

Article

Influencing Factors Analysis of Crude Oil Futures Price Volatility Based on Mixed-Frequency Data

Congxin Wu, Xinyu Wang *, Shan Luo, Jing Shan and Feng Wang *

School of Economics and Management, China University of Mining and Technology, No. 1, College Rd., Tongshan Dist., Xuzhou 221116, China; wucongxin@cumt.edu.cn (C.W.); ShanLuo97@cumt.edu.cn (S.L.); shanjing@cumt.edu.cn (J.S.)

* Correspondence: wangxinyu@cumt.edu.cn (X.W.); fengwang@cumt.edu.cn (F.W.)

Received: 8 October 2020; Accepted: 23 November 2020; Published: 25 November 2020



Abstract: This article takes into account the form of mixed data as well as the peak and thick tail characteristics contained in the data characteristics, expands the GARCH-MIDAS (Generalized Autoregressive Conditional Heteroskedasticity-Mixed Data Sampling) model, establishes a new GARCH-MIDAS model with the residual term of the skewed-t distribution, and analyzes the influence factors of crude oil futures price volatility, which can better explain the changing laws of crude oil price volatility. The results show the following: First, the low-frequency factors include crude oil production, consumption, inventory, and natural gas spot price, and the high-frequency factors include on-market trading volume and off-market spot price, which can significantly explain the volatility of oil price. Second, low-frequency factors include crude oil inventory, consumption, crude oil production, and speculative factors, and high-frequency factors include crude oil spot price and substitute prices. The increase in the volatility of trading volume is significantly positively correlated with oil price volatility, and the overall volatility model outperforms the horizontal effect model. Third, from the perspective of the combined effect of a single factor level and volatility, we find that supply and demand are the low-frequency factors; the trading volume of on-market factors, natural gas price, and crude oil spot price of off-market factors, among the high-frequency factors, are the most important factors affecting oil price volatility. Fourth, from the perspective of high-frequency and low-frequency effects combined, there is no significant difference between the various factor models, which shows that the mixed effect model of high and low frequency models has advantages in terms of the stability of the estimation results.

Keywords: crude oil futures; volatility; GARCH-MIDAS; mixed-frequency data

1. Introduction

The quantity and quality of statistical data are the cornerstones of macroeconomic model construction, estimation, and forecasting. How much more information can be extracted from economic data has always been a goal worthy of pursuit. As a commodity, crude oil plays an important role in the real economy and the financial market. As a major supplier of energy and raw materials for many industrial materials and agricultural sectors, it has gradually become an investment tool and can even impact the business cycle in the financial market [1].

The boom of commodities that began in 2004 saw international oil prices rise steadily until 2008, along with the increase of frequency and extent of oil price fluctuations. When the subprime crisis broke out in 2008, speculative funds left the market one after another and the dollar strengthened, which caused commodity prices around the world to fall sharply at the same time. The price of crude oil dropped by 70% from its peak in the latter half of 2018, which far exceeded the 40% decline in copper, corn and soybean futures, indicating that oil prices are more volatile. After the financial crisis,

the oil price fluctuated and reacted more to external factors. The psychological expectations of market participants and the existence of speculative factors aggravated the fluctuation range of the oil price, and the ensuing risks in the crude oil market and the crude oil futures market were also greater [2].

The volatility shock of crude oil price has an important impact on all markets [3,4]. The measurement and prediction of volatility are also of great significance for asset pricing and risk management, and the price volatility and volatility changes of commodities will also affect the pricing of commodity derivatives and the construction of the optimal hedging ratio [5]. Therefore, the study of fluctuation factors becomes the focus of research. Simple fundamental factors can no longer fully explain the volatility of oil prices. Existing studies have explored some factors that cause the volatility of crude oil prices and crude oil futures prices. The factors include OPEC's market and demand structure and demand elasticity [6], OPEC and the U.S. strategic petroleum reserve (SPR) announcement for crude oil spot and futures markets [7], the information efficiency of crude oil inventory positions [8,9], speculative factors [10], and irregular extreme events [2], natural disasters [11], geopolitical risks (GPRS) [12], even including the "long memory" of oil price fluctuation [13].

For oil futures, compared to the fundamentals of supply and demand, macroeconomic and financial data have become more important factors affecting oil futures prices [14,15]. Although traditional research methods have certain effects, they still have obvious disadvantages. As for the study of crude oil price volatility, the previous literature mainly included several methods. One was the GARCH (Generalized Autoregressive Conditional Heteroskedasticity)-class models [16,17]. After comparing nine different GARCH models (RiskMetrics, GARCH, IGARCH, GIR, EGARCH, APARCH, FIGARCH, FIAPARCH, and HYGARCH (Hyperbolic Memory Generalized Autoregressive Conditional Heteroscedasticity) models), and it was found that none of them were better than other models [18]. Although the addition of oil volatility index (OVX)-based implied volatility can greatly improve the forecast accuracy of crude oil prices compared with the simple GARCH-TYPE volatility models [17], most of these models are difficult to deal with the different frequencies between crude oil earnings volatility and macroeconomic covariables. Another method is that some researchers predicted the actual fluctuation of oil price based on the historical data in the past [19]. This method ignores various external factors and long-term stable conditions that affect the fluctuation of oil price. Previous studies have shown that the large fluctuations of WTI (West Texas Intermediate) crude oil market have great instability in both the short and long term, that is, the crude oil market is ineffective in the weak form [20], while the addition of macroeconomic uncertainties can improve the accuracy of oil price volatility assessment. The realized volatility model was to extract forecast information from historical volatility or prices [21]. Macro variables were not included in this model, and previous and current oil price fluctuations could not reflect and contain all available information [20].

The form of mixed data contained in the influencing factor data needs to be emphasized and studied in risk modeling. Due to many factors that affect crude oil price fluctuations, the complicated time series data and the various data sampling frequencies, as well as the factors involved in the data of supply and demand fundamentals, the macroeconomy, and the financial market, will change because of varying efficiency and difficulty of statistical work. Research must consider the form of mixed data [20]. However, the current research mostly uses the same frequency data modeling, which is a method of adding up or replacing. The high-frequency data are processed as low-frequency data, which causes the loss of sample information in the high-frequency data, and will artificially reduce the fluctuation of high-frequency financial data to a certain extent, and cannot meet the timeliness requirements. The high-frequency data are obtained by interpolating low-frequency data using mathematical methods that often lack economic theoretical support and cannot ensure accuracy. At the same time, in the current mixed modeling process, the statistical properties of the financial data, such as the peak and thick tail characteristics, are often been ignored, which may lead to the use of inappropriate models to analyze financial phenomena. Therefore, in the context of the complex characteristics of macroeconomic and financial market data, different data characteristics of different types of influencing factors need to be paid full attention in the research. This article will use the GARCH-MIDAS method,

considering the form of mixed data contained in crude oil futures data and the peak and thick tail characteristics of returns, to construct a mixed measurement model and to explore the influence of various factors with different data frequencies on crude oil futures price volatility.

2. Literature Review

The influencing factors of international crude oil price fluctuations can't be attributed solely to a single factor or a certain leading factor. Oil price fluctuations are a complex and changeable process and are the result of a combination of multiple factors. In the selection of factors affecting the fluctuation of international crude oil prices, some time series' adopt daily data, such as the US dollar exchange rate, but some adopt low-frequency monthly data, such as US crude oil inventories and OPEC crude oil production, so both high-frequency and low-frequency data appear simultaneously in the crude oil price volatility model. The processing method of most scholars has involved carrying out preprocessing on the data so that they could have the same frequency. For example, the mixed data are processed into the same frequency data by summing or interpolation, but this will cause the data to be unable to express all the information and finally affect the model results. Therefore, mixed data modeling can be used to reduce information loss caused by data processing.

