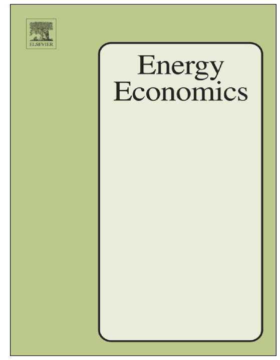
Accepted Manuscript



Forecasting the sign of U.S. oil and gas industry stock index excess returns employing macroeconomic variables

Jingzhen Liu, Alexander Kemp

PII: S0140-9883(19)30137-9
DOI: https://doi.org/10.1016/j.eneco.2019.04.023
Reference: ENEECO 4377

To appear in: *Energy Economics*

Received date: 6 August 2017
Revised date: 15 January 2019
Accepted date: 26 April 2019

Please cite this article as: J. Liu and A. Kemp, Forecasting the sign of U.S. oil and gas industry stock index excess returns employing macroeconomic variables, Energy Economics, https://doi.org/10.1016/j.eneco.2019.04.023

Forecasting the sign of U.S. oil and gas industry stock index excess returns employing macroeconomic variables

Jingzhen Liu[a,*], Alexander Kemp[a,*]

[a] *Aberdeen Centre for Research in Energy Economics and Finance, University of Aberdeen, Aberdeen, AB24 3QY, UK*

Abstract

In this study we propose a method of selecting the macroeconomic variables for forecasting the excess return signs of the U.S. oil and gas industry stock index by combining the Forward Sequential Variable Selection Algorithm and information criteria. We select predictors from a large monthly macroeconomic variable dataset designed by McCracken and Ng (2015). The method can adapt to the updated macroeconomic information and the possible time-varying relationship between the macroeconomic variables and the stock return signs. We also propose a method which can change the threshold value of the probit model automatically for considering the potential time-varying risk aversion level of the market participants. Further, we investigate the investment performance of an active trading strategy based on our forecasting model and compare it with a passive buy-and-hold trading strategy for different time periods.

Our study is important for both oil and gas industry investors and U.S. energy policy makers. The method that we used in this study offers a solution to the issue of selecting useful information from large datasets and absorbing updated market information.

Keywords: Excess stock return; U.S. Oil and gas industry; Probit model; Market timing; Big data
JEL codes: C53; C55; C58; G11; G17; E00

1. Introduction

In recent years, one frequently discussed topic in empirical finance studies is whether stock returns can be predicted. High forecasting accuracy of stock markets will help investors to increase

*Corresponding authors
Email addresses: jingzhen_liu1990@hotmail.com (Jingzhen Liu), a.g.kemp@abdn.ac.uk (Alexander Kemp)

their wealth management ability and also help policy makers when making industry-related policies. For the implementation of investment, the sign prediction of stock returns is important for asset allocation. Merton (1981) held the view in his market timing model that mutual fund managers concentrate more on the sign of return rather than the overall returns. Moreover, because the noise of the observed returns is too high to predict the overall return accurately, some studies found that the direction of stock returns is predictable to some extent. Specifically, Hong and Chung (2003) proposed a model-free statistical procedure to check whether the direction of change in 5 daily U.S. stock indices is predictable using historical self-information. Their empirical findings show that excess stock returns can be forecasted by using the information on historical excess stock returns such as volatility, skewness and kurtosis. Christoffersen and Diebold (2006) showed that volatility dependence produces sign dependence of past and future stock returns, and that it is statistically possible to have sign dependence even without conditional mean dependence. Thus, the signs of stock returns are perhaps predictable even when the return levels are unpredictable. We investigate whether macroeconomic variables have predictive power for future stock excess return signs for several reasons. First, many fundamentalist and long-term stock traders use new-found macroeconomic news to predict future stock performance and make investment decisions. Second, macroeconomic indicators to a great extent reflect past economic conditions and stimulate policy makers to make adjustments. The information about macroeconomic policy adjustments can also be used by stock market investors. Therefore, there should exist a channel for delivering information from the macroeconomic level to the stock market. Macroeconomic variables, including monetary policy variables and macro variables related to economic activity, are investigated in many studies of the predictability of stock markets.

Even if stock returns are supposed to be related to macroeconomic variables, they do not seem to add forecasting powers for stock returns superior to that provided by well-known financial ratios such as dividend yields, price-earnings ratios, book-to-market ratios, or short-term interest rates, and yield spreads. Çakmaklı and van Dijk (2016) summarized several complicating issues relating to the predictive powers of macroeconomic variables for stock returns. First, investors may consider a large macroeconomic variable set rather than selected variables which only contain limited information. Second, sound statistical reasons confirm that the predictive model will become more unstable when more variables are added. Third, the relationship between stock market performance and selected macroeconomic variables may be time-varying, which results in high instability of the predictive model. Thus, they conducted principle analysis on a large set of macroeconomic variables, and built a factor-augmented regression model for predicting monthly excess stock returns. However, it may not be necessary to use all macroeconomic information to forecast the future stock performance of specific industries, because different industries are exposed to different risks. For example, the oil price is a key risk factor for the oil and gas industry but not as important for other industries (see Moya-Martínez et al. (2014) and Shaeri et al. (2016)). Furthermore, the principle component analysis method is difficult to conduct economic inference, which confuses practitioners.

In this paper we study the predictive ability of macroeconomic variables for the signs of the excess returns of the U.S. oil and gas industry stock index. The oil and gas industry has been favoured by investors in recent years due to increasing oil prices during the period between 2008 and 2014. The number of mutual funds and exchange traded funds that invest in oil and gas industry companies also increased during this period (Ramos and Veiga (2011)). In developed countries such as the U.S. and U.K., and oil-exporting countries such as Canada and Brazil, the oil and gas industry occupies a large market capitalisation. However, the high volatility of oil and gas stocks has been very evident in the recent market turmoil, and more fundamental macroeconomic issues require consideration. Moreover, this industry is important because oil, as a major production input, determines costs of other industries by affecting input prices.

There are three main contributions of this paper. First, we focus on the out-of-sample forecasting ability of monthly macroeconomic variables and consider filtering information and selecting important predictors when predicting stock return signs. Second, we consider the possibility that the importance level of the macroeconomic information varies in different time periods and the potential time-varying relationship of macroeconomic variables and stock prices mentioned by Çakmaklı and van Dijk (2016). Our variable selection and prediction process are dynamic and can adapt to time-varying market conditions. Third, we propose a method which can automatically adjust the threshold value of the probit model. This practice removes the assumption that the risk aversion of the market participants is constant. To our knowledge, our paper is the first one that investigates a large set of macroeconomic variables constructed by McCracken and Ng (2015) for forecasting oil and gas stock index future signs. The procedure for selecting important predictors in this study can also be readily implemented for other industries, which is important for stock investors in terms of industry rotation strategy. The research on the out-of-sample prediction produces results which are important for both oil and gas industry investors and energy policy makers.

Other research, such as by Nyberg (2011), Pönkä (2017) and Pönkä (2016), used several predictors to forecast the composite stock index for different countries. In their research, the number and the combination of the predictors are fixed and they ignore the possibility that the important predictors may be different over time. They consider that the threshold value for stock return sign prediction is unchanged. The method that we propose can adapt to updated market information and account for the changing relationship between predictors and stock return signs. Moreover, we do not investigate the return sign prediction issue for the composite market index but for a specific industry. We obtain two novel findings. First, the importance of the macroeconomic variables for forecasting stock return signs may differ for different time periods. Second, changing the threshold value of the probit model can possibly increase the forecasting accuracy.

The rest of the paper is organized as follows. In Section 2, we discuss previous studies relating to the factors determining oil and gas industry stock returns and relevant research methodologies. In Section 3, we describe the static and dynamic probit models as well as the Forward Sequential Selection algorithm. The goodness-of-fit measures and statistical tests used for assessing sign forecasts are also described in this section. In Section 4 we introduce the dataset of the macroeconomic variables. Section 5 presents the empirical results. The conclusions are discussed in Section 6.

2. Literature Review

Three strands of the literature relate to our study. First, we review the literature about using macroeconomic variables to predict stock market performance. Second, we review the literature on the determinants of stock market returns of oil and gas companies. Third, we review the literature on the models and methodologies for forecasting the sign of stock returns.

2.1. Macroeconomic variables and stock market returns

Several studies indicate that some macroeconomic variables have predictive powers for stock returns. Lettau and Ludvigson (2001) found that fluctuations in the aggregate consumption-wealth ratio can be used for predicting both U.S. quarterly real stock returns and excess stock returns. They also found that this ratio is a better forecaster of future returns over short and intermediate horizons compared to the dividend yield and the dividend payout ratio. Piazzesi et al. (2007) considered a consumption asset pricing model which takes the composition risk relating to fluctuations in the relative share of housing into consideration, and suggested that the expenditure share of housing has predictive power for U.S. stock returns. Gomes et al. (2007) and Campbell and Diebold (2009) respectively found that both expenditures on durables as a fraction of its stock and survey-based measures of expected business conditions have predictive powers for stock returns. Recently, Çakmaklı and van Dijk (2016) used principle component analysis on a large set of macroeconomic variables and built a factor-augmented regression model for predicting monthly excess stock returns. Empirical results show that out-of-sample forecasting accuracy is significantly improved by adding macroeconomic information compared with the benchmark model which only includes valuation ratios and interest-related variables.

2.2. Determinants of stock returns of oil and gas companies

Sadorsky (2001), Boyer and Filion (2007), Sadorsky (2008), Ramos and Veiga (2011) and Bianconi and Yoshino (2014) found that on average oil price changes have simultaneous positive effects on the stock returns of oil and gas companies. They argued that changes in oil price, interest rate, and foreign exchange rates are other systematic risk factors except aggregate market price risk for oil and gas companies. The oil price determines both profits and operating costs of oil and gas companies, long-term interest rates relate to the investment costs, and the foreign exchange rate determines the input costs as well as profits, since the oil price is referenced by the U.S. dollar. Sadorsky (2001) suggested that an increase in the Canada-US exchange rates ($US/$C) or the term premium calculated as the premium between the annual yield on the 90-day Canadian Treasury bill and the yield on the 30-day Canadian Treasury bill decreases the returns to Canadian oil and gas stock prices. Boyer and Filion (2007) found that stock market returns of Canadian oil and gas companies are significantly negatively related to interest rates and the Canadian dollar exchange rate against the U.S. dollar. Sadorsky (2008) investigated the impact that global oil market risk factors have on the oil price risk of oil company stock prices, and confirmed that oil prices and market risk are both positive and statistically significant risk factors. Mohanty and Nandha (2011) studied the oil price risk exposures of the U.S. oil and gas sector using the Fama-French-Carhart four-factor asset pricing model augmented with oil price and interest rate factors. Their evidence shows that oil price changes are still statistically important for U.S. oil and gas sector companies even after controlling for the effects of market, book-to-market, and size factors, as well as the momentum factor. Ramos and Veiga (2011) explored the risk factors for oil and gas stock indices in 34 countries, and suggested that the oil and gas sector responds more strongly to oil price changes in developed countries than in emerging markets. Moreover, they also found that oil price rises have a greater impact than oil price falls.

