

Research Article

Forecasting Oil Price by Hierarchical Shrinkage in Dynamic Parameter Models

Yuntong Liu ¹, Yu Wei ¹, Yi Liu ² and Wenjuan Li ³

¹School of Finance, Yunnan University of Finance and Economics, Kunming, China

²Faculty of Transportation Engineering, Kunming University of Science and Technology, Kunming, China

³School of Statistics & Mathematics, Yunnan University of Finance and Economics, Kunming, China

Correspondence should be addressed to Yu Wei; weiyusy@126.com

Received 8 October 2020; Revised 14 November 2020; Accepted 20 November 2020; Published 3 December 2020

Academic Editor: Dehua Shen

Copyright © 2020 Yuntong Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The aim of this paper is to forecast monthly crude oil price with a hierarchical shrinkage approach, which utilizes not only LASSO for predictor selection, but a hierarchical Bayesian method to determine whether constant coefficient (CC) or time-varying parameter (TVP) predictive regression should be employed in each out-of-sample forecasting step. This newly developed method has the advantages of both model shrinkage and automatic switch between CC and TVP forecasting models; thus, this may produce more accurate predictions of crude oil prices. The empirical results show that this hierarchical shrinkage model can outperform many commonly used forecasting benchmark methods, such as AR, unobserved components stochastic volatility (UCSV), and multivariate regression models in forecasting crude oil price on various forecasting horizons.

1. Introduction

Crude oil price is one of the key indicators of the global macroeconomy and financial markets [1–6]. However, the oil price prediction is a complex process since various factors affect oil pricing [2] and the influence degree of these factors on oil price varies over time [7–11]. So, finding a proper oil price forecasting method, which is not merely able to select the important predictors but also reflect the dynamics of predictors impact, is of interest for a wide range of applications [12–19].

A vast of literatures [2, 4, 5, 11, 13, 18, 20–25] indicate that except for previous oil prices, other parameters such as basic oil supply, demand and oil stock effects, financial market forces, market sentiment and uncertainty, macroeconomy, and geopolitical influences are also main influencing factors. If adding all these explanatory variables into the multivariate regression or autoregression (AR) class framework, it may lead to overfitting and misspecification problems and thereby constrain the forecast accuracy [7, 26, 27]. Additionally, time-varying effect of these parameters should be also considered in oil price forecasting,

but drawing the time-varying effect into regression models would make the overfitting problem worse [7, 11, 28].

In this study, we introduce a prevailing Bayesian approach which not only overcomes overparametrization and misspecification problems in oil price prediction, but also discusses the time-varying properties of explanatory parameters in both short and long oil price forecasting horizons. This study mainly makes three contributions to the literature on oil price forecasting as follows.

First, we can estimate a large number of explanatory parameters with limited observations. Usually, low-frequency dataset is easier to access and process than high-frequency dataset; putting more informative explanatory variables into the model can help macroeconomists, politicians, and other market participants get more comprehensive information on the crude oil price. Further, we implement least absolute shrinkage and selection operator (LASSO) shrinkage method to handle all the considered endogenous and exogenous explanatory factors and select the most powerful influential factors automatically. Although previous studies [6, 29–33] simulate that LASSO-based approaches show better out-of-sample forecasts and

surpass both AR class models and time-varying parameter models, it is unclear whether LASSO operator is also outperforming other commonly used benchmark models in oil price forecasting. Examining the LASSO operator effectiveness may help oil market decision-makers identify significant influential indicators efficiently and seize investment opportunities.

Besides, for better explaining the oil price, we introduce more comprehensive exogenous (see Table 1) and endogenous variables (e.g., observations from previous time steps) as regression predictors. On the one hand, bringing previous oil prices into the regression enables comparison with autoregression models (AR) and time-varying vector autoregression (TVP-VAR) models, which are commonly used and proved models in energy price prediction that can generate accurate forecasts [18, 34–37]. On the other hand, we introduce a more comprehensive exogenous factors framework, which avoids model misspecification. Most of the oil price forecasting studies [3, 16, 34, 38–40] only focus on several key oil price predictors and ignore the rest due to the limited variables processing capacity; this leads to error of misspecification, while using the LASSO operator in this study can shrink the coefficient on unimportant explanatory variables to zero and include all the exogenous variables within the model without having to worry about multi-indicators' processing capacity.

Second, it has been well documented that the predictive ability of the forecast parameters on crude oil prices varies over time [7–11, 18, 41]. This motivates us to study the time-varying properties of the regression coefficients. Shrinkage model in time-varying parameters is described by [28] and is considered an effective forecasting method [32]. Accordingly, we apply LASSO for the time-varying regression model in the oil market and evaluate oil price forecasting performance. This Bayesian-based estimation method can predict both long-term and short-term forecast horizons via monthly information. With hierarchical shrinkage in oil price predictors, we can select the most relevant predictors and pick out time-varying parameters automatically. It is worth noting that few works investigate parameters dynamic properties incorporating a large set of predictors in a single model. Our study provides empirical evidence regarding the most powerful contributor in forecasting oil prices and judges its dynamic properties simultaneously.

Third, we extend our ideas for using the mean of the log predictive likelihood (MLPL) to check the entire of predictive distribution robustness, which fill gaps of the commonly used forecasting performance measurement—the mean of the squared forecast errors (MSFE) and the mean of the absolute value of the forecast errors (MAFE) [19, 23, 27, 28]—which can only judge the point forecasts. We also examine the forecasting performance by changing regressors, dependence variables, and rolling window estimation regimes for robustness check. Our out-of-sample evidence indicates that LASSO hierarchical shrinkage models outperform other competing models in most cases; LASSO can select informational variables automatically and efficiently.

The remainder of this paper is organized as follows: Section 2 presents the econometric approach and

comparison models. Section 3 introduces our data. Section 4 provides the out-of-sample empirical results and discussion. In Section 5, we present the robustness checks, and Section 6 concludes.

2. Empirical Models and Computation Processes

2.1. Empirical Models. Crude oil has both commodity and financial properties. As aforementioned, apart from the previous oil prices, we still have hundreds of influencing exogenous variables and seasonal adjustment should also be taken into account. In this case, the suitable full model for forecasting crude oil price is given by

$$\text{Oil}_{t+h} = c + \sum_{k=1}^k \beta_k x_{kt} + \sum_{r=0}^{p-1} \alpha_r \text{Oil}_{t-r} + \sum_{j=1}^{11} \gamma_j \text{dum}_j + \varepsilon_{t+h}, \quad (1)$$

where Oil_{t+h} is the future crude oil price we want to forecast at h -periods ahead, c is the intercept, and $\varepsilon_t \sim N(0, \sigma_t^2)$ is the error term. $\sum_{k=1}^k \beta_k x_{kt}$ represents the sum of exogenous variables part, k is the number of explanatory variables, and β_k is the k th regression parameter. $\sum_{r=0}^p \alpha_r \text{Oil}_{t-r}$ includes the sum of p lags of oil price; α_r is the r th lag coefficient. $\sum_{j=1}^{11} \gamma_j \text{dum}_j$ is the sum of 11 monthly dummies which is added for seasonal adjusting. γ_j is the j th dummy variable coefficient. In total, the number of potential independent variables should be $m = 1 + k + p + 11$.