On the study of oil price fluctuations and forecast, researchers believe that oil futures market shall be as a forecasting method. By applying four kinds of evaluation criteria, the futures market and the expert system for the EIA (Electronic Industries Alliance) were compared with each other from the aspects of prediction error and researchers finally found the NYMEX futures market situation is not inferior in terms of price forecast [22]. However, subsequent studies have reached the opposite result. The granger causality test method has been used to confirm that the forward curve is useful for one week and one month in advance in the daily forecast and the weekly forecast, respectively, but almost useless in the long-term forecast [23].

More researchers have used a variety of modeling methods for research. There have been studies chose a generalized regime switching model, which performs noticeably better than nonswitching models regardless of evaluation criteria [5]. Using the dynamic Nelson–Siegel model to explain the term structure of crude oil prices and using a generalized regression framework based on neural networks to forecast oil prices can be more accurately [24]. Developing three kinds of new HAR (Heterogeneous Autoregressive Model)-type models by incorporating investor sentiment and/or leverage effect in the corresponding original HAR-type models can be used to examine the predictable effect of investor sentiment and leverage on daily, as well as weekly and monthly fluctuations in the crude oil futures market [25]. Some scholars believe that using a simple autoregressive norm to simulate long memory, and using 11 predictive HAR-type time-series models to forecast realized variance can achieve the same effect as the more complex models [26]. More specifically, the TGARCH (Threshold Generalized Autoregressive Conditional Heteroskedasticity) model applies to the volatility of heating oil and natural gas, while the GARCH model applies to the volatility of crude oil and leaded gasoline. Simple moving average models can also work well if choosing the right order. Some complex models are even less effective than single-equation GARCH models. In the backtracking test, the nonparametric model for calculating risk measurement is superior to the parametric model in terms of transcendental number [27]. Some researchers studied the relationship between crude oil and natural gas prices, examined the behavior of natural gas and crude oil price volatility, and measured the extent to which crude oil volatility is displayed in long-term gas prices [28–31]. However, some researchers showed the evidence of causality from gas to oil [32].

The mixed data sampling method (MIDAS) was first proposed in 2007 [33]. Subsequently, a generalized autoregressive conditional heteroscedasticity model of mixed data sampling, namely GARCH-MIDAS, was proposed in 2013 [34]. The GARCH-MIDAS model divides the conditional variance into long-term and short-term variables, and low-frequency variables affect the conditional variance through long-term components. The effectiveness of the GARCH-MIDAS model in predicting the volatility of the stock market has been proved, and this model has been gradually

applied to studies [35]. Past researches have used the GARCH-MIDAS framework to re-examine the relationship between oil and the stock market by analyzing the long-term macroeconomic determinants of the daily stock market and crude oil price returns in the United States [36]. Some scholars also used the GARCH-MIDAS model to study the effects of oil supply and demand levels and fluctuations in the price fluctuations of WTI and Brent crude oil from 1986 to 2015 [37].

The GARCH-MIDAS model is also gradually used in predicting oil price. Some scholars forecasted oil market volatility using the probability (or intensity) of jump occurrences, the sign (e.g., positive or negative) of jumps, and the concurrence with stock market jumps, which are all under the mixed data sampling (MIDAS) modeling framework. The empirical results showed that this model can markedly improve the accuracy of oil market volatility forecasting [38]. Existing factors influencing oil price fluctuations to improve study and forecast changes in the oil market include speculation, economic fundamentals [39,40], basic (physical) predictors, financial predictors [41], uncertainty of geopolitical risk [42], economic policy uncertainty (EPU), and indicator variables [43,44]. In order to improve the accuracy of forecasts, scholars have suggested that the use of high-frequency financial data is more helpful in predicting oil prices [45]. Some scholars used MIDAS models and high-frequency data from four stock market indexes to predict WTI and Brent crude oil prices at lower frequencies. The results showed that the high frequency stock market index has a certain advantage over the low-frequency data in predicting the monthly crude oil price, and the MIDAS model of high-frequency data is superior to the common model [46].

However, most of the existing literature studies are based on the basic assumption of normal distribution, which does not meet the characteristics of financial data's spikes and thick tails, and these results cannot fully reflect market conditions. Compared with the normal distribution, a t-distribution can more accurately describe the characteristics of the "heavy tail" of the data, and thus can be closer to the actual distribution of financial data. At present, the t-distribution is mostly used in the research on the return and volatility of financial data such as stocks, but the t-distribution still has a relatively important reference for the study of crude oil price fluctuations.

In summary, the current researches on the volatility of the crude oil market are still not thorough and it can be manifested in two aspects: ① the traditional single-frequency data econometric model often ignores different frequency data information, which reduces the model's interpretation ability. ② from the perspective of the comprehensive characteristics of financial data, it lacks consideration of the complex characteristics of financial market data. In the process of modeling, the statistical properties of the financial data itself, such as peaks and thick tails, are often ignored, which may lead to the use of inappropriate models to analyze financial phenomena.

3. Model Construction of Crude Oil Futures Price Fluctuation Based on Mixed-Frequency Data

3.1. Improved GARCH-MIDAS Volatility Model

In order to re-examine the relationship between stock market volatility and macroeconomic activities and their volatility, the GARCH-MIDAS model was constructed by introducing the MIDAS method into the GARCH model [34]. Different information may have different effects on financial markets, depending on their short-term or long-term effects. In the GARCH-MIDAS model, the volatility of the high-frequency dependent variable is decomposed into the product of long-term low-frequency volatility components and short-term high-frequency volatility components; the low-frequency volatility component is then explained by the weighted level or weighted volatility of the low-frequency variable. The high-frequency volatility component is described by the traditional GARCH structure.

According to the research of Engle and Rangel, the GARCH-MIDAS model constructed in this paper is expressed as follows:

$$r_{i,t} = \mu + \sqrt{\tau_t \times g_{i,t}} \varepsilon_{i,t}, \quad \forall i = 1, \dots, N_t, \quad (1)$$

where $r_{i,t}$ represents the logarithmic return rate of crude oil futures prices on the i th day of month t ; $\varepsilon_{i,t}|\Phi_{i-1,t} \sim N(0,1)$, and $\Phi_{i-1,t}$ is the set of information available on the $(i-1)$ th day during t . Volatility has two components: the short-term dynamic component of daily volatility $g_{i,t}$ and the long-term component of volatility τ_t . $g_{i,t}$ is related to daily liquidity and other possible short-term factors, and τ_t is first related to expected future cash flows and future discount rates. Assume that the components $g_{i,t}$ follow the (day) GARCH (1,1) process:

$$g_{i,t} = (1 - \alpha - \beta) + \alpha(r_{i-1,t} - \mu)^2 / \tau_t + \beta g_{i-1,t}. \quad (2)$$

Furthermore, the MIDAS regression method is used to characterize the τ_t components based on the realized volatility (RV) of the rate of return:

$$\tau_t = m + \theta \sum_{k=1}^K \varphi_k(\omega_1, \omega_2) RV_{t-k}, \quad (3)$$

$$RV_t = \sum_{i=1}^{N_t} r_{i,t}^2, \quad (4)$$

where K represents the maximum lag order of the low-frequency variable, and RV_t represents the realized volatility in the month t . The weighting function in the equation adopts the Beta function, and the specific form is

$$\varphi_k(\omega) = \frac{(k/K)^{\omega_1-1} (1-k/K)^{\omega_2-1}}{\sum_{j=1}^K (j/K)^{\omega_1-1} (1-j/K)^{\omega_2-1}}. \quad (5)$$

Equation (1) to (5) together constitute a time-varying conditional variance GARCH-MIDAS model based on realized volatility and parameter space $\Theta = \{\mu, \alpha, \beta, m, \theta, \omega_1, \omega_2\}$. Compared with the basic GARCH model, this model is equivalent to a new type of component equation.