Lanza et al. (2005) modelled the determinants of the long-run dynamics of six major oil companies' stocks by exploring major financial variables including stock market indices, foreign exchange rates, and spot and future oil prices, over the period January 1998 to April 2003. Empirical findings of their multivariate cointegration and vector error correction models (VECM) confirmed the statistical significance of the major financial variables in explaining the long-run dynamics of oil company stock values. Giovannini et al. (2006) studied financial risk factors and several integrated oil company stock price returns based on VECM and DCC-GARCH methods. The financial risk factors include stock market indices, foreign exchange rates, and the difference between the 12-month futures price and spot price on the Brent oil spread. They found that market indices and the spread in oil prices are endogenous for the selected integrated oil corporations. Diaz and de Gracia (2016) investigated the impact of oil price shocks on the stock returns of four oil and gas corporations listed on the NYSE over the period January 1974 to December 2015. In the short-run, they found a significant positive impact of oil price shocks on stock returns, with the relationship becoming statistically significant during the post-1986 period. Kang et al. (2017) explored the effects of oil price shocks and economic policy uncertainty on the stock returns of oil and gas corporations. Their results suggested that oil demand-side shocks have a positive impact on stock returns, while policy uncertainty shocks have a negative impact. They also found that a well-diversified portfolio of oil companies is obtainable because their individual returns respond differently to structural shocks.

2.3. Methodologies of predicting the sign of stock excess returns

Another strand of the literature relates to the methodologies of predicting the sign of stock excess returns. The traditional linear classifier such as the static logit and probit models were considered by Leung et al. (2000). They found that the linear classifier performs better than level forecasting methods such as exponential smoothing, multivariate transfer function, vector autoregression with Kalman filter, and multilayered feedforward neural networks in forecasting stock index signs. Their input variables include short-term and long-term interest rates, lagged index returns, consumer price level, and industrial production. Their empirical evidence also indicates that the classification models can generate higher trading profits than level estimation models. Hong and Chung (2003), Rydberg and Shephard (2003), and Anatolyev and Gospodinov (2012) utilized the so-called autologistic model to forecast the stock return's direction. Specifically, Hong and Chung (2003) fit a MA(1)-Threshold-GARCH(1,1) model via maximum likelihood estimation (MLE) to the daily excess returns for five US stock indices, namely, DJIA, S&P500, NYSE, NASDAQ and S&P500F, respectively. Anatolyev and Gospodinov (2012) decomposed the observed returns of the NYSE/AMEX value-weighted index into a product of sign and absolute value components,and modeled the signs to capture the important nonlinearities in excess return dynamics. Their evidence suggests that, based on the autologistic model, this decomposition can improve out-of-sample sign forecasting for excess stock returns.

Nyberg (2011) and Nyberg et al. (2015) used dynamic and static bivariate probit models to investigate the sign predictability of U.S. stock returns. Recently, Çakmaklı and van Dijk (2016) employed a factor-augmented predictive regression model to forecast the return signs and level of the S&P 500 index. In summary, probit models and linear regression models are most commonly used by researchers for predicting stock returns. We employ the probit model in this study since it is easy to apply and interpret from an economic viewpoint.

3. Methodology

The aim of this study is to forecast the sign of the U.S. oil and gas industry stock index employing macroeconomic variables. We choose static and dynamic probit models which are commonly used for prediction. We then employ a sequential forward selection algorithm presented by Altinbas and Biskin (2015) for selecting relevant macroeconomic variables based on minimizing different kinds of information criteria.

3.1. Static and dynamic probit models

First, we transform the excess stock index return into a binary sign return indicator y_t which is used as the dependent variable:

$$y_t = \begin{cases} 1, & if\ the\ excess\ return\ is\ positive \\ 0, & otherwise, \end{cases} \tag{1}$$

where the excess return is the one-month excess return which is calculated by: $r_{e,t} = ln(\frac{S_t}{S_{t-1}}) - r_f$ where r_f is the one-month U.S. T-bill rate and S_t is the index price for month t. We specify a vector of explanatory variables as x_t, which in our case contains the selected lagged macroeconomic variables discussed in the next subsection. We denote the information set at time t as $\Omega_t = \sigma[(y_s, x_s), t \geq s]$ and y_t conditional on Ω_{t-1} being assumed to follow a Bernoulli distribution as follows:

$$y_t \mid \Omega_{t-1} \sim B(p_t). \tag{2}$$

The conditional expectation and probability based on information set Ω_{t-1} is specified as $E_{t-1}(\cdot)$ and P_{t-1} respectively. Then we have the conditional probability of y_t as follows:

$$p_t = E_{t-1}(y_t) = P_{t-1}(y = 1). \tag{3}$$

The conditional probability of positive excess stock returns p_t is assumed to follow a standard normal distribution which can be described as:

$$p_t = \Phi(\pi_t), \tag{4}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution, and π_t is a linear function of the lagged macroeconomic variables in Ω_{t-1}. The above model is a probit model and will be changed to a logit model if the logistic distribution is assumed instead of the standard normal distribution. McCullagh and Nelder (1989) indicated the strong similarity of logit and probit models. Hence we use the probit model since it was widely used in previous studies for sign prediction of stock excess returns. If we think of a static probit model, the latent variable π_t is only determined by the selected variables and specified as:

$$\pi_t = \omega + x_{t-1}^{'}\beta, \tag{5}$$

where x_{t-1} includes lagged values of the selected macroeconomic variables and ω is a constant. To investigate the impact of past signs on the signs of future stock excess returns, we also consider a dynamic probit model which adds the past signs into the explanatory variable set. In this case, the latent variable π_t is denoted as:

$$\pi_t = \omega + \delta y_{t-1} + x_{t-1}^{'}\beta. \tag{6}$$

Then the forecasting sign at time t is determined by

$$\hat{y}_t = \mathbf{1}[p_t > c], \tag{7}$$

where $\mathbf{1}[\cdot]$ is an indicator function and p_t is estimated by eqs. 5 and 6. All the parameters in the above two kinds of models can be estimated by the maximum loglikelihood estimation method.

3.2. Sequential forward selection method (SFS)

SFS is a bottom-up search method which starts with an empty set of macroeconomic variables. This subset is extended by adding new explanatory variables into the data set. At every iteration step, a new variable will be added into the subset if it increases the objective function more than other variables when it is used together with the variables in the subset. Therefore, at every iteration, only one variable is added to the subset. Let X and S represent the variable set and the selected variable subset respectively. Then we can describe the SFS process as follows:

Step 1: Iteration number is n=1: $S_n = \phi$ and $X_n = x_1, x_2, ..., x_N$

Step 2: Select the best variable, x_k, which increases the objective function more than other variables in the S_n after combining the selected variables that were chosen previously, $J(S_n \cup x_k)$. Then the new variable subset is updated by adding x_k to the present subset, namely, $S_{n+1} = S_n + x_k$, and x_k is deleted from the variable set X. Consequently, the new variable set is $X_{n+1} = X_n \setminus x_k$

Step 3: Repeat Step2

Step 4: Stop the iteration

The iteration number is updated as $n = n + 1$. If there does not exist a variable x_k which can increase the objective function, the SFS goes to Step 4, otherwise it goes to Step 3.

When the SFS stops, we assume that the most relevant macroeconomic variables are selected for the static and dynamic models. We choose Bayesian information criterion (BIC) as the objective function since it is widely used for econometric model selection in the literature on energy economic studies such as Nazlioglu and Soytas (2012) and Zagaglia (2010). Moreover, BIC is argued to be appropriate for selecting the ”true model” (i.e. the process that generated the data) from the set of candidate models (see Burnham and Anderson (2003), Vrieze (2012) and Aho et al. (2014)). Specifically, BIC will select the ”true model” with probability 1, as $n \to \infty$. We also investigate the out-of-sample forecasting performance by using Akaike and Hannan Quinn information criteria. The forecasting accuracy is lower than the forecasting accuracy when we use Bayesian information criterion. The definition of BIC is specified by the following equation:

$$BIC = -2ln(L) + kln(n), \tag{8}$$

where $ln(L)$ represents the maximized log-likelihood function value for the static and dynamic probit models specified by eq.5 and eq.6, respectively, k is the number of variables, and n is the sample size.

3.3. Threshold c selection

A natural threshold of $c = 0.5$ is usually selected for the stock return sign prediction task in the literature such as Nyberg (2011), Nyberg et al. (2015), Pönkä (2016) and Pönkä (2017). However, we consider a dynamic optimal threshold value which is determined every month. One reason is that a constant threshold value may not be optimal all the time for the binary classification task because we use the rolling window method so that different datasets are used for training the model. Another reason is that the threshold value may reflect the risk aversion level of oil and gas industry investors. For instance, if the number of positive returns is higher than the number of negative returns over a specific period, the best threshold value for classification may be lower than 0.5 because of the potential high risk aversion of investors during this period. In contrast, the period when the number of negative returns is higher than the positive returns indicates that the best threshold value for classification should be higher than 0.5. Also the investors’ risk aversion may be high during this period. Therefore, due to the different market states and economic conditions, the risk aversion may be time-varying. Thus, we consider a dynamic optimal threshold selection approach based on the Receiver Operating Characteristic (ROC) curve. The ROC curve is a graphical plot which demonstrates the classification ability of a binary classifier as its threshold value is changed. ROC analysis is usually used as a goodness-of-fit measure of classification accuracy in biostatistics and medical applications. However, it was introduced into economic applications by Berge and Jordà

(2011), Christiansen et al. (2014), and Nyberg et al. (2015). Based on the idea of signal prediction (eq. 10), we can define two measures of classification accuracy which are the true positive rate (TP) and the false positive rate (FP) coefficients:

$$TP(c) = P_{t-1}(p_t > c \mid y_t = 1), \tag{9}$$

$$FP(c) = P_{t-1}(p_t > c \mid y_t = 0), \tag{10}$$

for any threshold $0 \leq c \leq 1$.