Each part of the model (intercept $\sum_{j=1}^{11} \gamma_j \text{dum}_j$, $\sum_{r=0}^{p-1} \alpha_r \text{Oil}_{t-r}$, or $\sum_{k=1}^k \beta_k x_{kt}$) can be excluded from the model. Briefly, the computation steps can analyze models by adding different terms into the model, then judging the time-varying properties, and selectively do LASSO shrinkage for the variable parameters in both constant variance (homoskedasticity of σ_t) and stochastic variance (heteroskedasticity of σ_t) regimes. So, the model structure is diversifying to the following three restrictive forms:

(1) AR(p) model

$$\text{Oil}_{t+h} = c + \sum_{r=0}^{p-1} \alpha_r \text{Oil}_{t-r} + \sum_{j=1}^{11} \gamma_j \text{dum}_j + \varepsilon_{t+h}. \quad (2)$$

The model specifies that the future crude oil price depends linearly on its past values.

(2) Multivariate regression model

$$\text{Oil}_{t+h} = c + \sum_{k=1}^k \beta_k x_{kt} + \sum_{j=1}^{11} \gamma_j \text{dum}_j + \varepsilon_{t+h}. \quad (3)$$

The model considers the effect of several key exogenous variables but excludes endogenous variables' influence on the oil price.

TABLE 1: Variable definitions.

Category	Label	Definition	Unit	Data source
Crude oil price	WTI	WTI spot price	Dollars per barrel	Energy Information Administration (EIA)
	Brent	Europe Brent spot price		
Crude oil supply	OS_World	Crude oil production, world	Thousand barrels	
Crude oil demand	OD_cons	Total petroleum consumption	per day	
Crude oil stocks	OD_stocks	Total petroleum stocks	Million barrels	
Gold price	P_gold	Gold fixing price	US dollars per troy ounce	Federal Reserve Bank of St. Louis
Exchange rate	US_ex	Trade weighted US dollar index: broad, goods, index	Percentage	
Stock market price index	S&P500	S&P500 index		
Natural gas price	P_gas	Henry Hub Natural Gas Spot Price	Dollars per million btu	Energy Information Administration (EIA)
Market sentiment	VIX	VIX index		Federal Reserve Bank of St. Louis
Macroeconomy affecting factors	IP_total	Industrial production: total index	Percentage	
	Kilian	Kilian index		Kilian's website
Political change	EPU	Global policy uncertainty	Percentage	Economic policy uncertainty website
	Google	Google trend index	Percentage of popularity	Google trends

All the variables are calculated by the first log difference in order to make them stationary, except for the Kilian index, which is naturally stationary series. Neither measures of these variables are seasonally adjusted.

(3) The unobserved components stochastic volatility (UCSV) model

$$\text{Oil}_{t+h} = c + \sum_{j=1}^{11} \gamma_j \text{dum}_j + \varepsilon_{t+h}, \quad (4)$$

which assumes that the future oil price consists of components with a direct interpretation that cannot be observed.

These three models are commonly used and proved that they can generate relatively accurate linear regression prediction [18, 34–37, 42–45]. Same as the full model (equation (1)), these three restricted model versions can also do hierarchical parameter shrinkage and decide which variable parameter varies with time. In Section 4.2, we compare the full model and the three restricted models in prediction performance with the same prior choices and basic model structures. The specific econometric method computation processes are as follows.

2.2. Judging Time-Varying Properties and Forecasting Power of Predictors. To briefly describe the computation processes, the full variable model (equation (1)) can be simplified as

$$\begin{aligned} \text{Oil}_{t+h} &= \beta_t^* z_t + \varepsilon_{t+h}, \\ \beta_t^* &= \beta_{t-1}^* + v_t, \end{aligned} \quad (5)$$

where the variable of interest, Oil_{t+h} , can be defined as $\text{Oil}_{t+h} = \log(p_{t+h}) - \log(p_t)$. The variable matrix $z_t = [1, x_1, \dots, x_k, \Delta \log(p_t), \dots, \Delta \log(p_{t-p+1}), \text{dum}_1, \dots, \text{dum}_{11}]$, and the corresponding coefficients matrix of z_t is $\beta_t^* = [c, \beta_1, \dots, \beta_{k+p+1}]'$.

In equation (5), we assume $\varepsilon_t \sim N(0, \sigma_t^2)$ and $v_t \sim N(0, \Omega)$. σ_t^2 can be stochastic or constant volatility. The errors are assumed to be independent of each other and

independent at all leads and lags. Ω is of dimension $m \times m$, which can be large relative to the number of observations. To keep the model relatively brief, we assume Ω is a diagonal matrix, $\Omega = \text{diag}(\omega_1^2, \dots, \omega_m^2)$. Ω introduces shrinkage in the time variation then switches the constant coefficients to time-varying coefficients. If ω_i is zero, the i th ($i = 1, \dots, m$) coefficient is constant over time, and larger values of ω_i mean more time variation. In order to elicit ω_i , Belmonte et al. [28] separate the model into two parts, one part is constant (represented by βz_t) and the other part is time-varying (represented by $\beta_t z_t$). Equation (5) will change to

$$\begin{aligned} \text{Oil}_{t+h} &= \beta z_t + \beta_t z_t + \varepsilon_{t+h}, \\ \beta_t &= \beta_{t-1} + v_{t+h}, \\ \beta_0 &= 0, \end{aligned} \quad (6)$$

where $\beta = \beta_0^*$ and $\beta_t = \beta_0^* - \beta$. Then, let $\tilde{\beta}_{t,t} = \beta_{i,t}/\omega_i$ and transform equation (6) to

$$\begin{aligned} \text{Oil}_{t+h} &= \sum_{i=1}^r \beta_i z_{i,t} + \sum_{i=1}^r \omega_i \tilde{\beta}_{i,t} z_{i,t} + \varepsilon_{t+h}, \\ \tilde{\beta}_{i,t} &= \tilde{\beta}_{i,t-1} + v_{i,t}, \\ \tilde{\beta}_{i,0} &= 0, \end{aligned} \quad (7)$$

where $v_{i,t} \sim N(0, 1)$ for $i = 1, \dots, r$.

Through implementing LASSO in terms of equation (7), we can judge the time-varying properties and forecasting power of predictors. Four possible computation cases are discussed as follows:

- (1) ω_i shrank to 0, but β_i is not shrunk to 0; then, the i th variable parameter is constant over time
- (2) Both ω_i and β_i shrank to 0; then, the i th variable is irrelevant for forecasting the oil price

- (3) ω_i is not shrunk to 0, but β_i shrank to 0; then, the i th variable parameter has small time-varying characteristics (since $\tilde{\beta}_{i,0} = 0$, the coefficient will volatile around a value of zero)
- (4) Both ω_i and β_i are not shrunk to 0; then, the i th variable is relevant for forecasting the oil price and the time-varying coefficient is unrestricted around zero

2.3. Hierarchical Parameter Shrinkage. The parameters of interest are $\beta = (\beta_1, \dots, \beta_m)'$, $\tilde{\beta}_t = (\tilde{\beta}_{1,t}, \dots, \tilde{\beta}_{m,t})'$, and $\omega = (\omega_1, \dots, \omega_m)'$; we can use the Bayesian LASSO shrinkage priors to estimate these parameters. According to the study of [28], the LASSO shrinkage can be obtained by starting from normal hierarchical priors for β and ω .

Hierarchy shrinkage 1: for the constant coefficients, the priori for β_i ($i = 1, \dots, m$) is independent with $\beta_i | \tau_i^2 \sim N(0, \tau_i^2)$ and exponential mixing density $\tau_i^2 | \lambda \sim \exp(\lambda^2/2)$. λ is the shrinkage parameter for constant coefficients and we assume $\lambda^2 \sim \text{Gamma}(a_1, a_2)$. So, the first hierarchy is conditional on λ to estimate τ_i^2 then obtain β_i .