According to the foregoing analysis, the long-term component of crude oil futures price fluctuations has a strong correlation with the fluctuations of its influencing factors. Equation (3) is replaced, and GARCH-MIDAS models based on the level and volatility of their influencing factors are established, respectively. For horizontal effects,

$$\tau_t = m_l + \theta_l \sum_{k=1}^{K_l} \varphi_k(\omega_{1,l}, \omega_{2,l}) X_{l,t-k}^{mv}, \quad (6)$$

where K_l is the maximum lag order of the measured level of a certain macroeconomic variable and $X_{l,t-k}^{mv}$ represents the level value of the macroeconomic variable lagging k period relative to period t .

For the volatility effect, the volatility value of the explanatory variable must first be obtained. This article considers the use of AR(p) autoregressive model to obtain residual items, and the square of residuals is used to measure the volatility of macroeconomic variables. Thus,

$$\tau_t = m_v + \theta_v \sum_{k=1}^{K_v} \varphi_k(\omega_{1,v}, \omega_{2,v}) X_{v,t-k}^{mv}, \quad (7)$$

where K_v is the maximum lag order of the measured fluctuation value of a certain macroeconomic variable and $X_{v,t-k}^{mv}$ represents the fluctuation value of the macroeconomic variable relative to period t lagging by period k .

This article no longer considers the influence of the low-frequency volatility of the dependent variable.

In summary, when the long-term component equation contains only one influencing factor, Equation (1), (2), (5), and (6) together constitute the single-factor horizontal effect GARCH-MIDAS model, and Equation (1), (2), (5), and (7) together constitute the single factor volatility effect GARCH-MIDAS model.

When the influencing factors include two or more than two, the resulting mixed model is a multifactor GARCH-MIDAS model. The multifactor mixed model is different from the single-factor

mixed model in the setting of long-term components. It is a model that includes the level effect and fluctuation effect of all sequences at the same time:

$$\tau_t = m_{lv} + \theta_l \sum_{k=1}^K \varphi_k(\omega_{1,l}, \omega_{2,l}) X_{l,t-k}^{mv} + \theta_v \sum_{k=1}^K \varphi_k(\omega_{1,v}, \omega_{2,v}) X_{v,t-k}^{mv} \quad (8)$$

In this way, Equation (1), (2), (5), and (8) together constitute the multifactor GARCH-MIDAS model.

At the same time, in order to make full use of high-frequency influencing factor information, this paper considers the improvement of the classic GARCH-MIDAS model, adopts the method of realized volatility, and incorporates high-frequency macro-factors into the long-term component equation. Let $r_{Gi,t}$ be the logarithmic return rate of a high-frequency macro variable on day i of month t :

$$G_{tl} = \sum_{i=1}^{N_t} r_{Gi,t} \quad (9)$$

$$G_{tv} = \sum_{i=1}^{N_t} r_{Gi,t}^2, \quad (10)$$

Here, because the realized volatility method is used to convert the high-frequency macroeconomic sequence into low-frequency volatility, G_{tl} and G_{tv} represent the level value and volatility value of the high-frequency macro-variable returns, respectively. The corresponding long-term composition equation is as follows:

$$\tau_t = m_{l(v)} + \theta_{l(v)}^1 \sum_{k=1}^{K_{l(v)}} \varphi_k(\omega_{1,G}, \omega_{2,G}) G_{l(v),t-k} + \sum_{n=2}^N \theta_{l(v)}^n \sum_{k=1}^{K_{l(v)}} \varphi_k(\omega_{1,l(v)}, \omega_{2,l(v)}) X_{l(v),t-k}^{mv} \quad (11)$$

Therefore, Equation (1), (2), (5), and (9)–(11) together constitute a multifactor improved GARCH-MIDAS model incorporating high-frequency explanatory variables.

In particular, since there may be negative numbers in the macro data, the long-term component equation is taken in logarithmic form to ensure that the variance is nonnegative.

Thus, Equation (6) can be replaced by

$$\log \tau_t = m_l + \theta_l \sum_{k=1}^{K_l} \varphi_k(\omega_{1,l}, \omega_{2,l}) X_{l,t-k}^{mv} \quad (12)$$

Equation (7) can be replaced by

$$\log \tau_t = m_v + \theta_v \sum_{k=1}^{K_v} \varphi_k(\omega_{1,v}, \omega_{2,v}) X_{v,t-k}^{mv} \quad (13)$$

Equation (8) can be replaced by

$$\log \tau_t = m_{lv} + \theta_l \sum_{k=1}^K \varphi_k(\omega_{1,l}, \omega_{2,l}) X_{l,t-k}^{mv} + \theta_v \sum_{k=1}^K \varphi_k(\omega_{1,v}, \omega_{2,v}) X_{v,t-k}^{mv} \quad (14)$$

Equation (11) can be replaced by

$$\begin{aligned} \log \tau_t = m_{l(v)} + & \theta_{l(v)}^1 \sum_{k=1}^{K_{l(v)}} \varphi_k(\omega_{1,G}, \omega_{2,G}) G_{l(v),t-k} \\ & + \sum_{n=2}^N \theta_{l(v)}^n \sum_{k=1}^{K_{l(v)}} \varphi_k(\omega_{1,l(v)}, \omega_{2,l(v)}) X_{l(v),t-k}^{mv} \end{aligned} \quad (15)$$

3.2. GARCH-MIDAS-Skewed- t Model

In the traditional GARCH-MIDAS model, the residual items $\varepsilon_{i,t}$ obey the standard normal distribution; that is, $\varepsilon_{i,t} | \Phi_{i-1,t} \sim N(0,1)$ the traditional normal distribution of which does not well deal

with the “spikes and thick tails” and biases that exist in financial time series data. Hansen (1994) et al. introduced the skewed-t distribution into the ARCH family model and found that, when the residual term distribution is assumed to be a skewed-t distribution, the fitting effect significantly outperforms the fitting effect when the residual term is a normal distribution [47]. With reference to Hansen et al., this paper introduces the skewed-t distribution as the distribution state of the residual items in the GARCH-MIDAS model. The density function of the standard Skewed-t distribution is as follows:

When $Z < -a/b$, the density function of the skewed-t distribution is

$$g(z|\eta, \lambda) = bc \left(1 + \frac{1}{\eta - 2} \left(\frac{bz + a}{1 - \lambda} \right)^2 \right)^{-(\eta+1)/2} \quad (16)$$

When $Z \geq -a/b$, the density function of the skewed-t distribution is

$$g(z|\eta, \lambda) = bc \left(1 + \frac{1}{\eta - 2} \left(\frac{bz + a}{1 + \lambda} \right)^2 \right)^{-(\eta+1)/2} \quad (17)$$

where $2 < \eta < \infty$, and $-1 < \lambda < 1$.

$$a = 4\lambda c \left(\frac{\eta - 2}{\eta - 1} \right) \quad (18)$$

$$b^2 = 1 + 3\lambda^2 - a^2 \quad (19)$$

$$c = \frac{\Gamma(\frac{\eta+1}{2})}{\sqrt{\pi(\eta-2)}\Gamma(\frac{\eta}{2})} \quad (20)$$

where η is the degree of freedom. When λ is greater than 0, the density function is skewed to the right; when λ is less than 0, the density function is skewed to the left; if λ is equal to 0, the skewed-t distribution is the standard t distribution, the mean is 0, and the variance is 1.

