(3,9)	0.439	0.497	0.245	-1.940
LTY	4.964	1.740	0.067	-0.987	OBV(1,12)	0.457	0.499	0.173	-1.970
LTR	0.650	2.897	0.038	2.295	OBV(2,12)	0.475	0.500	0.101	-1.990
TMS	2.395	1.263	-0.189	-0.969	OBV(3,12)	0.484	0.500	0.065	-1.996
DFY	0.948	0.397	3.191	13.742	Panel F: Oil indicators				
DFR	-0.001	1.642	-0.443	7.227	WTI Oil	0.013	0.019	7.197	75.147
INFL	0.189	0.325	-1.001	5.716	Brent Oil	0.011	0.017	8.350	99.002

Notes: This table reports summary statistics for the log monthly stock volatility (in percent) and 7 market and investor sentiment indicators and 13 macroeconomic variables, 18 technical indicators, 2 oil price volatility indicators.

Table 2: Out-of-sample results for MCS evaluation method

Predictors	Loss function	MAE	Predictors	Loss function	MSE
Panel A: Uncertainty and sentiment variables (UMS)					
UI	0.6551	0.0000	UI	0.6773	0.0000
EPU	0.6599	0.0000	EPU	0.7059	0.0000
SENT	0.6349	0.0000	SENT	0.6859	0.0000
PDND	0.6526	0.0000	PDND	0.6974	0.0000
NIPO	0.6545	0.0000	NIPO	0.6994	0.0000
CEFD	0.6540	0.0000	CEFD	0.7023	0.0000
EMV	0.5904	0.0000	EMV	0.5628	0.0000
Panel B: Macroeconomic (MF) fundamentals					
DP	0.6622	0.0000	DP	0.7177	0.0000
DY	0.6529	0.0000	DY	0.7082	0.0000
EP	0.6660	0.0000	EP	0.7156	0.0000
DE	0.6628	0.0000	DE	0.7121	0.0000
BM	0.6643	0.0000	BM	0.7279	0.0000
NTIS	0.6554	0.0000	NTIS	0.7125	0.0000
TBL	0.6433	0.0000	TBL	0.6938	0.0000
LTY	0.6497	0.0000	LTY	0.6964	0.0000
LTR	0.6393	0.0000	LTR	0.6922	0.0000
TMS	0.6508	0.0000	TMS	0.7044	0.0000
DFY	0.6658	0.0000	DFY	0.7081	0.0000
DFR	0.6023	0.0000	DFR	0.6151	0.0000
INFL	0.6508	0.0000	INFL	0.7008	0.0000
Panel C: Technical (TI) indicators					
MA(1,9)	0.5637	0.0012	MA(1,9)	0.5324	0.0069
MA(2,9)	0.5899	0.0000	MA(2,9)	0.5691	0.0000
MA(3,9)	0.6005	0.0000	MA(3,9)	0.5838	0.0000
MA(1,12)	0.6120	0.0000	MA(1,12)	0.6062	0.0000
MA(2,12)	0.6222	0.0000	MA(2,12)	0.6184	0.0000
MA(3,12)	0.6350	0.0000	MA(3,12)	0.6571	0.0000
MOM(1)	0.5552	0.0087	MOM(1)	0.5216	0.0245
MOM(2)	0.5255	0.0873	MOM(2)	0.4700	0.0989
MOM(3)	0.5474	0.0054	MOM(3)	0.4924	0.0259
MOM(6)	0.6051	0.0000	MOM(6)	0.5950	0.0000
MOM(9)	0.6170	0.0000	MOM(9)	0.6333	0.0000
MOM(12)	0.6334	0.0000	MOM(12)	0.6528	0.0000
OBV(1,9)	0.5896	0.0000	OBV(1,9)	0.5949	0.0000
OBV(2,9)	0.6020	0.0000	OBV(2,9)	0.6315	0.0000
OBV(3,9)	0.6143	0.0000	OBV(3,9)	0.6377	0.0000
OBV(1,12)	0.6355	0.0000	OBV(1,12)	0.6717	0.0000
OBV(2,12)	0.6385	0.0000	OBV(2,12)	0.6750	0.0000

OBV(3,12)	0.6447	0.0000	OBV(3,12)	0.6831	0.0000
Panel D: Oil indicators					
WTI Oil	0.6424	0.0000	WTI Oil	0.6753	0.0000
Brent Oil	0.6458	0.0000	Brent Oil	0.6881	0.0000
Panel E: Shrinkage models					
Net	0.4760	1.0000	Net	0.3750	1.0000
Lasso	0.4792	0.7655	Lasso	0.3832	0.7434
Adaptive	0.4834	0.2749	Adaptive	0.3906	0.3143
Ridge	0.4765	0.9511	ridge	0.3787	0.7809

Notes: Out-of-sample forecasting results of shrinkage methods and univariate model including the value of loss functions and the MCS p-values. Considering two loss functions, MSE and MAE. Adaptive of Panel E refers to the adaptive lasso model.