The ROC curve is generated by drawing the true positive rate (eq. 9) against the false positive rate (eq.10) at various threshold (c) settings. A ROC curve beyond the 45-degree line suggests prediction accuracy superior to a coin toss, while the curves below the 45-degree line indicate perverse forecasts. The area under the ROC curve (AUC) is an alternative measure for evaluating the forecast performance. It is the integral of the ROC curve between zero and one. Thus, the AUC values are within the interval $[0, 1]$, and the value of 0.5 indicates a coin toss and the value of 1 represents perfect prediction. Therefore, we expect a value of AUC greater than 0.5, which indicates successful market timing ability. Figure 1 shows an example of the ROC curve. The ROC curve can also be used for determining the optimal operating point or the optimal threshold value for the classification task. This method accounts for the cost of misclassifying a negative class as a positive class and the cost of misclassifying a positive class as a negative one. To find the optimal operating point, we firstly find the slope on the ROC curve graph which considers both the classification accuracy and the costs of misidentification of negative returns and positive returns as follows:

$$S = \frac{Cost(P|N) - Cost(N|N)}{Cost(N|P) - Cost(P|P)} \times \frac{NN}{PN}, \tag{11}$$

where $Cost(P|N)$ is the cost of misclassifying a positive excess return as a negative one while $Cost(N|P)$ is the cost of misclassifying a negative excess return as a positive one, and NN and PN represent the observations of negative returns and positive returns in the sample. Here, we select the values of $Cost(P|N)$, $Cost(N|N)$, $Cost(N|P)$, and $Cost(P|P)$ as 0.5, 0, 0.5, and 0, which means that the costs are equal for misclassification of positive and negative excess returns. Then we find the optimal operating point by moving the straight line with slope S from the upper left corner of the ROC curve plot ($FP(c) = 0$, $TP(c) = 1$) down and to the right, until it intersects the

ROC curve.

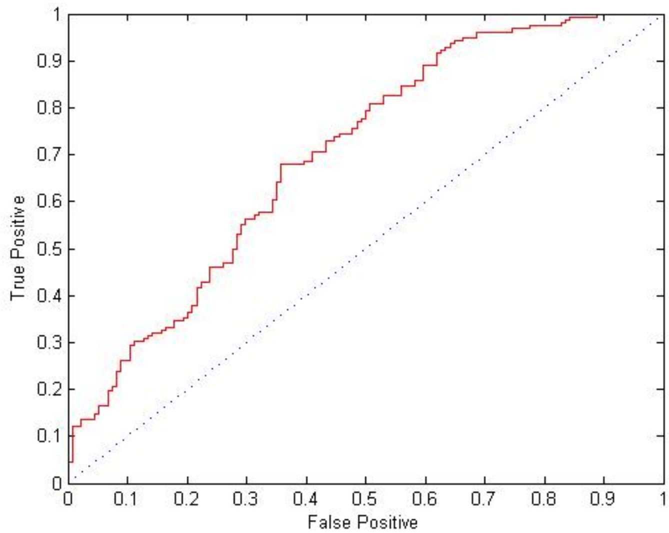


Figure 1: A sample of ROC curve plot

3.4. Goodness-of-fit measurement and sign forecasting

We use several measures to evaluate the out-of-sample forecasting ability of the prediction method that we propose.

One statistic that we consider for evaluating model performance is Quadratic Probability Score (QPS) which is defined as:

$$QPS = \frac{1}{T}\sum_{t=1}^{T} 2(y_t - p_t)^2 \tag{12}$$

where p_t is the conditional probability of positive stock excess returns estimated by the model at time t. This statistic is also widely used for evaluating probability predictions and can be viewed as a mean square error type of statistic for probit models. QPS values are within the interval $[0, 2]$ and score 0 represents perfect prediction.

Following previous studies such as Nyberg (2011) and Nyberg et al. (2015), we also report the success ratio (SR), which is defined as the percentage of correct forecasts. The results produced by the probit models are conditional probability forecasts of positive excess stock index returns in a given period. Therefore, we convert these probabilities into sign prediction signals based on a threshold by using the following equation:

$$\hat{y}_t = \mathbf{1}[p_t > c_t], \tag{13}$$

where $\mathbf{1}[\cdot]$ is an indicator function and p_t is estimated by eqs. 5 and 6. The threshold value at time t, c_t, is chosen using the approach described in Section 3.3. Given the threshold and the estimated forecasting signs, we can calculate the success ratio as follows:

$$SR = \frac{\hat{y}^{uu} + \hat{y}^{dd}}{\hat{y}^{uu} + \hat{y}^{du} + \hat{y}^{ud} + \hat{y}^{dd}}, \tag{14}$$

where the forecasts are grouped by the following characteristics:

$$\begin{aligned}
\hat{y}^{uu} &= \sum_{t=1}^{T} \mathbf{1}[\hat{y}_t = 1, y_t = 1], \\
\hat{y}^{ud} &= \sum_{t=1}^{T} \mathbf{1}[\hat{y}_t = 1, y_t = 0], \\
\hat{y}^{du} &= \sum_{t=1}^{T} \mathbf{1}[\hat{y}_t = 0, y_t = 1], \\
\hat{y}^{dd} &= \sum_{t=1}^{T} \mathbf{1}[\hat{y}_t = 0, y_t = 0],
\end{aligned}$$

In the above the superscript u is an ”up” signal and d is a ”down” signal. In addition, we can also calculate the success ratio of positive directions by

$$SR^u = \frac{\hat{y}^{uu}}{\hat{y}^{uu} + \hat{y}^{ud}} \tag{15}$$

and the success ratio of negative directions by

$$SR^d = \frac{\hat{y}^{dd}}{\hat{y}^{dd} + \hat{y}^{du}}. \tag{16}$$

Related to the success ratio, Pesaran and Timmermann (2009) have proposed a statistical test(denoted by PT) for evaluating the performance of directional forecasting accuracy allowing for serial correlation in y_t. Denote the forecasting signals by I^f, and E_N as a unit matrix. The statistic is defined by

$$PT = N \times Tr(S) \sim \chi^2(1), \tag{17}$$

where $I = (I_1, I_2, ..., I_N)^{'}$ which represents the real signal, $I^f = (I_1^f, I_2^f, ..., I_N^f)^{'}$, $M_\tau = E_N - \tau(\tau^{'}\tau)^{-1}\tau^{'}$, $\tau = (1, 1, ..., 1)^{'}$, N is the sample size, and S is for matrix trace. S can be calculated as follows:

$$\begin{aligned}
&S_{II^f} = S^{'}_{I^f I} = N^{-1} I^{'} M_\tau I^f, \\
&S_{II} = N^{-1} I^{'} M_\tau I, \\
&S_{I^f I^f} = N^{-1} I^{'}_f M_\tau I^f, \\
&S = S_{II}^{-1} S_{II^f} S_{I^f I^f}^{-1} S_{I^f I}.
\end{aligned}$$

The null hypothesis is I_i^f and I_i which are independent of each other, which indicates no directional prediction ability. Another independence test is based on the T statistic for a linear regression model. By using Bartlett weights as suggested by Newey and West (1987), we can construct the statistic as follows:

$$\tilde{t}_\beta = \frac{\hat{\beta}}{\sqrt{\hat{V}_{NW}(\hat{\beta})}}, \tag{18}$$

where $\hat{\beta}$ is the OLS estimator of the linear regression equation $I_i = \alpha + \beta I_i^f + u_i$, $i = 1, 2, ..., N$, $E(u_i \mid I_i^f, I_{i-1}^f, ...) = 0$.

3.5. Economic value investigation of the sign prediction

It was suggested by Leitch and Tanner (1991)and Cenesizoglu and Timmermann (2012) that statistically significant forecasting results usually do not indicate economic significance. Thus, a market timing ability test is required. Following Leung et al. (2000), Nyberg (2011) and Pönkä (2017), we consider a simple trading strategy based on the forecasted sign of excess stock index returns, and compare its out-of-sample performance with the passive buy-and-hold strategy. This will offer a direct investigation of whether investors can utilize past information of macroeconomic variables to make gains when they invest in the oil and gas industry. This can also be viewed as a check of market efficiency which argues that current stock prices reflect all the information and investors cannot use past information to outperform the market. In other words investors cannot construct active investment strategies by using past information to beat the buy-and-hold strategy.

For our trading strategy an investor is assumed to make a financial allocation between the oil and gas industry index and the one-month T-bill rate, namely a risky asset and a risk-free asset respectively. The weight of the stock index that the investor wants to put is represented as ω_t and the weight of the risk-free asset is denoted by $1-\omega_t$. Then, the return of the next period of the active trading strategy is calculated as, $r_{p,t+1} = \omega_t r_{index,t+1} + (1-\omega_t) r_{f,t+1}$, where $r_{index,t+1}$ is the realized stock index return in month $t+1$, and $r_{f,t+1}$ is the risk-free rate in month $t+1$. ω_t is dependent on the forecasted excess return sign of the stock index, namely, $\hat{y}_t$. If $\hat{y}_t = 1$, $\omega_t = \omega^{up}$, otherwise, if $\hat{y}_t = 0$, $\omega_t = \omega^{down}$. We consider two alternative scenarios. The first is more restrictive and does not allow for short-sales while under the second short-sales are allowed. We assume that ω^{up} is within the interval $[0.1, 1.5]$, while ω^{down} is within the range $[-1.0, 0]$. We consider different combinations of values of ω^{down} and ω^{up} to offer a comprehensive economic analysis.

Regarding the evaluation of the performance of the trading strategies, we use mean, standard deviation, annualized mean, the Sharpe ratio and the economic performance measure (EPM). The Sharpe ratio is the most widely used measure of risk-adjusted return and rank portfolio performance in the finance industry and is given by:

$$S = \frac{\overline{r_p} - \overline{r_f}}{\sigma_p}, \tag{19}$$

where $\overline{r_p}$ and σ_p are the average monthly return and standard deviation of a portfolio, and $\overline{r_f}$ refers to the average monthly risk-free rate, namely the one-month T-bill rate. Further, we also compare Sharpe ratios between the active investment portfolio and the buy-and-hold portfolio based on statistical inference. Statistical tests and inference for Sharpe ratios have been investigated in a variety of studies such as Jobson and Korkie (1981), Memmel (2003), DeMiguel and Nogales (2009). However, Ledoit and Wolf (2008) indicated that the test related to the Sharpe ratio by the above literature is not valid when returns have tails greater than the normal distribution or are of a time nature. Therefore, Ledoit and Wolf (2008) proposed a robust test for comparing Sharpe ratios based on a "studentized" time series bootstrap, and their empirical evidence suggests that their method is more robust than that used in earlier literature. Thus, in this study we use the approach of Ledoit and Wolf (2008) to test the following null hypothesis: $H0 : \Delta_s = 0$, where Δ_s is the difference in Sharpe ratios of two investment strategies given by $\Delta_s = S_1 - S_2$ (S_1 and S_2 refer to the Sharpe ratios of two strategies)[1]. Homm and Pigorsch (2012) proposed the EPM for accounting for the effects of higher moments on the investment performance evaluation which cannot be described by the Sharpe ratio. Hence, this measure offers a more comprehensive economic analysis and gives a more realistic representation of what investors consider practically when evaluating investment opportunities compared to the Sharpe ratio (Golec and Tamarkin (1998); Harvey and Siddique (2000); Chronopoulos et al. (2018)). The EPM is computed by dividing the mean return of the risky asset by its economic index of riskiness, which is a measure proposed by Aumann and Serrano (2008).