As far as the characteristics of the distribution function are concerned, although the skewed-t distribution can better characterize the “spikes and thick tails” of financial data and its biased phenomenon than the standard normal distribution, this conclusion is not yet verified in the research on the GARCH-MIDAS model. Therefore, it is necessary to introduce the skewed-t distribution of the residual term in the GARCH-MIDAS model.

4. Variables and Data Description

This section will first define the variables used in the empirical research of this article and describe their basic characteristics; the GARCH-MIDAS empirical model of the crude oil futures market will then be constructed, and empirical analysis will be conducted. The research results will be compared based on different data types (same frequency and mixed frequency). The GARCH model based on the same frequency data will be constructed, and the analysis of the volatility effect will be conducted. Finally, whether the analysis is based on the mixed data or the same frequency data, the overall analysis thinking from single factor to multiple factors will be followed.

4.1. Variable Definition

The volatility of crude oil futures prices refers to the rise and fall of crude oil futures prices within a certain period. The study of international crude oil price fluctuations should be based on the study of volatility factors, and the volatility effect which is the volatility risk of price returns should be further studied. At present, in the international crude oil pricing system, Brent crude oil futures and the Brent crude oil spot market have the most significant influence. About 70% of the crude oil spot transaction price reference is Brent crude oil. Two-thirds of the global crude oil market sets the price of Brent crude oil as pricing benchmarks, including China's refined oil price reference mechanism.

(1) Volatility variables

The paper selects Brent light crude oil futures (Brent) (USD/barrel), one of the most important international crude oil futures, as the representative of international crude oil futures. The data came from the website of the US Department of Energy Information Administration (EIA, <https://www.eia.doe.gov>). The daily closing price of Brent is used to construct the volatility indicator of the oil market, and the logarithmic return is taken as the explained variable (RF). The variable range is from January 2008 to December 2017, including multiple stages of the international crude oil spot price surge, plummet, recovery, high rise, and continuous decline in the past 10 years, which makes the research conclusions credible.

Regarding the choice of low-frequency and high-frequency explanatory variables, three aspects are mainly considered: theoretical relevance, existing literature practices, and data availability.

(2) Low-frequency explanatory variables

OPEC crude oil production (OPEC) and OECD crude oil consumption (OECD) were selected to represent crude oil supply and crude oil consumption demand, respectively [48,49]; US crude oil commercial inventory (Inventory, hereinafter referred to as US crude oil inventory) was selected to represent crude oil inventory level [50]; Australia's Newcastle Port Power Coal spot offshore price (Cspp, hereinafter referred to as coal price) and the Henry Hub natural gas spot price (GasSpotPrice, hereinafter referred to as natural gas price) were selected to represent alternative energy sources as explanatory variables [51]; WTI crude oil noncommercial arbitrage holdings (Wtioi, hereinafter referred to as noncommercial Sexual arbitrage positions) represent speculation as an explanatory variable [52].

Among the above variables, data on US crude oil inventories, OPEC crude oil production, OECD crude oil consumption, and Henry Hub natural gas spot price data came from the US Department of Energy Information Administration (EIA) website. The FOB (Free On Board) spot data of thermal coal of Australia's Newcastle Port and the WTI crude oil noncommercial arbitrage holdings are derived from Wind Database (<https://www.wind.com.cn>). The low frequency indicators are all monthly data, and the interval is from January 2008 to December 2017.

(3) High-frequency explanatory variables

In the process of selecting explanatory variables, in addition to considering low-frequency variables, high-frequency variables (daily data) are also considered in order to increase the sufficiency of the explanation. It needs to be pointed out that high frequency and low frequency are relative concepts, not absolute. Considering the relevance of the indicators and the availability of data, we selected the closing price of the US dollar index [53], the trading volume of Brent crude oil futures [54], the open interest of Brent crude oil futures (empty volume) [55], the Henry Hub Natural Gas Spot Price (Dollars per Million Btu) [32], and the Europe Brent Spot Price FOB as explanatory variables [56]. Among them, the US dollar index (DI), the Brent crude oil spot price (PS), and the natural gas price (GP) represent high-frequency factors outside the market, and volume (V) and open interest (OI) represent high-frequency factors on the market. The reasons why natural gas price might be an influencing factor of oil price worth mentioning. Researchers started the empirical analysis by splitting the data into subsamples, before and after the start of shale gas, and test for cointegration in the second subsample which was during the period from 2007 to 2013. Analyzing the impulse response functions of the three fitted models can note a relevant difference between the first and second sample. The impact of gas (or gas quantity) shocks is persistent on oil (or oil price) in the second sub-sample [32]. The data comes from Wind. The high-frequency data are all daily data, and the sample interval is from 1 January 2008 to 31 December 2017.

Variable description and data sources are shown in Table 1.

Table 1. Variable descriptions and data sources.

Variable Name	Variable Description	Unit	Data Sources	Meaning	Data Frequency
Volatility variable					
RF	Brent crude oil futures price daily closing price logarithmic return	USD/barrel	EIA, US Energy Information Administration	Volatility indicators in the crude oil market	Daily
Low-frequency explanatory variables (all monthly data)					
OPEC	OPEC crude oil production	Million barrels/day	EIA, US Energy Information Administration	Crude oil supply factors	Monthly
OECD	OECD crude oil consumption	Million barrels/day	EIA, US Energy Information Administration	Crude oil consumption demand factors	Monthly
Inventory	US crude oil commercial inventory	Million barrels	EIA, US Energy Information Administration	Crude oil inventory level factors	Monthly
Cspp	FOB spot price of thermal coal in Newcastle Port, Australia	USD/ton	Wind	Alternative energy factors	Monthly
GasSpotPrice	Henry Hub natural gas spot price	USD/Million British Thermal Units	EIA, US Energy Information Administration	Alternative energy factors	Monthly
Wtioi	WTI crude oil noncommercial arbitrage holdings	piece	Wind	Speculative factors	Monthly
High-frequency explanatory variables (all daily data)					
DI	Dollar index	Yuan	Wind-Yitian Finance	Transaction currency factor	Daily
PS	Brent crude oil spot price	USD/barrel	Wind	Off-site high-frequency factors	Daily
GP	Henry Hub Natural Gas Spot Price	USD/Million British Thermal Units	Wind	Off-site high-frequency factors	Daily
V	Brent crude oil futures trading volume	10,000 hands	Wind	High frequency factors in the field	Daily
OI	Brent crude oil futures open interest (short order)	10,000 hands	Wind	High frequency factors in the field	Daily

4.2. Data Processing and Basic Characteristic Analysis

4.2.1. Interpreted Variables

The rate of return of oil futures prices selects the logarithmic rate of return RF , that is, the logarithm of the current oil futures price minus the logarithm of the previous period's futures price. The calculation formula is

$$RF_t = \ln PF_t - \ln PF_{t-1} \quad (21)$$

where PF_t represents the Brent crude oil futures price on day t ; PF_{t-1} represents the Brent crude oil futures price on day $t-1$.