Hierarchy shrinkage 2: from equation (4), we can infer that the time-varying parameters $\tilde{\beta}_t$ (for $t = 1, \dots, T$) prior is of the form $\tilde{\beta}_t | \tilde{\beta}_{t-1} \sim N(\tilde{\beta}_{t-1}, I_r)$, where $\tilde{\beta}_0 = 0$. The hierarchical priori of ω_i , conditionally independent with $\omega_i | \xi_i^2 \sim N(0, \xi_i^2)$, is also with exponential mixing density $\xi_i^2 | \kappa \sim \exp(\kappa^2/2)$. The shrinkage parameter κ lies at the bottom of the hierarchy and we assume $\kappa^2 \sim \text{Gamma}(b_1, b_2)$. The second hierarchy is conditional on κ ; we can in turn to derive ξ_i^2 and ω_i , at last, judge whether $\tilde{\beta}$ is time-varying or not.

For the two hierarchy shrinkage processes mentioned above, we set the prior hyperparameters $a_1 = a_2 = b_1 = b_2 = 0.001$, which implies proper but very noninformative priors. For constant coefficients model, which removes the TVP part of the model, we set $b_1 = 100000$ to make ω_i shrink very close to zero and its prior variance is 0.1.

To complete these two hierarchical shrinkage computations, [28] provides Markov Chain Monte Carlo (MCMC) algorithm blocks and precise steps to draw the parameter posteriors. After using a nonparametric kernel smoothing algorithm on the parameter posteriors, we can obtain an approximation of the oil price predictive density.

As LASSO shrinkage can be applied to both constant coefficients and time-varying coefficients, the full model and restricted models can derive several versions for the following:

- (1) LASSO on constant coefficients and time-varying parameters: both constant and time-varying part use LASSO priors and do hierarchical shrinkage.
- (2) LASSO only on constant coefficients: this model omits the time-varying part ($\sum_{i=1}^r \omega_i \tilde{\beta}_{i,t} z_{i,t}$ in equation (7)) LASSO priors and uses a relatively noninformative and nonhierarchical normal prior for ω_i .
- (3) LASSO only on time-varying parameters: this model omits the constant part ($\sum_{i=1}^r \beta_i z_{i,t}$) LASSO priors

and uses a relatively noninformative and nonhierarchical normal prior for β_i .

- (4) TVP regression model: this model is traditional time-varying multivariate parameter model which does not hierarchical shrinkage for parameters. Use noninformative LASSO priors for both ω_i and β .
- (5) Constant coefficients model: this model removes the time-varying part ($\sum_{i=1}^r \omega_i \tilde{\beta}_{i,t} z_{i,t}$) by setting prior hyperparameters $b_1 = 100000, b_2 = 0.001$, which implies an extremely tight prior on ω_i with prior concentrated very close to 0.

2.4. Evaluation Criteria. The results of predictive density or forecasting points from the previous steps are useful to quantitatively compare the out-of-sample predictive performance among different models. Following the convention in the literature on prediction measurement, we use point forecasting loss functions of MAFE and MSFE to demonstrate the ranking of model forecasts [17, 23, 29, 31]. Further, since researchers and policymakers focus more on total distribution forecast uncertainty than just a point forecast, we also adopt the mean of the log predictive likelihood (MLPL) to evaluate the entire predictive distributions. The specific formulations of these three measuring statistics are listed below:

$$\text{MAFE} = \frac{1}{T-h-t_0+1} \sum_{t=t_0}^{T-h} |\text{Oil}_{t+h} - \text{Oil}_{t+h}^0|,$$

$$\text{MSFE} = \frac{1}{T-h-t_0+1} \sum_{t=t_0}^{T-h} (\text{Oil}_{t+h} - \text{Oil}_{t+h}^0)^2,$$

$$\text{MLPL} = \frac{1}{T-h-t_0+1} \sum_{t=t_0}^{T-h} \log [p(\text{Oil}_{t+h} = \text{Oil}_{t+h}^0 | \text{Data}_t)]. \quad (8)$$

Respectively, T is the end date, t_0 is the start date, h is prediction length, Oil_{t+h} is the predictive median of oil price, and Oil_{t+h}^0 is the corresponding real value. Smaller MAFE and MSFE and larger MLPL indicate stronger forecasting ability.

3. Data

This paper uses two prevailing proxies in crude oil pricing: the monthly spot price of Brent crude oil as dependent variable and West Texas Intermediate (WTI) oil futures for robustness check. Both datasets span from January 2004 to December 2018 yielding $t = 180$ observations; the out-of-sample evaluation period consists of the last 110 observations.

On the foundation of previous studies [3, 16, 34, 38–40], we select a relatively comprehensive predictors framework to forecast crude oil price and use available real-time data. The exogenous variable dataset consists of crude oil fundamentals (include crude oil supply, demand, and stocks),

capital market prices (gold price, exchange rate, and stock market price index), substitute product price (natural gas price), market sentiment index (volatility index), macroeconomic influencing factors (industrial production and Kilian indexes), and political change (global policy uncertainty and Google trend). This variable set not only captures the information in both the supply and demand of crude oil but also includes activities related to the financial market and macroeconomy. Accordingly, they are widely used variables for crude oil price forecasting.

The ADF and PP test in Table 2 indicate that no variables have unit roots after first-order logarithmic difference, which means all the series are stationary time series, so we can use these series for further econometric modeling. The two dependent variables—WTI and Brent—are left-skewed, leptokurtic, and nonnormal distribution. Within 20 lags, the Q-statistics of both WTI index and Brent spot price series show significant autocorrelation, which suggest that past oil prices have influences on the current oil price, so it is reasonable to include AR terms in the model.

To examine whether the current oil price is affected by the past oil prices, we further include the logged first difference of 12 lags of the Brent crude oil price index in the model. In addition, an intercept and 11 monthly dummies (omitting the January dummy) are designed to distinguish monthly or seasonal effects on the crude oil prices.

All the explanatory variables are standardized to have mean zero and variance one. The model can flexibly include an intercept, different numbers of lags, 11 monthly dummies, and 12 predictors listed above. In addition, it can forecast oil prices a month ahead (short term) and a year ahead (long term).

In summary, the full variable model includes 36 coefficients to estimate with fewer than 15 years of data, which is a relatively short dataset. Omitting 12 predictors and 11 dummies, the model leads to AR models or TVP-AR models. If the lags are further excluded, it leads to TVP models or multivariate regression models. If only 11 dummies are left in the model, model form changes to UCSV model. In total, for each sample size rolling window estimation, we compute 20 different versions of full models and 100 competing models to check the models' robustness.

4. Empirical Results

4.1. Time-Varying and Shrinkage Parameters Results. This section focuses on time-varying and shrinkage coefficients represented by ω_i^2 and τ_i^2 . ω_i close to zero means the i th ($i = 1, \dots, m$) coefficient is constant over time; larger values of ω_i allow for more time variation. While the smaller value of τ^2 ensures a higher degree of shrinkage, larger τ^2 indicates the prior is more dispersed and shrinkage is less. In order to better explain the time-varying and shrinkage properties, we post the full model (LASSO shrinkage on both constant coefficients and time-varying parameters) results for Brent oil as an instant.