3.6. Results under different market conditions

In this section we consider the forecasting performance and the performance of the active trading strategy based on the forecasting results in different market conditions. We account for the following three market conditions: 1) before and after the 2008 global financial crisis, and 2) bull and bear markets. The 2008 global financial crisis may change the dependence structure between the macroeconomic variables and stock index returns due to the changes of macroeconomic policy adjustments, behaviours of investors, and corporate capital structures, etc. Investigating the performance of the method that we propose in bear and bull markets allows us to know whether it can be used for risk management. Moreover, the results are also useful for policy makers who can affect the stock prices of oil and gas companies by implementing polices to adjust specific macroeconomic variables.

For investigating the impact of the 2008 global financial crisis, we divide the entire sample period into two sub-periods: May 1988–July 2008 and August 2008–October 2017 to consider the impact of the 2008 global financial crisis. Regarding the research on the effects of bull and bear markets, we select the following sub-periods: June 2003–August 2008, September 2008–September 2011, October 2011–December 2013, and January 2014–October 2017.

4. Data

In this study we utilize the monthly macroeconomic variable database constructed by McCracken and Ng (2015) which contains 134 macroeconomic variables. These variables are classified into 8

[1] For more details, see Ledoit and Wolf (2008). We appreciate that the authors offer the MATLAB codes of the test online.

groups, namely, (1). Output and Income, (2). Labor Market, (3). Consumption and Orders, (4). Orders and Inventories, (5). Money and Credit, (6). Interest rate and Exchange Rates, (7). Prices and (8). Stock Market. This database has several advantages. First, it is updated monthly using the FRED database. Second, it is publicly accessible and facilitates comparison with related research. Third, all the time series in the database are stationary. Since we focus on out-of-sample forecasting ability, we remove the variables whose historical records are revised to avoid the "look-ahead bias". And we only consider the variables whose release date is lagged for one month. Finally, 103 variables are left for the prediction task. The names of the macroeconomic variables and the data transformation method with respect to these variables are displayed in Appendix A.

We denote the one-month U.S. T-bill rate as the risk-free return and download the data from the Kenneth R. French Data Library[2].

We use the U.S. oil and gas industry stock index, one of the 19 supersector Datastream Global Equity Indices classified based on Industry Classification Benchmark (ICB). [3] The price of the index is computed by value weighting the prices of U.S. oil and gas companies which have relatively large capitalization values. Due to the data availability of stock index and macroeconomic variables, the sample period spans from February 1973 to October 2017. The stock index data are collected from Thomson Reuters Datastream. We use a rolling-window method for forecasting the one-month ahead sign of the stock index returns. This approach allows practitioners to capture the structural changes between explanatory variables and dependent variables by ignoring the information from distant time periods and only using recent information. One problem of this method relates to the window length selection. Too short a rolling window means that the variable selection and the estimation of model parameters are sensitive to data outliers, while too long a rolling window may not capture the time-varying characteristics of the relationship between predictors and stock index returns. We set the window length as 180 months to select variables and train the model every month[4]. Therefore, the out-of-sample forecasting period is from April 1988 to October 2017 for the static-BIC model and from May 1988 to October 2017 for the dynamic-BIC model. Specifically, in

[2]`http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`

[3]The Datastream Global Equity Indices include six levels. Level 1 corresponds to the total market index which is decomposed into 10 industries in Level 2 including: Oil and Gas, Basic Materials, Industrials, Consumer Goods, Healthcare, Consumer Services, Telecommunications, Utilities, Financials, and Technology. Level 2 classifications are subdivided in detail by level 3-6 which are super-sectors, sectors and sub-sectors.

[4]We also tried 60-month, 120-month data as the rolling window length but the results are not statistically significant and worse than using 180-month data

month t, we collect the stationary transformed data of 103 macroeconomic variables from month $t-181$ to month t and split them into two categories A and B: A contains variables whose release date is lagged for one month and B contains variables whose variables are readily available. Then we use the values of variables in group A from month $t-181$ to $t-2$ and the values of variables in group B from month $t-180$ to $t-1$ as the explanatory variables. We consider the return sign values from month $t-179$ to month t as the dependent variable. Based on this, we can confirm that the length of the data for training the models described by eqs.(5)–(6) is 180 months, approximately 15 years.

Figure 2 shows the price series of the US oil and gas industry stock index and return series during the period from April 1988 to October 2017. We can observe two clear price-increasing periods, from about 2002 to 2008 and from 2009 to 2014. We also observe two price-slashing periods, namely, 2008–2009 and 2014–2015, which correspond to the global financial crisis and the recent oil price crash, respectively. Compared to the relatively stable price-increasing period between 1988 and 2002, the volatility of the stock index price becomes higher after 2002. Statistics of the industry index return series show that the index experienced an increasing trend during the period between April 1988 and October 2017. The negative skewness value indicates that the return series distribution is asymmetric around its mean and the probability of negative returns is higher than suggested by a symmetric distribution. The kurtosis value and the statistic of the Jarque-Bera test both suggest that the return series distribution is non-normal. The reason for these characteristics may be the extreme events which pushed the boom or slump in the stock index price during the out-of-sample period.

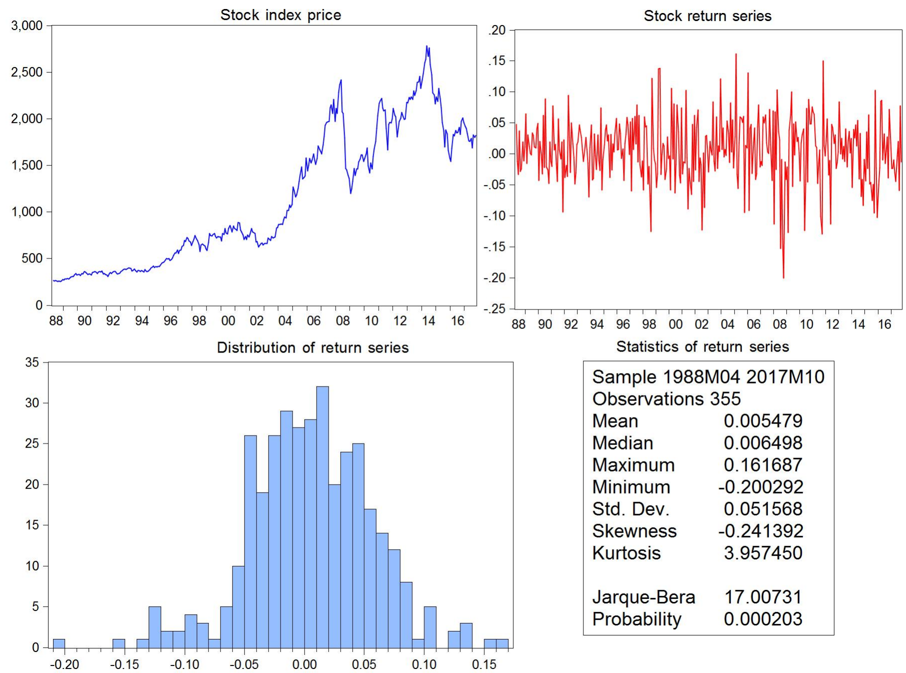


Figure 2: Time series plot and statistics of US oil and gas industry index during the period from April 1988 to October 2017

5. Results

Table 1 displays the out-of-sample forecasting results for the static and probit models based on the SFS variable selection algorithm discussed in Section 3. We find that the dynamic model with a time-varying threshold value, c_t, performs best for the sign prediction task during the whole out-of-sample period from May 1988 to October 2017. Based on the statistics of PT and T tests, the long-term forecasting accuracy is statistically significant. The forecasted return signs and the real return signs are dependent at 5% significance level during the out-of-sample period from May 1988 to October 2017. The total successful ratio is 0.554 while the prediction accuracy for the positive return signs, 0.622, is higher than the prediction accuracy for the negative return signs, 0.476. This phenomenon also exists in other models. The evidence indicates that the behaviour of macroeconomic factors may be a driver in pushing the prices of oil and gas industry stocks and thus partially explain the positive excess returns. One possible reason is that the profits of U.S. oil and gas companies are mainly affected by the supply chain of U.S. petroleum and relevant products. Therefore, the economic activities may reflect the potential demand for crude oil and other energy products, and thus the optimistic view of future oil price performance may stimulate investors to invest in the stocks of oil and gas companies.

More interestingly, comparing the two "Static" models, we find that the model with dynamic c_t performs better in the out-of-sample forecasting test. This result is consistent with the assumption that the risk aversion level of investors is time-varying so that an adaptive threshold value is more appropriate for modelling the probabilities of upward and downward movements of the stock index in different time periods. More importantly, adding the dynamic feature for the threshold value increase the forecasting accuracy of negative returns, which is crucial for risk management. Among the four models, the "Dynamic" model with "Dynamic" threshold value performs best in terms of out-of-sample forecasting.

Table 1: Out-of-sample forecasting results

	Static model $c_t = 0.5$	Dynamic model $c_t = 0.5$	Static model Dynamic c_t	Dynamic model Dynamic c_t
SR	0.530	0.528	0.541	0.554
SD	0.380	0.392	0.458	0.476
SU	0.661	0.649	0.614	0.622
QPS	0.551	0.549	0.551	0.549
PT	0.753	0.727	1.978	**3.341**
	[0.386]	[0.394]	[0.160]	**[0.068]**
T-statistic	0.843	0.818	**1.438**	**2.093**
	[0.200]	[0.207]	**[0.076]**	**[0.019]**

Notes: This table reports the out-of-sample forecasting results for the static and dynamic probit models based on the SFS variable selection algorithm in Section 3. The out-of-sample covers the period between May 1988 and October 2017. The first columns present the forecasting results for static and dynamic probit models under the constraint of constant threshold value, $c_t = 0.5$. The last two columns report the forecasting results for static and dynamic probit models where a dynamic threshold value c_t is used and the value is determined by the method described in Section 3.3. The bold text indicates statistically significant forecasting results.