It can be seen from Figure 1 that the trend of Brent crude oil futures prices and spot prices are almost the same, showing good consistency. From January 2008 to July 2008, the Brent crude oil futures price and the spot price were in the rising range and reached the highest point in history. From August 2008 to March 2009, they were in a downward trend. By July 2014, it entered the turbulent rise range, and then gradually fell, and rebounded after the end of 2017.

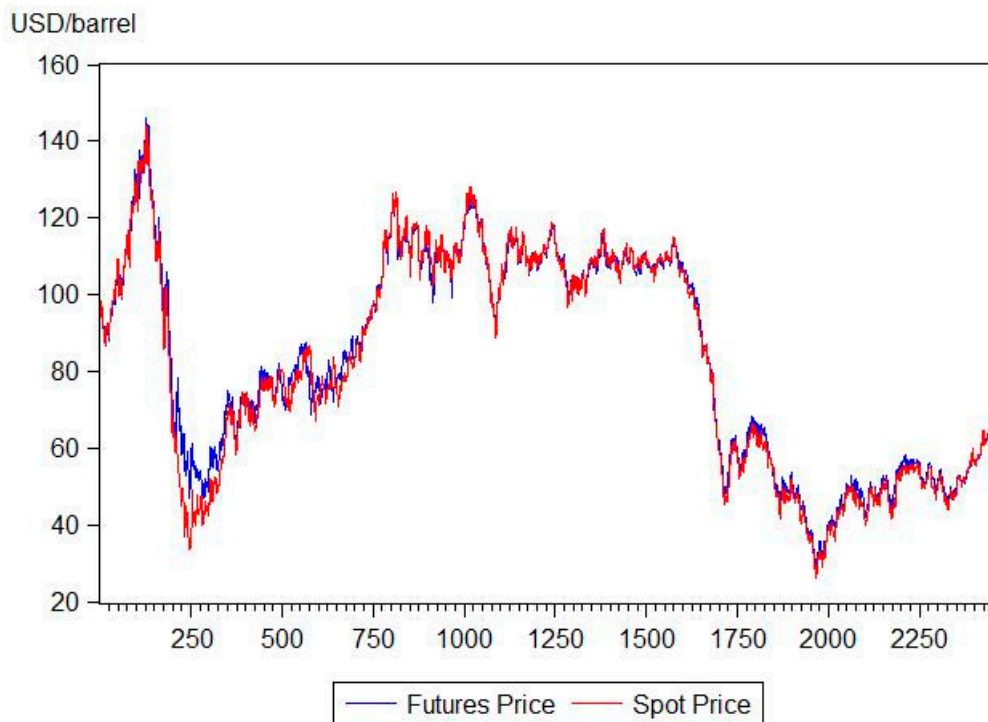


Figure 1. Trend chart of Brent crude oil futures price and spot price.

It can be seen from Figure 2 that the volatility of the Brent crude oil futures price yield shows a “clustering” phenomenon, and the volatility was small over a period of time, such as the data from 2013 to 2014, and it was very large in some time periods, such as year 2008 to 2009, and year 2015 to 2016.

Judging from the basic statistical characteristics of each series in Table 2:

① The average value of Brent crude oil futures returns is negative. In other words, for a long futures investor, the average investment income during the sample period is negative. In fact, if the transaction costs (such as transaction fees, the cost of funds occupied by margin, etc.) are considered, the futures yield is even lower. Obviously, the average value of Brent crude oil spot and futures price series is greater than zero.

② Nonnormality of sequence distribution. A common assumption in financial theory and empirical analysis is that the rate of return obeys a normal distribution. However, the actual financial rate of return often obeys a nonnormal distribution. The Brent crude oil futures prices, spot prices, and futures returns are no exception, the statistics in Table 2 illustrate this point. The skewness of the normal distribution should be equal to 0, that is positive skewness while the skewness value of the Brent crude oil futures return series is greater than 0, and that is negative skewness while the skewness of the Brent crude oil futures price and spot price series is less than 0. The kurtosis of spot prices and futures prices is less than 3, while the kurtosis of the Brent crude oil futures return sequence is greater than 3, indicating that it is a kurtosis distribution and fat-tailed distribution. The thick-tailed distribution means that there are many sample values that deviate from the sample average by a large margin, that is, the absolute value of the rate of return is larger than expected.

If the rate of return is normally distributed, then the Q-Q chart should be a straight line. In fact, the Q-Q diagram is an S-shaped curve, which shows that it has the characteristics of sharp peaks

and thick tails compared to the normal distribution. The test statistics of JB (Jarque–Bera) for all series are also much larger than the critical value (5.992). Based on the above results, it can be seen that the sequence has the characteristics of sharp peaks and thick tails, and does not obey the normal distribution.

③ From the perspective of coefficient of variation (CV), the volatility of Brent crude oil spot prices is slightly higher than that of futures prices. In addition, because the mean value of Brent crude oil futures returns is negative, its CV value is less than zero, but the absolute value of CV is relatively large, indicating that the fluctuations are relatively severe.

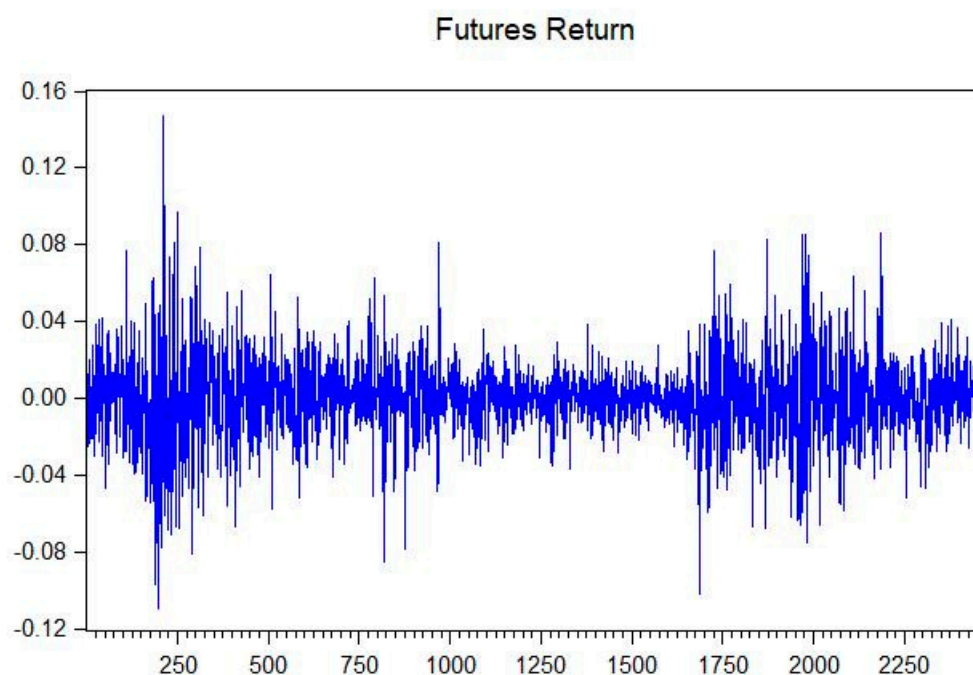


Figure 2. Price return trend chart of Brent oil futures.