These results show moderate shrinkage in most coefficients, but the shrinkage degree varies. Table 3 shows that in one-month ahead ($h=1$) forecasting, ω_i^2 for crude oil

consumption, gold price, and industrial production index tend to shrink more than the coefficients on other variables, which indicate that the influence on crude oil price from these three variables is relatively time-invariant. In contrast, in short-term forecasting, the impacts of crude oil production on oil prices vary over time. The τ_i^2 of gas price, industrial production index, and the Kilian index shrink most among all exogenous variables; this signifies that the role of substitute product of oil, production level, and macroeconomic factors will not exert a significant effect on crude oil price in the short term. Instead, the three representative market uncertainty variables—VIX, EPU, and Google trend—show low-level shrinkage, so the policy uncertainty, market sentiment, and topic heat have a greater effect on the oil prices in the short term.

In the long-run ($h=12$) forecasting, crude oil stocks, SP500, and the Kilian index show larger ω_i^2 than other variables, which means larger time variation in these coefficients. Moreover, the Kilian index presents the largest τ^2 , which indicates that Kilian's index is a powerful predictor for oil price long-term forecasting. In the contrary, trade-weighted US dollar index, gas price, and production level are relatively unimportant factors.

Table 4 exhibits that the half-year ago oil prices have big impact on current oil price in the short-run forecasting; the influences from the end and beginning of the quarter are moderate. Table 5 depicts that the crude oil prices bear little relationship to the cycle of the seasons, because all the monthly dummies shrink more than the most of the other predictors and lags.

4.2. Forecasting Results Evaluation. In the tables, all the results are presented relative to the corresponding full model (LASSO on both constant coefficients and TVPs); smaller MAFE or MSFE, or larger MLPL than full model statistics indicate that the restricted model is forecasting better than the benchmark model.

The upper metrics of Table 6 results indicate that in one-month ahead forecasting, there is evidence that LASSO on constant coefficients outperforms other restricted models in both stochastic and constant volatility, which meet the short-term forecasting expectation that the majority coefficients do not change over time. Table 3 results are consistent with Table 6 and proved our opinion again.

In terms of the latter forecast metrics—the annual forecasting horizon—coefficients tend to show more time variation, so the full model has the best performance.

It is worth noting that the TVP regression models and constant coefficients models produce the worst forecasts in both cases according to MLPL. The results verified again that the new Bayesian hierarchical LASSO outperforms the traditional counterparts and enhances the prediction accuracy. Additionally, the bad performance of LASSO only on time-varying parameters indicates that the inclusion of time-variant parameters in the model is necessary for the oil prices forecasting.

To sum up, all results exhibit the advantages of the Bayesian hierarchical shrinkage. Firstly, putting a LASSO prior allows

TABLE 2: Descriptive statistics.

	WTI	BRENT	OD_CONS	OD_STOCKS	OS_WORLD	P_GOLD	US_EX
Mean	0.0009	0.0015	-0.0001	0.0003	0.0004	0.0028	0.0003
Median	0.0058	0.0069	--0.0008	0.0005	0.0005	0.0027	0.0002
Maximum	0.0929	0.0851	0.0242	0.0090	0.0094	0.0537	0.0440
Minimum	-0.1442	-0.1351	-0.0264	-0.0086	-0.0064	-0.0904	-0.0178
Std. dev.	0.0389	0.0395	0.0107	0.0033	0.0029	0.0222	0.0070
Skewness	-0.8849***	-0.9662***	-0.0293	-2.2210	-0.0453	-0.4846***	1.4488***
Kurtosis (excess)	1.7698***	1.5476***	-0.5727	-0.0127	-0.1753	1.4502***	8.2127***
Jarque-Bera	46.7237***	45.7129***	2.4715	1.4584	0.2903	22.6911***	565.6776***
Q(20)	45.1300***	40.8931***	282.3031***	108.7411***	43.0371***	27.1741	26.0612
PP test	-9.4826***	-9.7966***	-20.9879***	-14.3289***	-13.5263***	-15.2111***	12.0550***
ADF	-6.6839***	-9.7414***	-3.2129**	-3.7359***	-7.0902***	-15.1254***	-11.9771***
	SP500	P_GAS	VIX	IP_TOTAL	EPU	KILIAN	Google
Mean	0.0019	-0.0010	0.0010	0.0004	0.0035	13.7941	0.0046
Median	0.0045	-0.0034	-0.0072	-0.0005	0.0041	2.1302	0.0000
Maximum	0.0444	0.1649	0.3703	0.0162	0.2831	187.8978	0.3358
Minimum	-0.0806	-0.1766	-0.2111	-0.0222	-0.2448	-163.4310	-0.2320
Std. dev.	0.0172	0.0549	0.0887	0.0072	0.0835	79.2186	0.1000
Skewness	-1.0534***	0.1095	0.6745***	0.1728	0.3511*	0.3115*	0.8297***
Kurtosis (excess)	3.1059***	1.1509***	1.5779***	0.0032	1.2777***	-0.7538**	1.0390***
Jarque-Bera	105.0568***	10.2378***	32.1443***	0.8914	15.8510***	7.1318***	28.6037***
Q(20)	40.0941***	24.4420	33.8112**	219.2882***	33.3670**	1659.5120***	18.3661
PP test	-11.2842***	-13.9759***	-16.6098***	-20.3007***	-15.4433***	-2.1650**	-14.1582***
ADF	-5.3398***	-13.8972***	-7.2381***	-3.9125***	-10.8053***	-2.6820***	-14.0785**

Symbols *, **, and *** denote rejections of the null hypothesis at the 10%, 5%, and 1% significance levels, respectively. The Jarque-Bera statistic is used to test the null hypothesis of the normal distribution. Q(20) is the Ljung-Box Q statistics with lag order of 20. ADF refers to the statistics from the augmented Dickey-Fuller unit root tests. The entire sample period is from January 2004 to December 2018.

TABLE 3: Posterior means and standard deviation of ω_i^2 and τ^2 for exogenous predictors.

Predictor	$h = 1$		$h = 12$	
	ω_i^2	τ^2	ω_i^2	τ^2
INTERCEPT	1.691E-03	1.721E-02	5.852 E-02	4.832E-03
	4.552E-03	3.507E-02	1.455E-02	7.005E-03
OD_CONS	1.223E-03	1.668E-02	2.247E-04	5.939E-03
	2.203E-03	3.193E-02	3.202E-04	8.661E-03
OD_STOCKS	1.558E-03	2.320E-02	1.132E-03	5.094E-03
	2.308E-03	4.347E-02	2.140E-03	8.442E-03
OS_PROD	4.510 E-03	2.039E-02	3.591E-04	4.553E-03
	6.680E-03	3.548E-02	7.997E-04	6.569E-03
GOLD	6.845E-03	1.688E-02	3.195E-04	4.943E-03
	1.034E-02	3.506E-02	4.571E-04	7.695E-03
US_EX	1.507E-03	1.884E-02	6.058E-04	4.508E-03
	2.977E-03	3.919E-02	1.079E-03	8.013E-03
SP500	1.162E-03	2.338E-02	1.441E-03	5.481E-03
	2.443E-03	5.809E-02	1.869E-03	8.684E-03
GAS	9.481E-04	1.380E-02	6.770E-04	4.342E-03
	1.799E-03	2.929E-02	1.408E-03	8.098E-03
VIX	2.877E-03	2.454 E-02	4.482E-04	9.406 E-03
	5.570E-03	5.135E-02	5.983E-04	1.401E-02
IP	8.177E-04	1.453E-02	4.197E-04	4.817E-03
	1.542E-03	2.883E-02	8.323E-04	7.580E-03
EPU	2.595E-03	2.019E-02	3.585E-04	6.666 E-03
	6.928E-03	3.991E-02	1.467E-03	8.442E-03
KILIAN	2.884E-03	1.426E-02	1.337 E-03	1.239 E-02
	7.256E-03	3.152E-02	1.705E-03	1.370E-02
Google	3.367E-03	2.978 E-02	9.316E-04	7.450E-03
	6.438E-03	5.414E-02	9.836E-04	9.335E-03

Note. The bold text noted indicates relatively larger value among all ω_i^2 and τ^2 , while the underlined text represents values relatively smaller ones.