Table 3: MCS results for alternative GRACH model

Predictors	Loss	MAE	Predictors	Loss	MSE
Panel A: Uncertainty and sentiment variables (UMS)					
UI	0.6551	0.0000	UI	0.6773	0.0000
EPU	0.6599	0.0000	EPU	0.7059	0.0000
SENT	0.6349	0.0000	SENT	0.6859	0.0003
PDND	0.6526	0.0000	PDND	0.6974	0.0000
NIPO	0.6545	0.0000	NIPO	0.6994	0.0000
CEFD	0.6540	0.0000	CEFD	0.7023	0.0000
EMV	0.5904	0.0000	EMV	0.5628	0.0005
Panel B: Macroeconomic (MF) fundamentals					
DP	0.6622	0.0000	DP	0.7177	0.0000
DY	0.6529	0.0000	DY	0.7082	0.0003
EP	0.6660	0.0000	EP	0.7156	0.0000
DE	0.6628	0.0000	DE	0.7121	0.0000
BM	0.6643	0.0000	BM	0.7279	0.0000
NTIS	0.6554	0.0000	NTIS	0.7125	0.0000
TBL	0.6433	0.0000	TBL	0.6938	0.0003
LTY	0.6497	0.0000	LTY	0.6964	0.0000
LTR	0.6393	0.0000	LTR	0.6922	0.0000
TMS	0.6508	0.0000	TMS	0.7044	0.0000
DFY	0.6658	0.0000	DFY	0.7081	0.0000
DFR	0.6023	0.0000	DFR	0.6151	0.0005
INFL	0.6508	0.0000	INFL	0.7008	0.0000
Panel C: Technical (TI) indicators					
MA(1,9)	0.5637	0.0014	MA(1,9)	0.5324	0.0061
MA(2,9)	0.5899	0.0000	MA(2,9)	0.5691	0.0005
MA(3,9)	0.6005	0.0000	MA(3,9)	0.5838	0.0005
MA(1,12)	0.6120	0.0000	MA(1,12)	0.6062	0.0005
MA(2,12)	0.6222	0.0000	MA(2,12)	0.6184	0.0003
MA(3,12)	0.6350	0.0000	MA(3,12)	0.6571	0.0000
MOM(1)	0.5552	0.0094	MOM(1)	0.5216	0.0217
MOM(2)	0.5255	0.0889	MOM(2)	0.4700	0.0911
MOM(3)	0.5474	0.0074	MOM(3)	0.4924	0.0221
MOM(6)	0.6051	0.0000	MOM(6)	0.5950	0.0005
MOM(9)	0.6170	0.0000	MOM(9)	0.6333	0.0005
MOM(12)	0.6334	0.0000	MOM(12)	0.6528	0.0000
OBV(1,9)	0.5896	0.0000	OBV(1,9)	0.5949	0.0005
OBV(2,9)	0.6020	0.0000	OBV(2,9)	0.6315	0.0005
OBV(3,9)	0.6143	0.0000	OBV(3,9)	0.6377	0.0005
OBV(1,12)	0.6355	0.0000	OBV(1,12)	0.6717	0.0000
OBV(2,12)	0.6385	0.0000	OBV(2,12)	0.6750	0.0000
OBV(3,12)	0.6447	0.0000	OBV(3,12)	0.6831	0.0000
Panel D: Oil indicators					

WTI Oil	0.6424	0.0000	WTI Oil	0.6753	0.0000
Brent Oil	0.6458	0.0000	Brent Oil	0.6881	0.0000
Panel E: Shrinkage models					
Net	0.4760	1.0000	Net	0.3750	1.0000
Lasso	0.4792	0.7689	Lasso	0.3832	0.7439
Adaptive	0.4834	0.2991	Adaptive	0.3906	0.2921
Ridge	0.4765	0.9420	Ridge	0.3787	0.7747
Panel F: models					
GRACH	0.5922	0.0003	GRACH	0.5866	0.0021

Notes: Out-of-sample forecasting results of shrinkage methods, GRACH model and univariate model including the value of loss functions and the MCS p-values. Considering two loss functions, MSE and MAE. Adaptive of Panel E refers to the adaptive lasso model.