Then, we will further discuss the basic statistical characteristics of Brent crude oil futures prices, spot prices, and futures yields. The results are shown in Table 2.

Table 2. Basic statistical characteristics of Brent oil futures price, spot price and futures return.

Index	Futures Price (FP)	Spot Price (PS)	Futures Return (RF)
Mean	82.71533	81.49046	−0.000147
Median	81.47500	79.82500	0
Maximum	146.0800	143.9500	0.147017
Minimum	27.79000	26.01000	−0.109455
Std. Dev.	27.05621	28.29182	0.020872
Skewness	−0.038434	−0.044386	0.177744
Kurtosis	1.630533	1.582079	6.520205
CV	0.3271	0.34718	−141.9864
Jarque–Bera	192.8384 [0]	206.8840 [0]	1359.791 [0]

Note: JB (Jarque–Bera) normal test statistics are: $JB = \frac{T}{6} \left[S^2 + \frac{1}{4}(K - 3)^2 \right]$, where S is skewness and K is kurtosis; [], the value in brackets is the probability value (p value); coefficient of variation (CV) = Std. Dev./Mean.

The stationarity of each high-frequency original sequence is then tested, and it is found that they are all first-order single integer sequences. As the time series put into the GARCH or ARMA model must be a stationary series, this article calculates the logarithmic difference of each series according to the processing method of the crude oil futures return sequence, that is, the return rate is modeled.

In the extraction of volatility, this article adopts the Schwert (1989) method, using explanatory variable logarithmic difference sequence to construct an AR(p) model, and the residuals obtained are then squared and used as the volatility of each variable. The wave dynamic potential of each variable is captured, and the AIC information criterion and the SC information criterion should be comprehensively considered as for the lag order p of the AR(p) model.

4.2.2. Low-Frequency Explanatory Variables

This article analyzes the influence of various factors on the price fluctuation of crude oil market from multiple aspects such as level effect, volatility effect, and the level and volatility effects combined. Therefore, it is necessary to process the data of each low-frequency explanatory variable. In addition, when constructing the GARCH-MIDAS model, the seasonal effect adjustment (X12 addition adjustment) was carried out before modeling because monthly data was selected.

In the extraction of volatility, this paper adopts the Schwert (1989) method, using the original data of explanatory variables (high-frequency variables) or seasonally adjusted data (low-frequency variables) to construct AR(p) models, and the residuals obtained are squared and used as the volatility of each variable to capture the wave dynamics of each variable. As for the lag order p of the AR(p) model, the AIC and the SC are comprehensively considered.

In addition, the statistics of the JB test of all sequences are statistically significant at the 10% significance level, indicating that all variables do not conform to the normal distribution. Therefore, when constructing the GARCH-MIDAS model, a standard GARCH-MIDAS model based on the normal distribution and the GARCH-MIDAS-skewed-t model with multiple explanatory factors (hereinafter referred to as GARCH-MIDAS-t) should be constructed. On the basis of the estimated results, a comparative analysis should be made.

In order to further determine the accurate form of the error distribution, this section first uses the GARCH (1,1)-n model to generate the standard residuals, and analyzes and judges the residuals and the normal distribution Q-Q diagram. Through a comprehensive analysis, the ideal distribution of the error distribution of the GARCH model is the T distribution. Therefore, the GARCH (1,1)-t model has improved accuracy and simplicity for fitting the volatility of Brent futures return series. Therefore, it is reasonable to choose the GARCH (1,1)-t model as the benchmark model for this empirical research.

5. Empirical Study on Fluctuation Effects of Crude Oil Futures Price from a Mixed-Frequency Data Sampling Perspective

5.1. Analysis of the Single-Factor Hierarchical Effect of Level and Fluctuation

5.1.1. Level Effect of Single Factor

(1) Level effect of low-frequency variables

Tables 3 and 4 give the estimation results of the mixed model for the level value of a single factor. θ_l reflects the level effect of each influencing factor on the long-term component of the high-frequency volatility, and ω_l reflects the optimal estimation weight in the single-factor mixed level value model. The coefficients θ_l examine the impact of the level of individual factors or the increase in the rate of return on the volatility of the crude oil market; that is, when θ_l is positive and statistically significant, the increase of the factor will promote the Brent crude oil futures market fluctuation. When θ_l is negative and statistically significant, the increase in this factor will reduce the volatility of the Brent crude oil futures market.

Table 3. Estimation results of low-frequency and single-factor level effect.

Variable	Inventory	OPEC	OECD	GasSpot Price	Cspp	Wtioi
μ	0.0001 [0.6055]	0.0001 [0.8525]	-2.295×10^{-5} [0.8732]	0.0001 [0.8595]	0.0001 [0.9288]	0.0001 [0.8473]
α	0.0507 *** [0]	0.0368 *** [0]	0.0456 [0]	0.0373 *** [0.0000]	0.0427 ** [0.0193]	0.0334 [0.6205]
β	0.9493 *** [0]	0.9604 *** [0]	0.9486 [0]	0.9598 *** [0]	0.9547 *** [0]	0.9665 *** [0]
m	-6.8425 *** [0]	-5.6824 *** [0]	-25.7218 [0]	-9.8229 *** [0]	-9.4151 * [0.0764]	-14.3512 *** [0]
θ_l	0.0027 *** [0.0013]	-0.0001 *** [0]	0.3463 *** [0]	0.0859 *** [0.0088]	0.5520 [0.6574]	0.0001 [0.4744]
ω_l	11.1236 *** [0]	1.0009 [0.8049]	12.2889 [0]	1.0788 [0.7594]	18.651 *** [0.0000]	2.9059 [0.8192]
g_0	0.0044 * [0.2016]	1.5761 [0.1654]	1.46719 [0]	1.1479 [0.1408]	0.8735 [0.8407]	162.3080 * [0.0631]
η	6.4348 [0]	6.4259 [0]	6.4214 [0]	6.4382 [0]	6.4193 [0.371422]	6.4290 [0]
λ	-0.0403 [0.2115]	-0.0402 [0.2152]	-0.0401 [0.2148]	-0.0399 [0.2181]	-0.0406 [0.2356]	-0.0400 [0.2143]
BIC	-7.044	-7.035	-7.0398	-7.044	-7.0498	-6.8641
LLH	-4973.9	-4967.6	-4970.9	-4974.0	-4968.9	-4951.1
H	0.0182	0.0195	0.0195	0.0188	0.0155	0.0212
MAE	0.0005	0.0005	0.0005	0.0005	0.0005	0.0004
RMSE	0.0008	0.0009	0.0009	0.0008	0.0008	0.0007

Note: ***, **, and * indicate statistically significant at the significance level of 1%, 5%, and 10%, respectively; p values are in brackets []; MAE means the mean value of absolute forecast error of realized volatility; RMSE means the square root of mean value of absolute forecast error of realized volatility.

Table 4. Estimation results of high-frequency and single-factor level effect.