TABLE 4: Posterior means and standard deviation of ω_i^2 and τ^2 for lags.

Lags	$h = 1$		$h = 12$	
	ω_i^2	τ^2	ω_i^2	τ^2
1	4.292 E - 03	2.350E - 02	4.613E - 04	5.335E - 03
	8.873E - 03	4.947E - 02	1.124E - 03	7.308E - 03
2	1.407E - 03	1.517E - 02	7.182E - 04	4.762E - 03
	4.063E - 03	2.832E - 02	9.439E - 04	6.510E - 03
3	1.484E - 03	1.716E - 02	6.976E - 04	8.312E - 03
	3.116E - 03	3.761E - 02	1.287E - 03	1.249E - 02
4	1.040E - 03	1.587E - 02	4.530E - 04	4.833E - 03
	2.353E - 03	3.271E - 02	9.666E - 04	7.340E - 03
5	1.268E - 03	1.499E - 02	9.913E - 04	4.767E - 03
	3.672E - 03	3.030E - 02	1.152E - 03	7.669E - 03
6	9.445E - 04	1.849E - 02	7.573E - 04	4.690E - 03
	1.875E - 03	3.977E - 02	1.374E - 03	7.282E - 03
7	1.563E - 03	2.183 E - 02	3.827E - 04	4.349E - 03
	3.164E - 03	4.417E - 02	6.040E - 04	6.625E - 03
8	1.746E - 03	2.010E - 02	6.014E - 04	5.303E - 03
	3.673E - 03	3.964E - 02	1.525E - 03	8.904E - 03
9	1.387E - 03	1.336E - 02	3.603E - 04	4.652E - 03
	2.672E - 03	2.807E - 02	6.767E - 04	6.961E - 03
10	1.591E - 03	1.191E - 02	8.393E - 04	5.798 E - 03
	2.848E - 03	2.392E - 02	1.430E - 03	1.116E - 02
11	4.017E - 03	1.635E - 02	5.315E - 04	5.441E - 03
	7.294E - 03	3.271E - 02	1.024E - 03	7.460E - 03
12	1.307E - 03	1.288E - 02	1.031 E - 03	5.542E - 03
	2.731E - 03	2.699E - 02	1.667E - 03	7.646E - 03

Note. The bold text noted indicates relatively larger value among all ω_i^2 and τ^2 , while the underlined text represents values relatively smaller ones.

TABLE 5: Posterior means and standard deviation of ω_i^2 and τ^2 for monthly dummies.

Dummies	$h = 1$		$h = 12$	
	ω_i^2	τ^2	ω_i^2	τ^2
1	1.137 E - 02	3.553 E - 02	1.498E - 03	5.688E - 03
	2.505E - 02	7.022E - 02	2.571E - 03	1.037E - 02
2	1.972E - 03	1.617E - 02	1.490E - 03	6.749E - 03
	6.910E - 03	3.400E - 02	3.898E - 03	1.077E - 02
3	1.996E - 03	1.851E - 02	6.993E - 04	5.422E - 03
	5.533E - 03	4.477E - 02	1.383E - 03	8.975E - 03
4	1.788E - 03	1.732E - 02	8.008E - 04	5.116E - 03
	5.006E - 03	3.710E - 02	1.551E - 03	7.011E - 03
5	1.322E - 03	1.895E - 02	1.404E - 03	5.145E - 03
	3.293E - 03	4.164E - 02	2.564E - 03	7.389E - 03
6	7.115E - 03	1.769E - 02	1.425E - 03	6.803E - 03
	1.534E - 02	3.484E - 02	2.453E - 03	9.165E - 03
7	2.157E - 03	1.658E - 02	7.892E - 04	5.889E - 03
	5.331E - 03	3.614E - 02	1.423E - 03	9.500E - 03
8	1.415E - 03	1.627E - 02	3.182E - 03	5.242E - 03
	3.236E - 03	3.636E - 02	7.964E - 03	8.526E - 03
9	3.228E - 03	1.909E - 02	1.314E - 03	6.328E - 03
	8.431E - 03	3.732E - 02	3.359E - 03	9.615E - 03
10	2.274E - 03	1.932E - 02	6.981E - 04	5.538E - 03
	5.342E - 03	4.034E - 02	1.450E - 03	8.144E - 03
11	4.155E - 03	1.898E - 02	1.020E - 03	5.402E - 03
	9.748E - 03	5.034E - 02	1.884E - 03	7.023E - 03

Note. The bold text noted indicates relatively larger value among all ω_i^2 and τ^2 , while the underlined text represents values relatively smaller ones.

the data to decide whether the coefficients are time-varying and by how much they vary and restricts the TVP regression models coefficients wandering too widely which can obtain a better forecast performance. Secondly, in allusion of the misspecification problem, LASSO priors can automatically discover the lack of time variation in coefficients and shrinking the coefficients of unnecessary variables to zero, which improve the prediction accuracy and solve misspecification efficiently. Thirdly, hierarchical shrinkage in time-varying series facilitates researchers' start with a very flexible model with a relatively short dataset; the model results allow researchers and practitioners identify the most powerful predictors more efficiently then make the right investment decisions.

To investigate whether forecast performance varies over time, we present Figure 1, which uses the model with LASSO prior to both constant coefficients and time-varying parameters (TVPs) with forecasting horizon $h = 1$ (similar patterns are found with the other computation results).

From (a) and (b) in Figure 1, it can be seen that the constant and stochastic volatility versions of the model forecast roughly as well as each other; however, many conflicts occur during the time of the shale oil revolution in 2014. MAFE, MSFE, and MLPL will have a similar pattern for most of the time, but inconsistent during periods of oil price intense volatility. What is happening is that the heteroskedastic version includes too much increase in volatility which began with the shale oil revolution since MLPL measures the whole distribution prediction performance. This has little impact on the point forecasts MAFE and MSFE which do not differ by much between the constant and stochastic versions of the model.

5. Robustness Checks

5.1. *Robustness to Different Models' Specification.* Firstly, we conduct the robustness check by changing the variable set; the out-of-sample performance of AR, multivariate, and UCSV models are shown in the following tables.

Tables 7–9 indicate that, like the full model, smaller MAFE and MSFE and larger MLPL are also observed in LASSO on constant coefficients and LASSO on both constant and TVPs in AR, multivariate, and UCSV models. These results suggest that hierarchical shrinkage method can also outperform other competing models even with changes in the model structures.

5.2. *Robustness Check by Alternative Estimation Window.* In this section, we change the estimation window from the recursive rolling window to the rolling window; the results are shown in Table 10. LASSO on constant coefficients and TVPs and LASSO only on constant results are qualitatively similar in both rolling window and recursive rolling window.

Further, we change three in-sample window sizes suggested by [6, 46, 47] to check the robustness of hierarchical shrinkage models. In 40%, 50%, and 60% different out-of-sample evaluation periods, the results show that models with LASSO shrinkage exhibit lower MAFE and MSFE and higher

TABLE 6: Measures of forecast performance for log return of Brent with the full models.