Table 4: MCS results for alternative CBOE VIX

Predictors	Loss	MAE	Predictors	Loss	MSE
Panel A: Uncertainty and sentiment variables (UMS)					
UI	0.6135	0.0001	UI	0.5987	0.0010
EPU	0.6090	0.0001	EPU	0.5958	0.0010
SENT	0.6068	0.0001	SENT	0.5930	0.0010
PDND	0.6081	0.0001	PDND	0.5906	0.0010
NIPO	0.6043	0.0003	NIPO	0.5930	0.0010
CEFD	0.6118	0.0001	CEFD	0.5978	0.0010
EMV	0.6069	0.0001	EMV	0.5941	0.0010
Panel B: Macroeconomic (MF) fundamentals					
DP	0.6108	0.0001	DP	0.5999	0.0010
DY	0.6068	0.0003	DY	0.5988	0.0010
EP	0.6108	0.0001	EP	0.6037	0.0010
DE	0.6113	0.0001	DE	0.5950	0.0010
BM	0.6055	0.0003	BM	0.5990	0.0010
NTIS	0.6066	0.0003	NTIS	0.5988	0.0010
TBL	0.6118	0.0001	TBL	0.5966	0.0010
LTY	0.6153	0.0001	LTY	0.5815	0.0010
LTR	0.6054	0.0001	LTR	0.5916	0.0010
TMS	0.6075	0.0003	TMS	0.5960	0.0010
DFY	0.6124	0.0001	DFY	0.5978	0.0010
DFR	0.6056	0.0003	DFR	0.5896	0.0010
INFL	0.6130	0.0001	INFL	0.5961	0.0010
Panel C: Technical (TI) indicators					
MA(1,9)	0.5989	0.0003	MA(1,9)	0.5826	0.0010
MA(2,9)	0.6086	0.0001	MA(2,9)	0.5987	0.0010
MA(3,9)	0.6110	0.0001	MA(3,9)	0.5981	0.0010
MA(1,12)	0.6133	0.0001	MA(1,12)	0.6009	0.0010
MA(2,12)	0.6106	0.0001	MA(2,12)	0.5958	0.0010
MA(3,12)	0.6116	0.0001	MA(3,12)	0.5945	0.0010
MOM(1)	0.5410	0.4944	MOM(1)	0.5020	0.4223
MOM(2)	0.5477	0.4588	MOM(2)	0.5097	0.3250
MOM(3)	0.5695	0.0500	MOM(3)	0.5475	0.0113
MOM(6)	0.6073	0.0001	MOM(6)	0.5900	0.0010
MOM(9)	0.6105	0.0001	MOM(9)	0.5957	0.0010
MOM(12)	0.6122	0.0001	MOM(12)	0.5963	0.0010
OBV(1,9)	0.6006	0.0003	OBV(1,9)	0.5870	0.0010
OBV(2,9)	0.6008	0.0003	OBV(2,9)	0.5877	0.0010
OBV(3,9)	0.6047	0.0001	OBV(3,9)	0.5918	0.0010
OBV(1,12)	0.6092	0.0001	OBV(1,12)	0.5955	0.0010
OBV(2,12)	0.6092	0.0001	OBV(2,12)	0.5921	0.0010
OBV(3,12)	0.6083	0.0001	OBV(3,12)	0.5904	0.0010
Panel D: Oil indicators					

WTI Oil	0.6119	0.0001	WTI Oil	0.5940	0.0010
Brent Oil	0.6127	0.0001	Brent Oil	0.5924	0.0010
Panel E: Shrinkage models					
Net	0.5258	0.7494	Net	0.4690	0.8880
Lasso	0.5235	1.0000	Lasso	0.4674	1.0000
ALR	0.5295	0.4944	ALR	0.4761	0.4223
ridge	0.5286	0.7494	ridge	0.4720	0.8880

Notes: Out-of-sample forecasting results of shrinkage methods and univariate model to predict CBOE VIX including the value of loss functions and the MCS p-values. Considering two loss functions, MSE and MAE. Adaptive of Panel E refers to the adaptive lasso model.

Table 5: Portfolio performance.