Table B1 in Appendix B reports the variables which are selected using the combination of dynamic probit model and the SFS algorithm during the out-of-sample period. We find that the variables are mainly concentrated in the following categories: output and income, labor market, prices, interest rate, exchange rates, money and credit, and stock market. The 10 most frequently selected variables include two output and income variables, two labor market variables, four variables related to interest rate and exchange rates, one money and credit variable, and one variable which reflects the commodity price level. The monthly real personal income estimate is an important macroeconomic indicator which is emphasized by the National Bureau of Economic Research when dating the business cycle[5]. It measures the incomes obtained by individuals from participation in production, from government and business transfers, and from holding interest-bearing securities and corporate stocks. The civilian labor force reflects the employment situation of people who are not employed with any government or military institution. This indicator reflects the possibility of capital expansion of the companies in the U.S., which in turn reflects the potential demand for energy. The changes of real personal income and civilian labor force affect stock returns of the US oil and gas companies for the following reasons. First, with rising personal incomes there is an increased demand for petroleum products such as gasoline as the income elasticity of demand for these is positive and significant for US (see e.g. Haas and Schipper (1998), Dees et al. (2007) and Fournier et al. (2013)). Second, the increased employment means increased incomes and thus increased demand for petroleum products. Table 2 reports the results of the Granger causality test proposed by Granger (1969) for the changes of real personal income and the excess returns of the US oil and gas industry index based on the vector autoregressive model[6]. We discover that, using this test, the changes in real personal income cause the excess returns of the US oil and gas industry to be statistically significantly at 10% level. This evidence also suggests that historical value real personal income carries useful information for predicting US oil and gas stock excess returns.

We also observe that the macroeconomic variables related to money, interest rate, and credit are frequently selected for the prediction task. Oil and gas companies generally require large debt finance to invest in projects related to the extraction, drilling, and transportation of crude oil and derived products. Thus, monetary policies and interest rate levels closely relate to finance costs and cash flows.

[5] `http://www.nber.org/cycles/recessions.html`

[6] The Granger causality test results at 10% significance level for other variables are reported in B1 in Appendix B. We find that not all selected macroeconomic variables cause the excess stock return of the US oil and gas index. This may be because that Granger causality test is based on a linear static model without considering time-varying and nonlinear relationships. This issue will be accounted for in our future research

Figure 3 shows the number of selected variables and the time-varying threshold value for classifying negative and positive excess returns during the period from May 1988 to October 2017. We find that the number of selected variables is not constant, which suggests that macroeconomic variables do not always contain much useful information for predicting future excess returns of the US oil and gas industry. Furthermore, the number of selected variables experiences a growing trend, which indicates that the macroeconomic fundamentals of the U.S. oil and gas industry have become more complex, and practitioners need to consider more macroeconomic information when justifying the future performance of the industry. The threshold value shows a dynamic characteristic, which means that a constant threshold value such as 0.5 may not be optimal for the sign prediction task. On the other hand, this evidence also suggests that the risk aversion level of market participants may be time-varying.

Table 2: Causality test results for real personal income and excess returns of the US oil and gas industry index

Causality relationship	VAR lag	F-statistic	Fc (10%)	Fc (5%)	Fc (1%)
$R_{RPI} \Rightarrow R_{Index}$	1	3.487	2.715	3.859	6.683
$R_{Index} \Rightarrow R_{RPI}$	1	0.065			

Notes: This table reports the Granger causality test based on the vector autoregressive model (VAR). R_{RPI} and R_{Index} represent the changes and returns of monthly real personal income and excess returns of the US oil and gas industry index. The lags are selected based on minimizing the value of Bayesian Information Criteria. The critical values of F-statistic at 10% (Fc (10%)), 5% (Fc (10%)) and 1% (Fc (10%)) level are reported.

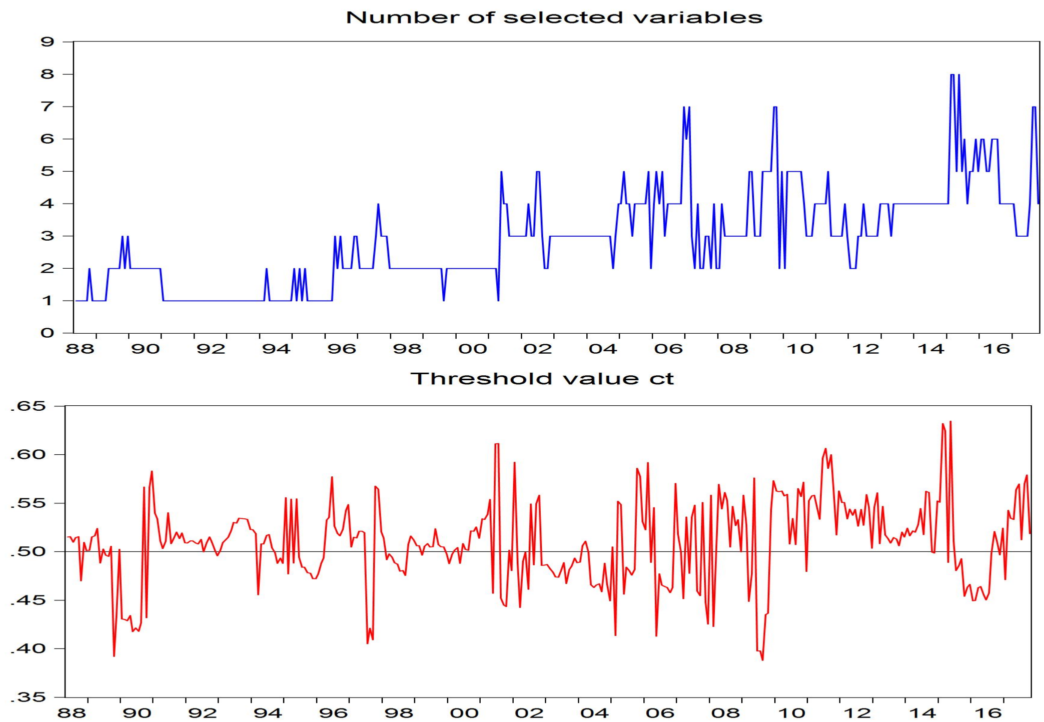


Figure 3: Number of selected variables and time-varying threshold value c_t during the period from May 1988 to October 2017

Table 3 displays the forecasting performance of the dynamic probit model in different sub–periods which represent different market conditions. Before the 2008 global financial crisis, the forecasting accuracy is statistically significant while after the crisis it is not significant. This result is potentially due to the increased volatility of energy markets after the financial crisis. We also find that the forecasting ability of the model does not depend on bull and bear market conditions. During the period from June 2003 to August 2008, the null hypothesis that the forecasted signs and real return signs are independent can only be rejected at 10% level based on the T-value. The accuracy is relative low and insignificant over the period between October 2011 and December 2013. The oil and gas industry stock index price experienced an increasing trend from October 2011 to December 2013 but also experienced a higher volatility compared to the period between June 2003 to August 2008. Comparing the two bear market periods, September 2008–September 2011 and January 2014–-October 2017, we find that the forecasting accuracy is higher in the recent oil price downward-moving period.

Table 3: Out-of-sample forecasting performance in different periods

Sub–periods	SR	SU	SD	QPS	PT		T–value	
May 1988–July 2008	0.564	0.636	0.477	0.526	**3.040**	**0.081**	**2.071**	**0.020**
Aug 2008–Oct 2017	0.532	0.589	0.473	0.599	0.299	0.584	0.690	0.246
Jun 2003–Aug 2008	0.583	0.632	0.500	0.516	0.906	0.341	**1.313**	**0.097**
Sep 2008–Sep 2011	0.538	0.571	0.500	0.695	0.229	0.632	0.400	0.346
Oct 2011–Dec 2013	0.481	0.647	0.200	0.523	0.231	0.631	-0.849	0.798
Jan 2014–Oct 2017	0.565	0.556	0.571	0.557	0.554	0.457	1.005	0.160

Notes: This table reports the forecasting performance of the dynamic probit model with time-varying threshold value c_t in different sub–periods.

Figure 4 shows the annualized Sharpe ratios of the active trading strategy described in Section 3.5 during the out-of-sample period between May 1988 and October 2017 when we use different combinations of ω^{up} and ω^{down} values. ω^{up} ranges from 0.1 to 1.5 while ω^{down} is within the interval between -1 and 0. The changing step size is set as 0.1. Following the studies of Lee and Mathur (1996b), Phan et al. (2015), Lee and Mathur (1996a), Narayan et al. (2013) and Szakmary and Mathur (1997), we assume that the transaction cost is constant and equal to 0.1% for every transaction[7]. The area of ”warmer” colour indicates that the annualized Sharpe ratio value is high while the area of ”colder” colour means low Sharpe ratio values. It shows that relaxing investment restrictions can significantly improve the risk-adjusted return of the trading strategy. However, as the ω^{up} value becomes higher, the Sharpe ratio converges to a constant value which is a little higher than 0.2. It is close to the annualized Sharpe ratio value of the buy-and-hold strategy, namely, 0.191 which is shown in Table 4. Moreover, the Sharpe ratio does not increase much if we only decrease the ω^{down} but keep the ω^{up} value unchanged. This evidence is consistent with the empirical findings of the forecasting accuracy test which indicates that forecasting accuracy is higher for positive return signs but lower for negative return signs.

[7]We also investigated the impacts of different transaction costs including 0, 0.2%, 0.3%, and 0.5%. The results are quite similar

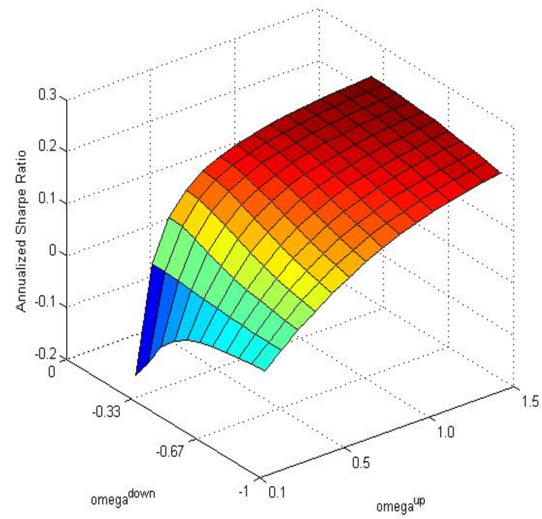
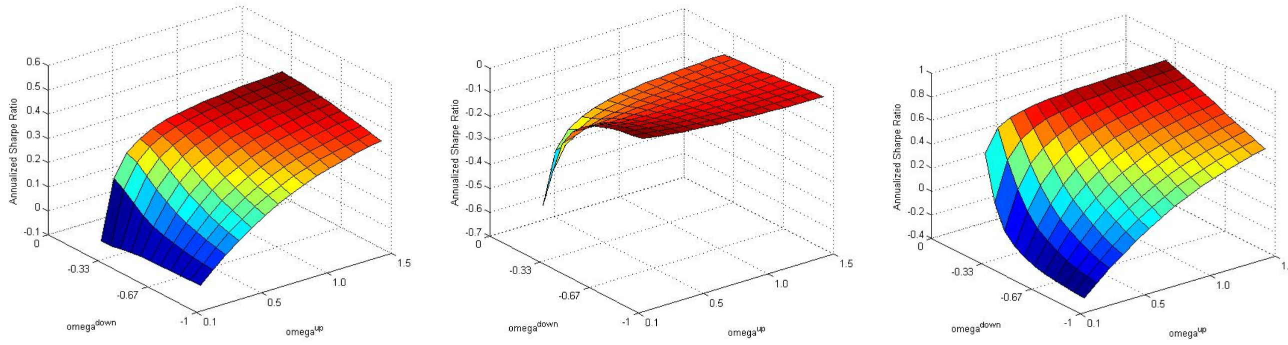
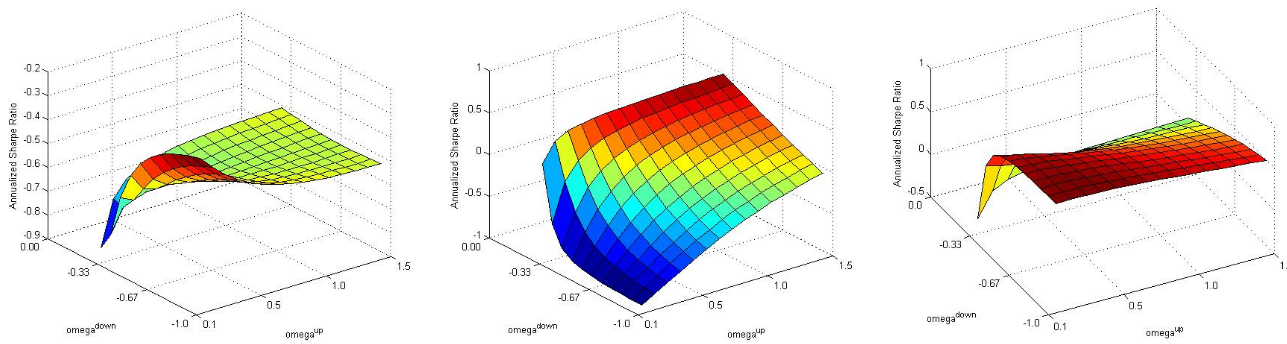