	Constant variance			Stochastic variance		
	MAFE	MSFE	MLPL	MAFE	MSFE	MLPL
<i>Model ($h = 1$)</i>						
LASSO on constant and TVPs	0.072	0.009	2.886	0.088	0.013	2.126
LASSO only on constant coeff.	0.061	0.007	3.377	0.075	0.010	2.896
LASSO only on TVPs	0.112	0.021	1.970	0.107	0.018	1.696
TVP regression model	0.100	0.016	2.222	0.114	0.019	1.589
Constant coeff. model	0.100	0.016	2.189	0.106	0.017	1.658
<i>Model ($h = 12$)</i>						
LASSO on constant and TVPs	0.373	0.223	0.545	0.407	0.265	0.414
LASSO only on constant coeff.	0.389	0.224	0.468	0.463	0.340	0.325
LASSO only on TVPs	0.678	0.809	0.308	0.710	0.842	0.232
TVP regression model	0.648	0.707	0.295	0.709	0.833	0.226
Constant coeff. model	0.658	0.717	0.289	0.689	0.809	0.227

Note. The value noted in bold and underlined text indicates a model performing the best out of all models, while the bold and italic text represents a model performing the worst. MSFE, MAFE, and MLPL refer to the mean squared forecast error, mean absolute forecast error, and mean log predictive likelihood, respectively.

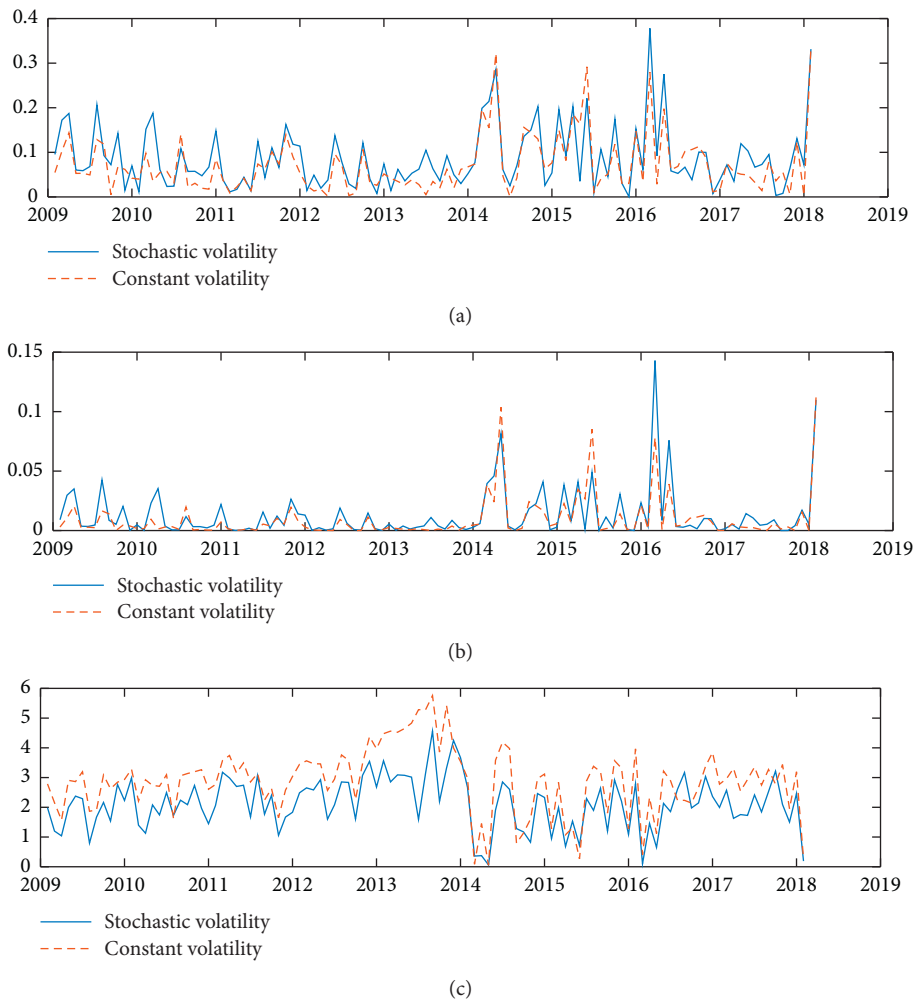


FIGURE 1: Forecasting performance measurement of models with LASSO prior on constant and time-varying coefficients, $h = 1$. (a) is the absolute forecast errors, (b) is the forecast errors squared, and (c) is the log predictive likelihood.

MLPL in most cases, suggesting the out-of-sample results of the hierarchical shrinkage are robust to different computation algorithms and sample sizes.

5.3. *Robustness to Alternative Dependent Variable.* Table 11 reports the main out-of-sample forecasting results of another prevailing proxy of crude oil prices, WTI. The

TABLE 7: Measures of forecast performance for log return of Brent with autoregression (AR) models.

	Constant variance			Stochastic variance		
	MAFE	MSFE	MLPL	MAFE	MSFE	MLPL
<i>Model (h = 1)</i>						
LASSO on constant and TVPs	0.067	0.008	2.825	0.089	0.014	1.969
LASSO only on constant coeff.	0.061	0.007	3.209	0.079	0.010	2.722
LASSO only on TVPs	0.098	0.015	1.989	0.104	0.016	1.822
TVP regression model	0.090	0.013	2.299	0.100	0.015	1.960
Constant coeff. model	0.091	0.013	2.275	0.097	0.014	1.996
<i>Model (h = 12)</i>						
LASSO on constant and TVPs	0.293	0.163	0.673	0.480	0.385	0.359
LASSO only on constant coeff.	0.337	0.207	0.502	0.462	0.321	0.313
LASSO only on TVPs	0.549	0.485	0.358	0.637	0.684	0.299
TVP regression model	0.557	0.523	0.337	0.586	0.533	0.281
Constant coeff. model	0.557	0.520	0.337	0.576	0.506	0.275

Note. The value noted in bold and underlined text indicates a model performing the best out of all models, while the bold and italic text represents a model performing the worst. MSFE, MAFE, and MLPL refer to the mean squared forecast error, mean absolute forecast error, and mean log predictive likelihood, respectively.

TABLE 8: Measures of forecast performance for log return of Brent with multivariate models.

	Constant variance			Stochastic variance		
	MAFE	MSFE	MLPL	MAFE	MSFE	MLPL
<i>Model (h = 1)</i>						
LASSO on constant and TVPs	0.061	0.007	3.350	0.077	0.011	2.696
LASSO only on constant coeff.	0.059	0.006	3.556	0.071	0.008	3.203
LASSO only on TVPs	0.069	0.009	3.022	0.081	0.011	2.383
TVP regression model	0.072	0.008	3.037	0.077	0.010	2.953
Constant coeff. model	0.072	0.008	3.029	0.076	0.010	2.874
<i>Model (h = 12)</i>						
LASSO on constant and TVPs	0.320	0.168	0.663	0.386	0.236	0.470
LASSO only on constant coeff.	0.328	0.165	0.538	0.303	0.153	0.546
LASSO only on TVPs	0.363	0.213	0.566	0.395	0.244	0.452
TVP regression model	0.376	0.221	0.480	0.314	0.175	0.482
Constant coeff. model	0.376	0.223	0.479	0.343	0.197	0.439

Note. The value noted in bold and underlined text indicates a model performing the best out of all models, while the bold and italic text represents a model performing the worst. MSFE, MAFE, and MLPL refer to the mean squared forecast error, mean absolute forecast error, and mean log predictive likelihood, respectively.

TABLE 9: Measures of forecast performance for log return of Brent ($h = 1$) with UCSV models.