Variable	V	OI	DI	GP	PS
μ	-0.0001 [0.8707]	0.0001 [0.9392]	0.0001 [0.9384]	0.0002 [0.9481]	0.0001 [0.9311]
α	0.0360 *** [0]	0.0421 *** [0]	0.0426 *** [0]	0.0421 *** [0.0008]	0.0445 *** [0.0001]
β	0.9609 *** [0]	0.9553 *** [0]	0.9542 *** [0]	0.9554 *** [0]	0.9449 *** [0]
m	-9.9266 *** [0]	-9.4256 *** [0]	-9.5483 *** [0]	-9.4099 *** [0]	-8.3016 *** [0]
θ_l	0.0002 *** [0.0005]	-0.9962 [0.1779]	-0.4062 [0.7135]	0.9857 [0.5246]	-0.0009 *** [0.0002]
ω_l	9.9950 [0.6718]	19.9031 *** [0]	11.5354 *** [0]	19.6919 *** [0]	8.168 *** [0]
g_0	2.2192 ** [0.0280]	0.9230 [0.5270]	1.0414 [0.1634]	0.8355 [0.6825]	0.9594 * [0.0984]
η	6.4295 [0]	6.4633 [0]	6.4737 [0]	6.4405 [0]	6.4656 [0]
λ	-0.0402 [0.2182]	-0.0397 [0.2211]	-0.0402 [0.2116]	-0.0397 [0.2190]	-0.0395 [0.2206]
BIC	-7.0422	-7.0439	-7.0431	-6.9623	-7.0214
LLH	-4972.63	-4973.83	-4973.23	-6508.49	-4972.35
H	0.0188	0.0188	0.0188	0.0195	0.0111
MAE	0.0005	0.0005	0.0005	0.0005	0.0005
RMSE	0.0008	0.0008	0.0008	0.0008	0.0008

Note: ***, **, and * indicate statistically significant at the significance level of 1%, 5%, and 10%, respectively; p values are in brackets [].

From the estimated results in Table 3, we can see the following:

① From the point of view of H value: based on the level effect model of US crude oil inventory (Inventory) and Australia's Newcastle port spot offshore price of thermal coal (Cspp), the H value is 0.0182 and 0.0155, respectively, which is closer to 1%. That is to say, the level effect model based on the above two variables is more stable. The H value of WTI crude oil noncommercial arbitrage holdings is the largest, and the estimation result is poor.

② From the perspective of BIC value, the BIC value of Inventory, GasSpotPrice, and Cspp models is the smallest. From the perspective of LLH (Log likelihood), the Inventory and GasSpotPrice models have the smallest LLH value, indicating that the Inventory and GasSpotPrice models have improved accuracy and simplicity.

From the results of coefficient θ_i estimation, we can see the following:

① The coefficients θ_i of US crude oil inventory (Inventory), OPEC crude oil production (OPEC), OECD crude oil consumption (OECD), and the Henry Hub natural gas spot price (GasSpotPrice) are statistically significant. The coefficients θ_i of the FOB spot price of thermal coal in Port Canius, Australia (Cspp) and WTI crude oil noncommercial arbitrage holdings (Wtioi) are not significant. The θ_i values of the US crude oil inventories, the OECD crude oil consumption, the natural gas spot prices, the spot price of thermal coal at the Port of Cannes, Australia, and the WTI crude oil noncommercial arbitrage positions are greater than zero, while the θ_i value of OPEC crude oil production is less than zero.

② The crude oil inventory coefficient is significantly greater than zero, and the US crude oil inventory level is positively correlated with the volatility of crude oil futures prices. Crude oil inventory is a part of the crude oil demand and the potential supply of crude oil next year. The higher its value, the higher the volatility of crude oil futures prices. Similarly, the OECD is the main demand representative of the crude oil market. The increase in OECD crude oil consumption will cause a significant increase in crude oil market volatility, which is consistent with economic theory.

③ As a kind of clean energy, natural gas has a wider scope of promotion. As one of the important substitutes for oil, the increase in natural gas prices is also one of the factors driving the fluctuation of crude oil prices. Similar to natural gas, coal is also an important substitute for oil. It has a positive impact on the fluctuation of crude oil prices. Due to coal's substitution role for crude oil requires conditions, its advantage of replacing crude oil will not exist when the price of coal gradually increases. Therefore, it can be seen from the estimation results that the impact is not sufficiently significant, the coal price level has which on the fluctuation of crude oil prices in the long run.

④ Crude oil noncommercial arbitrage holdings represent speculative factors in the crude oil futures market. From the estimation results, the coefficient θ_i of Wtioi is greater than zero, but the statistics are not significant, indicating that speculative factors are not important factors for international oil price fluctuations for the global crude oil market.

⑤ Unlike others, the coefficient θ_i of OPEC crude oil production is less than zero and statistically significant, indicating that the increase in OPEC crude oil production will weaken the volatility of Brent crude oil futures. OPEC crude oil production, as the main supplier of the oil market, has a significant impact on price fluctuations in the international crude oil market. When the OPEC crude oil production value is high, the global crude oil market is adequately supplied, and international oil prices which has the effect of stabilizing oil prices will be more stable.

Based on the above analysis, in terms of horizontal effects, crude oil production, consumption, inventory, and natural gas prices are the main factors that have a major impact on the fluctuation of Brent crude oil futures prices after the financial crisis in addition to coal spot prices and speculative factors.

(2) Level effect analysis of high-frequency variables

In the process of selecting explanatory variables, in addition to considering low-frequency variables, the influence of high-frequency variables is also considered in order to increase the sufficiency of the explanation.

This article combines indicator correlation and data availability and selects five high-frequency variables including the closing price of the US dollar index (USD), Brent crude oil futures trading volume (10,000 lots), Brent crude oil futures open interest (short orders) (10,000 lots), Henry Hub Natural Gas Spot Price (Dollars per Million Btu), and Europe Brent Spot Price FOB (Dollars per Barrel) as explanatory variables. Among them, the US dollar index, Brent crude oil spot prices, and natural gas prices represent high-frequency factors outside the market; trading volume and open interest represent high-frequency factors on the market.

From the estimated results in Table 4, we can see the following:

① From the point of view of H value: the H value of the Brent crude oil spot price (PS) model is the lowest, at 0.0111, and the H value of the natural gas price (GS) model is the highest, at 0.0195, while that of the Brent crude oil futures volume (V) and the dollar index (the H value of DI) and Brent crude oil futures' open interest (OI) is 0.0188. Thus, the spot price model is relatively more stable, while the natural gas price model is the worst.

② Based on MAE and RMSE criteria, the estimation effect of each model is basically the same.

③ The BIC and LLH values of the Brent crude oil futures trading volume (V), US dollar index (DI), and Brent crude oil futures open interest (OI) are better than other variables.

Judging from the θ_l coefficient estimation results in Table 4, the following can be concluded:

① The θ_l values of the Brent crude oil futures trading volume and Brent crude oil spot price is statistically significant. Other statistics such as the Brent crude oil futures open interest, the US dollar index, and the natural gas daily price are not statistically significant. The θ_l values of the spot price, the open interest, and the dollar index are negative, and the θ_l values of the other two high-frequency factors are positive.

② Trading volume is a high-frequency factor in the field affecting the market, reflecting market liquidity and trading activity. The trading volume coefficient is greater than zero and statistically significant, indicating that the trading volume on the market is an important factor affecting price fluctuations. However, when the daily trading volume is high, the volatility of Brent crude oil futures prices becomes stronger, and vice versa. Unlike trading volume, Brent crude oil futures open interest has a nonsignificant negative effect on the volatility of Brent crude oil futures. This is consistent with the results of the same frequency analysis. Different from the securities market, open interest is a statistical quantity unique to the futures market, reflecting the flow of funds on and off the market. An increase in open interest indicates that funds have entered the market, and a decrease in open interest indicates that funds have flowed out of the market. At the same time, there is a complicated correspondence between trading volume and open interest. The higher the open interest, the larger the market capital, and the weaker the relative market volatility.