Figure 4: The values of annualized Sharpe ratio for different combinations of values of ω^{up} and ω^{down} under 0.1% transaction cost during the period from May 1988 to October 2017

Figure 5 shows the relationship among annualized Sharpe ratios, ω^{up} and ω^{down} in different sub-periods. The relationship is dynamic for different periods due to the different forecasting performance of the model. We can observe that the Sharpe ratio is quite stable and relatively high when $\omega^{up} = 1.5$ and $\omega^{down} = 0$. This evidence suggests that investors can increase the active trading strategy by increasing investment leverage when the forecasted return sign is positive but should not short sell when the forecasted return sign is negative.



(a) May 1988–July2008 (b) Aug 2008–Oct 2017 (c) Jun 2003–Aug 2008



(d) Sep 2008–Sep 2011 (e) Oct 2011–Dec 2013 (f) Jan 2014–Oct 2017

Figure 5: The values of annualized Sharpe ratio for different combinations of values of ω^{up} and ω^{down} under 0.1% transaction cost for different sub-periods

Table 4 displays the out-of-sample performance of the active trading strategy and the buy-and-hold trading strategy in different time periods. The values of ω^{up} and ω^{down} are selected as 1.5 and 0. This means that the investor is supposed not to short-sell the stock index but can borrow money to invest in the stock index. The volatility and average monthly return of the active trading strategy based on the dynamic probit model are lower than the buy-and-hold strategy in every considered period. This evidence indicates that low volatility requires the sacrifice of return, which is consistent with traditional financial theory. Moreover, the statistics of the test proposed by Ledoit and Wolf (2008) also show that the differences in Sharpe ratios for the two strategies are not significant, which means that the investors cannot obtain significant higher risk-adjusted returns by simply using the method that we propose. However, for the period between January 2014 and

October 2017, the active trading strategy experienced higher returns and lower risk compared to the buy-and-hold strategy. This result is consistent with the evidence shown in Table 3 that the forecasting accuracy of downward and upward movements is relatively high in this period. On the other hand, Figure 4 also shows that the number of selected macroeconomic variables in this period is relatively high. Therefore, we can infer that the number of macroeconomic indicators for determining the fundamentals of the U.S. oil and gas industry increased recently. In other words, practitioners should consider more economic indicators and their behaviour when making macroeconomic policies relating to the energy market or investing in the oil and gas industry. The performance of the active trading strategy and the buy-and-hold strategy is also evaluated by the values of the EPM measure which considers the impacts of the higher moments. The results are consistent with the findings reflected by the Sharpe ratios.

Table 4: Performance of active trading strategy and buy-and-hold strategy in different periods.

Sub–periods	A–Ret (%)	A–Std (%)	A–SR	E–SRs test p–value	EPM*100
May 1988–Oct 2017					
Active trading	7.564	13.191	0.230	0.853	0.865
Buy-and-hold	6.425	17.873	0.191		0.601
May 1988–July 2008					
Active trading	11.987	18.500	0.414	0.760	2.902
Buy-and-hold	10.067	16.756	0.343		1.944
Aug 2008–Oct 2017					
Active trading	-2.130	22.076	-0.104	0.944	0.000
Buy and hold	-1.545	19.991	-0.086		0.000
Jun 2003–Aug 2008					
Active trading	20.137	21.717	0.793	0.973	10.756
Buy-and-hold	19.951	19.337	0.881		11.426
Sep 2008–Sep 2011					
Active trading	-14.959	29.645	-0.510	0.973	0.000
Buy-and-hold	-6.841	24.984	-0.280		0.000
Oct 2011–Dec 2013					
Active trading	9.235	18.292	0.503	0.174	3.758
Buy-and-hold	17.200	15.736	1.091		17.403
Jan 2014–Oct 2017					
Active trading	1.437	16.615	0.073	0.370	0.090
Buy-and-hold	-8.098	17.640	-0.471		0.000

Table 4 (Continued)

Sub–periods	A–Ret (%)	A–Std (%)	A–SR	E–SRs test p–value	EPM*100

Notes: This table reports the performance of the active trading strategy based on the dynamic probit model and the buy-and-hold trading strategy in different out-of-sample periods. Here, $\omega^{up} = 1.5$ and $\omega^{down} = 0$. The average trading cost is assumed to be 0.1% for every transaction. A–Ret, A–Std, and A–SR represent annualized return, annualized standard deviation, and annualized Sharpe ratios, respectively. The column titled "E–SRs test p–value" reports the p–values of the statistical test which is proposed by Ledoit and Wolf (2008) for the null hypothesis that the Sharpe ratios of the active trading strategy and the buy-and-hold trading strategy are equal. The last column reports the EPM values for the two strategies. Following Homm and Pigorsch (2012) and Chronopoulos et al. (2018), we set the EPM measure equal to zero if the average excess returns of a portfolio are negative.

6. Conclusions

In this paper we study the predictive ability of macroeconomic variables for the signs of the excess returns of the U.S. oil and gas industry stock index. In contrast to previous studies which focus on the simultaneous effects of risk factors and predictions of return levels, we concentrate on the direction of movement of excess returns. The sign predictability in stock returns is important for investors in asset allocation and risk management. It is also useful for policy makers relating to the energy sector. We make several contributions to the literature.

First, we consider a framework of searching and using important information embedded in a large monthly macroeconomic dataset. As the quantity of information increases, how to utilize the information effectively becomes more complex. We use a forward sequential selection algorithm combined with the Bayesian information criterion to select the most important macroeconomic variables for the prediction task. This accounts for the forecasting accuracy and over-fitting problem simultaneously.

Second, we consider the time-varying relationship between macroeconomic variables and stock market performance. We also examine the impacts of dynamic market conditions and possible changing risk aversion of market participants when selecting predictors and threshold values for sign prediction. Our results show that the number of macroeconomic variables that can be used for the forecasting task changes over time. Breaking the assumption of the constant threshold value of the dynamic probit model can improve the forecasting accuracy.

We find that useful information for predicting the future performance of the oil and gas industry is mostly reflected in the variables grouped into the following categories: output and income, labor market, prices, interest rate and foreign exchange rates, money and credit, and stock market behaviour. Of these, real personal income is always selected. This result supports the findings of Haas and Schipper (1998), Dees et al. (2007) and Fournier et al. (2013) that the income elasticity of demand for petroleum products such as gasoline is positive and significant for US. Our results also show that the number of selected variables increases over time, and the forecasting accuracy in sub–periods suggests that macroeconomic fundamentals should be concentrated in the recent bear oil market period. However, the active trading strategy based on the forecasting results of the model and method that we use in this study cannot help investors to obtain significant higher risk-adjusted returns compared to the buy-and-hold strategy even if the investment restrictions are relaxed.

Acknowledgment

The authors are grateful for the funding support of the University of Aberdeen Elphinstone Scholarship for this research. The authors thank the Associate Editor and two referees for their critical comments that have led to significant improvements of the paper.

'Conflicts of interest: none'.

References

Aho, K., Derryberry, D., Peterson, T., 2014. Model selection for ecologists: the worldviews of aic and bic. Ecology 95, 631–636.

Altinbas, H., Biskin, O.T., 2015. Selecting macroeconomic influencers on stock markets by using feature selection algorithms. Procedia Economics and Finance 30, 22–29.

Anatolyev, S., Gospodinov, N., 2012. Modeling financial return dynamics via decomposition. Journal of Business & Economic Statistics .

Aumann, R.J., Serrano, R., 2008. An economic index of riskiness. Journal of Political Economy 116, 810–836.

Berge, T.J., Jordà, Ò., 2011. Evaluating the classification of economic activity into recessions and expansions. American Economic Journal: Macroeconomics 3, 246–277.

Bianconi, M., Yoshino, J.A., 2014. Risk factors and value at risk in publicly traded companies of the nonrenewable energy sector. Energy Economics 45, 19–32.

Boyer, M.M., Filion, D., 2007. Common and fundamental factors in stock returns of canadian oil and gas companies. Energy Economics 29, 428–453.

Burnham, K.P., Anderson, D.R., 2003. Model selection and multimodel inference: a practical information-theoretic approach. Springer Science & Business Media.

Çakmaklı, C., van Dijk, D., 2016. Getting the most out of macroeconomic information for predicting excess stock returns. International Journal of Forecasting 32, 650–668.

Campbell, S.D., Diebold, F.X., 2009. Stock returns and expected business conditions: Half a century of direct evidence. Journal of Business & Economic Statistics 27, 266–278.

Cenesizoglu, T., Timmermann, A., 2012. Do return prediction models add economic value? Journal of Banking & Finance 36, 2974–2987.

Christiansen, C., Eriksen, J.N., Møller, S.V., 2014. Forecasting us recessions: The role of sentiment. Journal of Banking & Finance 49, 459–468.

Christoffersen, P.F., Diebold, F.X., 2006. Financial asset returns, direction-of-change forecasting, and volatility dynamics. Management Science 52, 1273–1287.