Predictor	CER gain	CER gain (50 bps)	Predictor	CER gain	CER gain (50 bps)
UI	0.676	0.722	MA(1,9)	4.307	4.296
EPU	-0.147	-0.129	MA(2,9)	3.540	3.598
SENT	0.451	0.485	MA(3,9)	2.559	2.625
PDND	-0.139	-0.147	MA(1,12)	2.193	2.254
NIPO	0.188	0.204	MA(2,12)	1.978	2.064
CEFD	0.314	0.348	MA(3,12)	1.672	1.741
EMV	0.191	0.238	MOM(1)	6.208	5.663
DP	0.043	0.049	MOM(2)	5.972	5.663
DY	0.110	0.133	MOM(3)	4.674	4.487
EP	0.618	0.650	MOM(6)	2.789	2.835
DE	-0.339	-0.352	MOM(9)	1.728	1.724
BM	-0.522	-0.511	MOM(12)	1.418	1.444
NTIS	-0.286	-0.239	OBV(1,9)	2.935	2.855
TBL	-0.309	-0.336	OBV(2,9)	2.762	2.740
LTY	0.355	0.418	OBV(3,9)	1.023	1.067
LTR	0.194	0.182	OBV(1,12)	1.136	1.198
TMS	-0.574	-0.622	OBV(2,12)	0.550	0.544
DFY	-0.230	-0.237	OBV(3,12)	0.197	0.202
DFR	0.551	0.280	WTI Oil	0.439	0.410
INFL	-0.028	-0.042	Brent Oil	0.307	0.285
Panel B: shrinkage approaches					
Net-UMS	0.665	0.804	Adaptive -UMS	0.607	0.722
Net-MF	0.595	0.368	Adaptive -MF	0.719	0.507
Net-TI	6.943	6.708	Adaptive -TI	6.869	6.666
Net-Oil	0.259	0.273	Adaptive -Oil	0.248	0.225
Net-ALL	6.367	6.271	Adaptive -ALL	6.240	6.168
LR-UMS	0.431	0.545	RR-UMS	0.794	1.037
LR-MF	0.807	0.610	RR-MF	0.564	0.303
LR-TI	6.831	6.629	RR-TI	6.826	6.677
LR-Oil	0.222	0.196	RR-Oil	0.254	0.439
LR-ALL	6.264	6.178	RR-ALL	6.016	5.972

Notes: Performance of portfolio. This table provides the performance of portfolio formed by volatility forecasts from 2001 to 2018. Each period a mean–variance investor allocates wealth between stock and T-bill based on return and volatility forecasts. In this framework, we use the popular historical average return forecasts and use $\gamma = 3$. The optimal weight of stock is restricted between 0 and 1.5. The portfolio performance is evaluated based on certainty equivalent return (CER). Adaptive of Panel B refers to the adaptive lasso model.

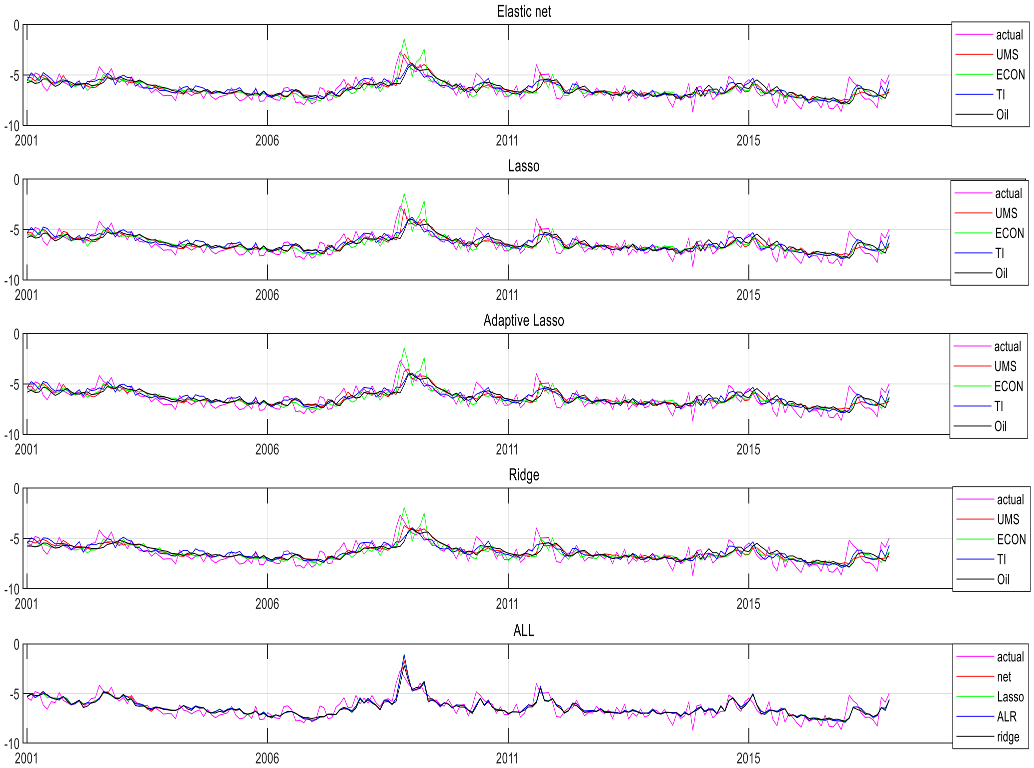

Figure 1: Stock volatility based on shrinkage regressions and recursive windows. “actual”, represents true stock volatility, others indicators names correspond to shrinkage regression forecasts from recursive windows.