	Constant variance			Stochastic variance		
	MAFE	MSFE	MLPL	MAFE	MSFE	MLPL
<i>Model (h = 1)</i>						
LASSO on constant and TVPs	0.064	0.007	3.111	0.076	0.010	2.969
LASSO only on constant coeff.	0.060	0.007	3.229	0.066	0.007	3.474
LASSO only on TVPs	0.069	0.009	2.879	0.075	0.010	2.750
TVP regression model	0.070	0.008	2.898	0.069	0.008	3.071
Constant coeff. model	0.070	0.008	2.909	0.068	0.008	3.148
<i>Model (h = 12)</i>						
LASSO on constant and TVPs	0.280	0.137	0.807	0.283	0.138	0.629
LASSO only on constant coeff.	0.297	0.157	0.530	0.269	0.117	0.774
LASSO only on TVPs	0.281	0.137	0.778	0.275	0.135	0.768
TVP regression model	0.298	0.160	0.511	0.276	0.121	0.706
Constant coeff. model	0.301	0.160	0.517	0.283	0.124	0.799

Note. The value noted in bold and underlined text indicates a model performing the best out of all models, while the bold and italic text represents a model performing the worst. MSFE, MAFE, and MLPL refer to the mean squared forecast error, mean absolute forecast error, and mean log predictive likelihood, respectively.

TABLE 10: Measures of forecast performance for log return of Brent with recursive rolling window results.

	Constant variance			Stochastic variance		
	MAFE	MSFE	MLPL	MAFE	MSFE	MLPL
<i>Model (h = 1)</i>						
LASSO on constant and TVPs	0.066	0.008	3.273	0.083	0.012	2.438
LASSO only on constant coeff.	0.061	0.007	3.460	0.076	0.010	2.955
LASSO only on TVPs	0.078	0.010	2.750	0.088	0.013	2.423
TVP regression model	0.074	0.009	3.003	0.079	0.010	2.625
Constant coeff. model	0.074	0.009	2.991	0.077	0.010	2.696
<i>Model (h = 12)</i>						
LASSO on constant and TVPs	0.315	0.170	0.712	0.391	0.260	0.463
LASSO only on constant coeff.	0.306	0.166	0.592	0.328	0.206	0.702
LASSO only on TVPs	0.363	0.228	0.597	0.439	0.318	0.406
TVP regression model	0.411	0.270	0.481	0.392	0.253	0.486
Constant coeff. model	0.412	0.274	0.480	0.370	0.223	0.479

Note. The value noted in bold and underlined text indicates a model performing the best out of all models, while the bold and italic text represents a model performing the worst. MSFE, MAFE, and MLPL refer to the mean squared forecast error, mean absolute forecast error, and mean log predictive likelihood, respectively.

TABLE 11: Measures of forecast performance for log return of Brent with WTI forecasting.

	Constant variance			Stochastic variance		
	MAFE	MSFE	MLPL	MAFE	MSFE	MLPL
<i>Model (h = 1)</i>						
LASSO on constant and TVPs	0.075	0.009	2.629	0.093	0.014	1.883
LASSO only on constant coeff.	0.063	0.007	3.131	0.080	0.011	2.476
LASSO only on TVPs	0.137	0.029	1.626	0.129	0.029	1.390
TVP regression model	0.115	0.022	1.891	0.111	0.021	1.413
Constant coeff. model	0.114	0.022	1.912	0.112	0.021	1.432
<i>Model (h = 12)</i>						
LASSO on constant and TVPs	0.333	0.202	0.612	0.388	0.265	0.422
LASSO only on constant coeff.	0.340	0.192	0.529	0.482	0.427	0.419
LASSO only on TVPs	0.642	0.732	0.308	0.696	0.859	0.218
TVP regression model	0.630	0.766	0.291	0.651	0.812	0.208
Constant coeff. model	0.629	0.771	0.293	0.673	0.856	0.204

Note. The value noted in bold and underlined text indicates a model performing the best out of all models, while the bold and italic text represents a model performing the worst. MSFE, MAFE, and MLPL refer to the mean squared forecast error, mean absolute forecast error, and mean log predictive likelihood, respectively.

results are quite close to Brent oil, which provides further support for the superiority of the hierarchical shrinkage method in alternative proxy of crude oil price forecasting.

6. Conclusions

In this paper, we predict the crude oil price based on the Bayesian hierarchical shrinkage method with a relatively short dataset and comprehensive variables framework. This method avoids overfitting and misspecification problems faced by linear regression prediction and improves the oil price forecasting accuracy. It also takes parameters dynamic properties into account. So, practitioners or policymakers can easily identify the most powerful indicators and do appropriate strategies during different periods.

The point and distribution forecasting performance statistics suggest that the hierarchical shrinkage models exhibit significantly better out-of-sample forecasting performance than other competing models in both homoskedasticity and heteroskedasticity versions. Our results are robust to a wide range of model settings, including various

model structures, different out-of-sample sizes, alternative estimation rolling windows, and crude oil proxies. Therefore, our study provides evidence regarding which indicators are informative and powerful to improve forecasting accuracy in the oil market.

Data Availability

The Brent and WTI crude oil price data are openly available on the website of EIA at https://www.eia.gov/dnav/pet/pet_pri_spt_s1_d.htm.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

References

- [1] L. Kilian, A. Rebucci, and N. Spatafora, "Oil shocks and external balances," *Journal of International Economics*, vol. 77, no. 2, pp. 181–194, 2009.