Chronopoulos, D.K., Papadimitriou, F.I., Vlastakis, N., 2018. Information demand and stock return predictability. Journal of International Money and Finance 80, 59–74.

Dees, S., Karadeloglou, P., Kaufmann, R.K., Sanchez, M., 2007. Modelling the world oil market: Assessment of a quarterly econometric model. Energy Policy 35, 178–191.

DeMiguel, V., Nogales, F.J., 2009. Portfolio selection with robust estimation. Operations Research 57, 560–577.

Diaz, E.M., de Gracia, F.P., 2016. Oil price shocks and stock returns of oil and gas corporations. Finance Research Letters .

Fournier, J.M., Koske, I., Wanner, I., Zipperer, V., 2013. The price of oil–will it start rising again? .

Giovannini, M., Grasso, M., Lanza, A., Manera, M., 2006. Conditional correlations in the returns on oil companies stock prices and their determinants. Empirica 33, 193–207.

Golec, J., Tamarkin, M., 1998. Bettors love skewness, not risk, at the horse track. Journal of political economy 106, 205–225.

Gomes, J.F., Kogan, L., Yogo, M., 2007. Durability of output and expected stock returns. Technical Report. National Bureau of Economic Research.

Granger, C.W., 1969. Investigating causal relations by econometric models and cross-spectral methods. Econometrica: Journal of the Econometric Society , 424–438.

Haas, R., Schipper, L., 1998. Residential energy demand in oecd-countries and the role of irreversible efficiency improvements. Energy economics 20, 421–442.

Harvey, C.R., Siddique, A., 2000. Conditional skewness in asset pricing tests. The Journal of Finance 55, 1263–1295.

Homm, U., Pigorsch, C., 2012. Beyond the sharpe ratio: An application of the aumann–serrano index to performance measurement. Journal of Banking & Finance 36, 2274–2284.

Hong, Y., Chung, J., 2003. Are the directions of stock price changes predictable? statistical theory and evidence. Manuscript, Cornell University .

Jobson, J.D., Korkie, B.M., 1981. Performance hypothesis testing with the sharpe and treynor measures. The Journal of Finance 36, 889–908.

Kang, W., de Gracia, F.P., Ratti, R.A., 2017. Oil price shocks, policy uncertainty, and stock returns of oil and gas corporations. Journal of International Money and Finance 70, 344–359.

Lanza, A., Manera, M., Grasso, M., Giovannini, M., 2005. Long-run models of oil stock prices. Environmental Modelling & Software 20, 1423–1430.

Ledoit, O., Wolf, M., 2008. Robust performance hypothesis testing with the sharpe ratio. Journal of Empirical Finance 15, 850–859.

Lee, C.I., Mathur, I., 1996a. A comprehensive look at the efficacy of technical trading rules applied to cross-rates. The European Journal of Finance 2, 389–411.

Lee, C.I., Mathur, I., 1996b. Trading rule profits in european currency spot cross-rates. Journal of Banking & Finance 20, 949–962.

Leitch, G., Tanner, J.E., 1991. Economic forecast evaluation: profits versus the conventional error measures. The American Economic Review , 580–590.

Lettau, M., Ludvigson, S., 2001. Consumption, aggregate wealth, and expected stock returns. the Journal of Finance 56, 815–849.

Leung, M.T., Daouk, H., Chen, A.S., 2000. Forecasting stock indices: a comparison of classification and level estimation models. International Journal of Forecasting 16, 173–190.

McCracken, M.W., Ng, S., 2015. Fred-md: A monthly database for macroeconomic research. Journal of Business & Economic Statistics .

McCullagh, P., Nelder, J.A., 1989. Generalized linear models, no. 37 in monograph on statistics and applied probability.

Memmel, C., 2003. Performance hypothesis testing with the Sharpe ratio. Finance Letters 1.

Merton, R.C., 1981. On market timing and investment performance. i. an equilibrium theory of value for market forecasts. Journal of business , 363–406.

Mohanty, S.K., Nandha, M., 2011. Oil risk exposure: the case of the us oil and gas sector. Financial Review 46, 165–191.

Moya-Martínez, P., Ferrer-Lapeña, R., Escribano-Sotos, F., 2014. Oil price risk in the spanish stock market: An industry perspective. Economic Modelling 37, 280–290.

Narayan, P.K., Narayan, S., Sharma, S.S., 2013. An analysis of commodity markets: what gain for investors? Journal of Banking & Finance 37, 3878–3889.

Nazlioglu, S., Soytas, U., 2012. Oil price, agricultural commodity prices, and the dollar: A panel cointegration and causality analysis. Energy Economics 34, 1098–1104.

Newey, W.K., West, K.D., 1987. Hypothesis testing with efficient method of moments estimation. International Economic Review , 777–787.

Nyberg, H., 2011. Forecasting the direction of the us stock market with dynamic binary probit models. International Journal of Forecasting 27, 561–578.

Nyberg, H., Pönkä, H., et al., 2015. International sign predictability of stock returns: The role of the united states. CREATES Research Papers 20, 2015.

Pesaran, M.H., Timmermann, A., 2009. Testing dependence among serially correlated multicategory variables. Journal of the American Statistical Association 104, 325–337.

Phan, D.H.B., Sharma, S.S., Narayan, P.K., 2015. Stock return forecasting: some new evidence. International Review of Financial Analysis 40, 38–51.

Piazzesi, M., Schneider, M., Tuzel, S., 2007. Housing, consumption and asset pricing. Journal of Financial Economics 83, 531–569.

Pönkä, H., 2016. Real oil prices and the international sign predictability of stock returns. Finance Research Letters 17, 79–87.

Pönkä, H., 2017. Predicting the direction of us stock markets using industry returns. Empirical Economics 52, 1451–1480.

Ramos, S.B., Veiga, H., 2011. Risk factors in oil and gas industry returns: International evidence. Energy Economics 33, 525–542.

Rydberg, T.H., Shephard, N., 2003. Dynamics of trade-by-trade price movements: decomposition and models. Journal of Financial Econometrics 1, 2–25.

Sadorsky, P., 2001. Risk factors in stock returns of canadian oil and gas companies. Energy Economics 23, 17–28.

Sadorsky, P., 2008. The oil price exposure of global oil companies. Applied Financial Economics Letters 4, 93–96.

Shaeri, K., Adaoglu, C., Katircioglu, S.T., 2016. Oil price risk exposure: A comparison of financial and non-financial subsectors. Energy 109, 712–723.

Szakmary, A.C., Mathur, I., 1997. Central bank intervention and trading rule profits in foreign exchange markets. Journal of International Money and Finance 16, 513–535.

Vrieze, S.I., 2012. Model selection and psychological theory: a discussion of the differences between the akaike information criterion (aic) and the bayesian information criterion (bic). Psychological methods 17, 228.

Zagaglia, P., 2010. Macroeconomic factors and oil futures prices: a data-rich model. Energy Economics 32, 409–417.

Appendix A. Macroeconomic variables

Table A1: Macroeconomic variables

Tcode	FRED	Description	Release date
5	RPI	Real Personal Income	Last day
5	W875RX1	Real personal income excluding current transfer receipts	Last day
5	DPCERA3M086SBEA	Real personal consumption expenditures (chain-type quantity index)	Last day
5	INDPRO	Industrial Production Index	15th
5	IPFPNSS	Industrial Production: Final Products and Nonindustrial Supplies	15th
5	IPFINAL	Industrial Production: Final Products (Market Group)	15th
5	IPCONGD	Industrial Production: Consumer Goods	15th
5	IPDCONGD	Industrial Production: Durable Consumer Goods	15th
5	IPNCONGD	Industrial Production: Nondurable Consumer Goods	15th
5	IPBUSEQ	Industrial Production: Business Equipment	15th
5	IPMAT	Industrial Production: Materials	15th
5	IPDMAT	Industrial Production: Durable Materials	15th
5	IPNMAT	Industrial Production: Nondurable Materials	15th
5	IPMANSICS	Industrial Production: Manufacturing (SIC)	15th
5	IPB51222S	Industrial Production: Residential utilities	15th
5	IPFUELS	Industrial Production: Fuels	15th
2	CUMFNS	Capacity Utilization: Manufacturing (SIC)	15th
5	CLF16OV	Civilian Labor Force	8th
5	CE16OV	Civilian Employment Level	8th
2	UNRATE	Civilian Unemployment Rate	8th
2	UEMPMEAN	Average (Mean) Duration of Unemployment	8th
5	UEMPLT5	Number of Civilians Unemployed for Less Than 5 Weeks	8th
5	UEMP5TO14	Number of Civilians Unemployed for 5 to 14 Weeks	8th
5	UEMP15OV	Number of Civilians Unemployed for 15 Weeks and Over	8th
5	UEMP15T26	Number of Civilians Unemployed for 15 to 26 Weeks	8th
5	UEMP27OV	Number of Civilians Unemployed for 27 Weeks and Over	8th
5	PAYEMS	All Employees: Total Nonfarm Payrolls	8th
5	USGOOD	All Employees: Goods-Producing Industries	8th
5	CES1021000001	All Employees: Mining and Logging: Coal Mining	8th
5	USCONS	All Employees: Construction	8th
5	MANEMP	All Employees: Manufacturing	8th
5	DMANEMP	All Employees: Durable Goods	8th
5	NDMANEMP	All Employees: Nondurable goods	8th
5	SRVPRD	All Employees: Service-Providing Industries	8th
5	USTPU	All Employees: Trade, Transportation and Utilities	8th
5	USWTRADE	All Employees: Wholesale Trade	8th
5	USTRADE	All Employees: Retail Trade	8th
5	USFIRE	All Employees: Financial Activities	8th
5	USGOVT	All Employees: Government	8th
1	CES0600000007	Average Weekly Hours of Production and Nonsupervisory Employees: Goods-Producing	8th
2	AWOTMAN	Average Weekly Overtime Hours of Production and Nonsupervisory Employees: Manufacturing	8th
1	AWHMAN	Average Weekly Hours of Production and Nonsupervisory Employees: Manufacturing	8th
4	HOUST	Housing Starts: Total: New Privately Owned Housing Units Started	19th
4	HOUSTNE	Housing Starts in Northeast Census Region	19th

Table A1 (Continued)