- [2] K. Lang and B. R. Auer, "The economic and financial properties of crude oil: a review," *The North American Journal of Economics and Finance*, vol. 52, 2019.
- [3] J. Li, S. Zhu, and Q. Wu, "Monthly crude oil spot price forecasting using variational mode decomposition," *Energy Economics*, vol. 83, pp. 240–253, 2019.
- [4] R. A. Ratti and J. L. Vespignani, "Oil prices and global factor macroeconomic variables," *Energy Economics*, vol. 59, pp. 198–212, 2016.
- [5] P. Zagaglia, "Macroeconomic factors and oil futures prices: a data-rich model," *Energy Economics*, vol. 32, no. 2, pp. 409–417, 2010.
- [6] Y. Zhang, F. Ma, and Y. Wang, "Forecasting crude oil prices with a large set of predictors: can lasso select powerful predictors?" *Journal of Empirical Finance*, vol. 54, pp. 97–117, 2019.
- [7] Y. Wang, L. Liu, and C. Wu, "Forecasting the real prices of crude oil using forecast combinations over time-varying parameter models," *Energy Economics*, vol. 66, pp. 337–348, 2017.
- [8] M. Balçilar and Z. A. Ozdemir, "The nexus between the oil price and its volatility risk in a stochastic volatility in the mean model with time-varying parameters," *Resources Policy*, vol. 61, pp. 572–584, 2019.
- [9] S. A. Ozdemir, I. John, K. Ivanovski, and R. Smyth, "Dynamics of oil price, precious metal prices and the exchange rate in the long-run," *Energy Economics*, vol. 84, p. 104508, 2019.
- [10] M. Guidolin, E. Hansen, and M. Pedio, "Cross-asset contagion in the financial crisis: a bayesian time-varying parameter approach," *Journal of Financial Markets*, vol. 45, pp. 83–114, 2019.
- [11] R. Kruse and C. Wegener, "Time-varying persistence in real oil prices and its determinant," *Energy Economics*, vol. 85, Article ID 104328, 2019.
- [12] R. Alquist, L. Kilian, and J. Robert, "Forecasting the price of oil," *Handbook of Economic Forecasting*, vol. 427–507, 2013.
- [13] J. Chai, J.-E. Guo, L. Meng, and S.-Y. Wang, "Exploring the core factors and its dynamic effects on oil price: an application on path analysis and bvar-tvp model," *Energy Policy*, vol. 39, no. 12, pp. 8022–8036, 2011.
- [14] L. Lei, Y. Shang, Y. Chen, and Y. Wei, "Does the financial crisis change the economic risk perception of crude oil traders? A midas quantile regression approach," *Finance Research Letters*, vol. 30, pp. 341–351, 2019.
- [15] R. Ma, C. Zhou, H. Cai, and C. Deng, "The forecasting power of epu for crude oil return volatility," *Energy Reports*, vol. 5, pp. 866–873, 2019.
- [16] J. Wang, G. Athanasopoulos, R. J. Hyndman, and S. Wang, "Crude oil price forecasting based on internet concern using an extreme learning machine," *International Journal of Forecasting*, vol. 34, no. 4, pp. 665–677, 2018.
- [17] J. Wang, X. Li, T. Hong, and S. Wang, "A semi-heterogeneous approach to combining crude oil price forecasts," *Information Sciences*, vol. 460–461, pp. 279–292, 2018.
- [18] F. Wen, M. Zhang, M. Deng, Y. Zhao, and J. Ouyang, "Exploring the dynamic effects of financial factors on oil prices based on a tvp-var model," *Physica A: Statistical Mechanics and Its Applications*, vol. 532, p. 121881, 2019.
- [19] Y. Zhang, F. Ma, and Y. Wei, "Out-of-Sample prediction of the oil futures market volatility: a comparison of new and traditional combination approaches," *Energy Economics*, vol. 81, pp. 1109–1120, 2019.
- [20] A. Jadidzadeh and A. Serletis, "The global crude oil market and biofuel agricultural commodity prices," *The Journal of Economic Asymmetries*, vol. 18, Article ID e00094, 2018.
- [21] M. S. Kim, "Impacts of supply and demand factors on declining oil prices," *Energy*, vol. 155, pp. 1059–1065, 2018.
- [22] W. Thorbecke, "Oil prices and the U.S. Economy: evidence from the stock market," *Journal of Macroeconomics*, vol. 61, p. 103137, 2019.
- [23] Y. Wei, J. Liu, X. Lai, and Y. Hu, "Which determinant is the most informative in forecasting crude oil market volatility: fundamental, speculation, or uncertainty?" *Energy Economics*, vol. 68, pp. 141–150, 2017.
- [24] Y. Wei, S. Qin, X. Li, S. Zhu, and G. Wei, "Oil price fluctuation, stock market and macroeconomic fundamentals: evidence from China before and after the financial crisis," *Finance Research Letters*, vol. 30, pp. 23–29, 2019.
- [25] Q. Wang, Y.-H. Chiu, and C.-R. Chiu, "Driving factors behind carbon dioxide emissions in China: a modified production-theoretical decomposition analysis," *Energy Economics*, vol. 51, pp. 252–260, 2015.
- [26] J. H. Stock and M. W. Watson, "Generalized shrinkage methods for forecasting using many predictors," 2011, http://www.princeton.edu/%7Emwatson_generalized_shrinkage_February_2011.pdf%200–62.
- [27] Y. Yi, F. Ma, Y. Zhang, and D. Huang, "Forecasting the prices of crude oil using the predictor, economic and combined constraints," *Economic Modelling*, vol. 75, pp. 237–245, 2018.
- [28] M. A. G. Belmonte, G. Koop, and D. Korobilis, "Hierarchical shrinkage in timevarying parameter models," *Journal of Forecasting*, vol. 94, pp. 80–94, 2011.
- [29] Y. Zhang, Y. Wei, Y. Zhang, and D. Jin, "Forecasting oil price volatility: forecast combination versus shrinkage method," *Energy Economics*, vol. 80, pp. 423–433, 2019.
- [30] H. Miao, S. Ramchander, T. Wang, and D. Yang, "Influential factors in crude oil price forecasting," *Energy Economics*, vol. 68, pp. 77–88, 2017.
- [31] J. Li and W. Chen, "Forecasting macroeconomic time series: lasso-based approaches and their forecast combinations with dynamic factor models," *International Journal of Forecasting*, vol. 30, no. 4, pp. 996–1015, 2014.
- [32] G. Kapetanios and F. Zikes, "Time-varying lasso," *Economics Letters*, vol. 169, pp. 1–6, 2018.
- [33] T. Park and G. Casella, "The bayesian lasso," *Journal of the American Statistical Association*, vol. 103, no. 482, pp. 681–686, 2012.
- [34] J. Chai, L.-M. Xing, X.-Y. Zhou, Z. G. au, and J.-X. Li, "Forecasting the WTI crude oil price by a hybrid-refined method," *Energy Economics*, vol. 71, pp. 114–127, 2018.
- [35] Q. Zhang, S. Li, and R. Li, "China's dependency on foreign oil will exceed 80% by 2030: developing a novel nmgm-arima to forecast China's foreign oil dependence from two dimensions," *Energy*, vol. 163, pp. 151–167, 2018.
- [36] M. Matyjaszek, P. Riesgo Fernández, A. Krzemień, K. Wodarski, and G. Fidalgo Valverde, "Forecasting coking coal prices by means of arima models and neural networks, considering the transgenic time series theory," *Resources Policy*, vol. 61, pp. 283–292, 2019.
- [37] E. A. Fidalgo Valverde, A. N. Çatik, and M. Balçilar, "The impact of oil prices on the stock returns in Turkey: a Tvp-var approach," *Physica A: Statistical Mechanics and Its Applications*, vol. 535, p. 122392, 2019.
- [38] T. Yao and Y.-J. Zhang, "Forecasting crude oil prices with the Google index," *Energy Procedia*, vol. 105, pp. 3772–3776, 2017.
- [39] S. Bekiros, R. Gupta, and A. Paccagnini, "Oil price forecastability and economic uncertainty," *Economics Letters*, vol. 132, pp. 125–128, 2015.

- [40] A. Lanza, M. Manera, and M. Giovannini, "Modeling and forecasting cointegrated relationships among heavy oil and product prices," *Energy Economics*, vol. 27, no. 6, pp. 831–848, 2005.
- [41] V. K. Singh, P. Kumar, and S. Nishant, "Feedback spillover dynamics of crude oil and global assets indicators: a system-wide network perspective," *Energy Economics*, vol. 80, pp. 321–335, 2019.
- [42] Z. Mandalinci, "Forecasting inflation in emerging markets: an evaluation of alternative models," *International Journal of Forecasting*, vol. 33, p. 22, 2017.
- [43] Bo Zhang, J. C. C. Chan, and J. L. Cross, "Stochastic volatility models with arma innovations: an application to G7 inflation forecasts," *International Journal of Forecasting*, vol. 36, no. 4, 2020.
- [44] J. C. C. Chan, "Moving average stochastic volatility models with application to inflation forecast," *Journal of Econometrics*, vol. 176, p. 10, 2013.
- [45] G. V. Moura and E. Douglas Turatti, "Efficient estimation of conditionally linear and Gaussian state space models," *Economics Letters*, vol. 124, p. 5, 2014.
- [46] B. Rossi and A. Inoue, "Out-of-Sample forecast tests robust to the choice of window size," *Journal of Business & Economic Statistics*, vol. 30, pp. 432–453, 2011.
- [47] A. Inoue, L. Jin, and B. Rossi, "Rolling window selection for out-of-Sample forecasting with time-varying parameters," *Journal of Econometrics*, vol. 196, no. 1, pp. 55–67, 2017.