Tcode	FRED	Description	Release date
4	HOUSTMW	Housing Starts in Midwest Census Region	19th
4	HOUSTS	Housing Starts in South Census Region	19th
4	HOUSTW	Housing Starts in West Census Region	19th
4	PERMIT	New Private Housing Units Authorized by Building Permits	19th
4	PERMITNE	New Private Housing Units Authorized by Building Permits in the Northeast Census Region	19th
4	PERMITMW	New Private Housing Units Authorized by Building Permits in the Midwest Census Region	19th
4	PERMITS	New Private Housing Units Authorized by Building Permits in the South Census Region	19th
4	PERMITW	New Private Housing Units Authorized by Building Permits in the West Census Region	19th
6	M1SL	M1 Money Stock	14th
6	M2SL	M2 Money Stock	14th
5	M2REAL	Real M2 Money Stock	15th
6	AMBSL	St. Louis Adjusted Monetary Base	8th
6	TOTRESNS	Total Reserves of Depository Institutions	7th
7	NONBORRES	Reserves of Depository Institutions, Nonborrowed	7th
6	BUSLOANS	Commercial and Industrial Loans, All Commercial Banks	15th
6	REALLN	Real Estate Loans, All Commercial Banks	15th
5	S&P 500	S&P's Common Stock Price Index: Composite	Readily available
5	S&P: indust	S&P's Common Stock Price Index: Industrials	Readily available
2	FEDFUNDS	Effective Federal Funds Rate	1st
2	TB3MS	3-Month Treasury Bill: Secondary Market Rate	1st
2	TB6MS	6-Month Treasury Bill: Secondary Market Rate	1st
2	GS1	1-Year Treasury Constant Maturity Rate	1st
2	GS5	5-Year Treasury Constant Maturity Rate	1st
2	GS10	10-Year Treasury Constant Maturity Rate	1st
2	AAA	Moody's Seasoned Aaa Corporate Bond Yield	5th
2	BAA	Moody's Seasoned Baa Corporate Bond Yield	5th
1	TB3SMFFM	3-Month Treasury Bill Minus Federal Funds Rate	1st
1	TB6SMFFM	6-Month Treasury Bill Minus Federal Funds Rate	1st
1	T1YFFM	1-Year Treasury Constant Maturity Minus Federal Funds Rate	1st
1	T5YFFM	5-Year Treasury Constant Maturity Minus Federal Funds Rate	1st
1	T10YFFM	10-Year Treasury Constant Maturity Minus Federal Funds Rate	1st
1	AAAFFM	Moody's Seasoned Aaa Corporate Bond Minus Federal Funds Rate	5th
1	BAAFFM	Moody's Seasoned Baa Corporate Bond Minus Federal Funds Rate	5th
5	TWEXMMTH	Trade Weighted U.S. Dollar Index: Major Currencies	Readily available
5	EXSZUSx	Switzerland / U.S. Foreign Exchange Rate	Readily available
5	EXJPUSx	Japan / U.S. Foreign Exchange Rate	Readily available
5	EXUSUKx	U.S. / U.K. Foreign Exchange Rate	Readily available
5	EXCAUSx	Canada / U.S. Foreign Exchange Rate	Readily available
6	WPSFD49207	Producer Price Index by Commodity for Final Demand: Finished Goods	12th
6	WPSFD49502	Producer Price Index by Commodity for Final Demand: Personal Consumption Goods (Finished Consumer Goods)	12th
6	WPSID61	Producer Price Index by Commodity for Intermediate Demand by Commodity Type: Processed Goods for Intermediate Demand	12th
6	WPSID62	Producer Price Index by Commodity for Intermediate Demand by Commodity Type: Unprocessed Goods for Intermediate Demand	12th
6	OILPRICEx	Crude Oil Prices: West Texas Intermediate (WTI) - Cushing, Oklahoma	Readily available
6	PPICMM	Producer Price Index by Commodity Metals and metal products: Primary nonferrous metals	12th
6	CPIAUCSL	Consumer Price Index for All Urban Consumers: All Items	13th
6	CPIAPPSL	Consumer Price Index for All Urban Consumers: Apparel	13th
6	CPITRNSL	Consumer Price Index for All Urban Consumers: Transportation	13th
6	CPIMEDSL	Consumer Price Index for All Urban Consumers: Medical Care	13th

Table A1 (Continued)

Tcode	FRED	Description	Release date
6	CUSR0000SAC	Consumer Price Index for All Urban Consumers: Commodities	13th
6	CUSR0000SAD	Consumer Price Index for All Urban Consumers: Durables	13th
6	CUSR0000SAS	Consumer Price Index for All Urban Consumers: Services	13th
6	CPIULFSL	Consumer Price Index for All Urban Consumers: All Items Less Food	13th
6	CUSR0000SA0L2	Consumer Price Index for All Urban Consumers: All items less shelter	13th
6	CUSR0000SA0L5	Consumer Price Index for All Urban Consumers: All items less medical care	13th
6	CES0600000008	Average Hourly Earnings of Production and Nonsupervisory Employees: Goods-Producing	8th
6	CES2000000008	Average Hourly Earnings of Production and Nonsupervisory Employees: Construction	8th
6	CES3000000008	Average Hourly Earnings of Production and Nonsupervisory Employees: Manufacturing	8th
6	MZMSL	MZM Money Stock	15th
6	INVEST	Securities in Bank Credit at All Commercial Banks	15th

Notes: This table reports the candidate macroeconomic variables which are used for predicting the sign of future oil and gas industry returns. The variables come form the dataset of monthly macroeconomic variables constructed by McCracken and Ng (2015). The column TCODE denotes the following data transformation for a series x: (1) no transformation; (2) Δx_t; (3) $\Delta^2 x_t$; (4)$log(x_t)$; (5) $\Delta log(x_t)$; (6) $\Delta^2 log(x_t)$; (7) $(x_t/x_{t-1} - 1.0)$. The FRED column gives mnemonics in FRED followed by a short description. Δ refers to difference once and Δ^2 refers to difference twice. The release dates of these variables are lagged for one month except several ones which are accessible readily. Namely, for most macroeconomic variables, the practitioners can only obtain the value for month $t-1$ in month t. We report the release date of the values of the variables for previous month in the last column of the table. Readily available means the variable's data can be obtained in current month.

Appendix B.

Table B1: Selected variables and selection frequency

Variable name	Categories	T CODE	Frequency (%)	Granger Test(10%)
Real personal income	Output and income	(5)	100.00	Yes
Civilian Labor Force	Labor market	(5)	22.88	No
St. Louis Adjusted Monetary Base	Money and credit	(6)	20.06	No
Consumer Price Index for All Urban Consumers: Commodities	Prices	(6)	15.82	Yes
Total Reserves of Depository Institutions	Interest rate and exchange rates	(6)	12.71	Yes
Moody's Seasoned Aaa Corporate Bond Yield	Interest rate and exchange rates	(1)	12.43	No
All Employees: Government	Labor market	(5)	11.58	No
Producer Price Index by Commodity Metals and metal products: Primary nonferrous metals	Interest rate and exchange rates	(6)	10.73	Yes
Moody's Seasoned Baa Corporate Bond Minus Federal Funds Rate	Interest rate and exchange rates	(1)	10.17	No
Industrial Production: Durable Materials	Output and income	(5)	9.60	Yes
S&P's Common Stock Price Index: Composite	Stock market	(5)	7.06	Yes
Industrial Production: Durable Consumer Goods	Output and income	(5)	5.65	No
All Employees: Retail Trade	Labor market	(5)	5.65	No
S&P's Common Stock Price Index: Industrials	Stock market	(5)	5.08	Yes
All Employees: Wholesale Trade	Labor market	(5)	3.67	No
Trade Weighted U.S. Dollar Index: Major Currencies	Interest rate and exchange rates	(5)	3.67	Yes
U.S. / U.K. Foreign Exchange Rate	Interest rate and exchange rates	(5)	3.39	Yes
Consumer Price Index for All Urban Consumers: Durables	Prices	(6)	3.39	No
Consumer Price Index for All Urban Consumers: All Items	Prices	(6)	3.11	Yes
Real M2 Money Stock	Interest rate and exchanges	(5)	2.82	No
Crude Oil Prices: West Texas Intermediate (WTI)–Cushing, Oklahoma	Prices	(6)	2.54	Yes
Industrial Production: Materials	Output and income	(5)	1.98	Yes
Consumer Price Index for All Urban Consumers: Transportation	Prices	(6)	1.98	Yes
Canada / U.S. Foreign Exchange Rate	Interest rate and exchange rates (5)		1.41	Yes
Civilian Unemployment Rate	Labor market	(2)	1.13	No
Average Weekly Hours of Production and Nonsupervisory Employees: Manufacturing	Labor market	(2)	1.13	No
Reserves of Depository Institutions, Nonborrowed	Money and credit	(7)	1.13	No
All Employees: Mining and Logging: Coal Mining	Labor market	(5)	0.85	No
M1 Money Stock	Money and credit	(6)	0.85	
Switzerland / U.S. Foreign Exchange Rate	Interest rate and exchange rates	(5)	0.85	Yes

Table B1: Macroeconomic variables

Variable name	Categories	T CODE	Frequency (%)	Granger Test(10%)
All Employees: Service-Providing Industries	Labor market	(5)	0.56	No
Producer Price Index by Commodity for Intermediate Demand by Commodity Type: Unprocessed Goods for Intermediate Demand	Prices	(6)	0.56	Yes
Real personal consumption expenditures (chain-type quantity index)	Consumption, orders and inventories	(5)	0.28	Yes
Industrial Production Index	Output and income	(5)	0.28	Yes
Industrial Production: Nondurable Materials	Output and income	(5)	0.28	No
Industrial Production: Residential utilities	Output and income	(5)	0.28	No
All Employees: Total Nonfarm Payrolls	Labor market	(5)	0.28	No
All Employees: Goods-Producing Industries	Labor market	(5)	0.28	No
Moody's Seasoned Aaa Corporate Bond Minus Federal Funds Rate	Interest rate and exchange rates	(1)	0.28	No
Consumer Price Index for All Urban Consumers: Apparel	Prices	(6)	0.28	No

Notes: This table reports the selected variables using the combination of dynamic probit model and SFS variable selection algorithm in Section 3. The column TCODE denotes the following data transformation for a series x: (1) no transformation; (2) Δx_t; (3) $\Delta^2 x_t$; (4)$log(x_t)$; (5) $\Delta log(x_t)$; (6) $\Delta^2 log(x_t)$; (7) $(x_t/x_{t-1} - 1.0)$. We also report their categories and frequencies during the out-of-sample period. The table reports the results of Granger causality test for the selected variables in the last column based on the full-sample data which spans the period between February 1973 and October 2017. The variables which are labelled "Yes" are the ones cause the excess stock return at 10% significance level.

Highlights

- Propose a method of selecting the macroeconomic variables for out-of-sample prediction of the excess return signs of the U.S. oil and gas industry stock index by combining the Forward Sequential Variable Selection Algorithm and information criteria.

- Propose a method which can change the threshold value of the probit model automatically for considering the potential time-varying risk aversion level of the market participants.

- Real personal income is identified as an important predictor for stock excess return changing direction, and we investigate the economic reason of this evidence.

- The economic value of the identified probit forecasting model is investigated by constructing an active trading strategy.