③ The θ_l value of natural gas price is greater than zero, but it is not statistically significant. The analysis result of the daily price level effect of natural gas is similar to the analysis result of the monthly low-frequency data of natural gas, and both are nonsignificant positive relationships.

④ The θ_l value of the US dollar exchange rate is less than zero. The US dollar exchange rate represents the price of the US dollar. As a unit of calculation of crude oil futures, it has a certain impact on the price of crude oil. From the estimation results, when the exchange rate of the US dollar decreases, the volatility of oil prices increases as the marked price is only the nominal price of the crude oil, while the actual value of crude oil remains unchanged. When the value of the US dollar decreases, the price of crude oil will increase. When the price of the US dollar increases, the price unit of crude oil will decrease, so it has the effect of stabilizing oil prices.

⑤ The Brent futures price is representative of the expected Brent spot price, and the spot price is the basis of the futures price. Based on the estimation results, the Brent crude oil spot price income has a significant negative effect on the futures price fluctuation; that is, the spot price is negatively correlated to the volatility of the futures market, especially when the spot price is high, the futures price will also be high, and the relative volatility will be weaker.

Therefore, from the analysis results of high-frequency factors, the trading volume in the high-frequency factors on the market and the spot price in the high-frequency factors outside the market have a significant explanatory effect on the price volatility of the Brent crude oil futures market.

In addition, based on the estimation results in Tables 3 and 4, the Brent crude oil futures market has a strong GARCH effect. Various GARCH-MIDAS models reflect the adaptability of the mixed level model in describing the fluctuations of the Brent crude oil futures market.

5.1.2. Single Factor Volatility Effect

Tables 5 and 6 are the estimation results of the volatility effect of low-frequency and high-frequency variables, respectively. The coefficients θ_v in each equation examine the volatility of a single factor or the impact of increased volatility on the volatility of the crude oil futures market; that is, when θ_v is positive and statistically significant, the increase in the volatility of this factor will promote the volatility of the crude oil futures market. When θ_v is negative and statistically significant, the increase in the volatility of this factor will reduce the volatility of the crude oil futures market.

Table 5. Estimation results of low-frequency and single-factor fluctuation effect.

Variable	Inventory	OPEC	OECD	GasSpot Price	Cspp	Wtioi
μ	−0.0001 [0.9812]	0.0001 [0.8709]	0.0001 [0.9131]	0.0002 [0.9246]	0.0001 [0.9400]	0.0001 [0.9187]
α	0.0391 *** [0]	0.0396 *** [0]	0.0382 *** [0]	0.0442 *** [0]	0.0480 *** [0]	0.0392 *** [0]
β	0.9548 *** [0]	0.9603 *** [0]	0.9615 *** [0]	0.9530 *** [0]	0.9519 *** [0]	0.9576 *** [0]
m	−9.8328 *** [0]	−5.1434 *** [0.0002]	−7.7499 *** [0]	−9.4440 *** [0]	−4.2320 *** [0.0060]	−9.9527 *** [0]
θ_v	0.0858 *** [0]	7.9951 *** [0]	6.7618 *** [0]	0.0421 [0.8623]	0.6696 [0.5779]	0.4715 *** [0.0005]
ω_l	1.0009 ** [0.0130]	9.9957 [0.4502]	9.5686 [0.4756]	20.0000 *** [0]	1.2839 [0.5163]	1.1609 [0.1731]
g_0	1.9882 ** [0.0261]	0.0203 [0.4947]	0.2993 [0.4274]	1.0295 [0.2990]	0.0026 [0.3271]	0.6517 * [0.0687]
η	6.4349 [0]	6.4301 [0]	6.4029 [0]	6.4466 [0.0003]	6.43101 [0]	6.4202 [0]
λ	−0.0403 [0.2130]	−0.0403 [0.2098]	−0.0402 [0.2126]	−0.03846 [0.2404]	−0.0430 [0.2048]	−0.0407 [0.2100]
BIC	−7.0422	−7.0439	−7.0431	−6.9623	−7.0252	−7.0438
LLH	−4972.63	−4973.83	−4973.23	−6508.49	−4903.42	−4973.76
H	0.0188	0.0188	0.0188	0.0175	0.0185	0.0195
MAE	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005
RMSE	0.0008	0.0008	0.0008	0.0008	0.0009	0.0008

Note: ***, ** indicate statistically significant at the significance level of 1%, 5%, and 10%, respectively; p values are in brackets [].

Table 6. Estimation results of high-frequency and single-factor fluctuation effect model.

Variable	V	Oi	DI	GP	PS
μ	0.0001 [0.9449]	0.0002 [0.2377]	0.0001 [0.8790]	0.0001 [0.8442]	0.0001 [0.9965]
α	0.0332 *** [0]	0.3999 *** [0]	0.0375 [0.3911]	0.0362 *** [0]	0.0384 *** [0]
β	0.9667 *** [0]	0.6000 *** [0]	0.9621 *** [0]	0.9605 *** [0]	0.9547 *** [0]
m	−6.5572 *** [0]	−1.0000 [0.9948]	−8.0806 [0.9469]	−9.5688 *** [0]	−9.2252 *** [0]
θ_v	0.0001 *** [0]	1.8822 [0.9845]	2.6597 [0.9825]	7.9867 *** [0]	0.0001 *** [0]
ω_l	7.8997 [0.4096]	6.6237 [0.9970]	5.7794 [0.9992]	4.3075 [0.5402]	9.3012 [0.4894]
g_0	0.0404 ** [0.0181]	1.9986 [0.9948]	0.4172 [0.9934]	1.8398 [0.2056]	7.9645 * [0.0664]
η	6.4305 [0]	6.4242 [0]	6.4221 [0]	6.4206 [0]	6.4223 [0]
λ	−0.0408 [0.2078]	−0.0403 [0.2123]	−0.0403 [0.2115]	−0.0401 [0.2141]	−0.0403 [0.2123]
BIC	−7.0428	−7.0423	−7.0418	−6.946	−7.0423
LLH	−4973.02	−4972.65	−4972.32	−6774.82	−4972.69
H	0.0173	0.0174	0.0175	0.0173	0.0188
MAE	0.0005	0.0005	0.0005	0.0005	0.0005
RMSE	0.0008	0.0008	0.0009	0.0008	0.0008

Note: ***, **, and * indicate statistically significant at the significance level of 1%, 5%, and 10%, respectively; p values are in brackets [].

(1) Analysis of volatility effects of low-frequency variables

The volatility effects of low-frequency variables are estimated as shown in Table 5.

From the estimated results in Table 5, we can see the following:

① The $\alpha + \beta$ values of each model are very close to 1, indicating that each model has a relatively strong ARCH effect and volatility persistence.

② In terms of BIC and LLH standards, except for the volatility effect model of the spot offshore price of thermal coal (Cspp) in the Port of Cannes, Australia, the comparison results of other models are relatively consistent. Relatively speaking, the estimated accuracy of the OPEC, OECD, and Wtioi models is slightly higher. Regarding the H value, the estimation accuracy of the thermal coal spot offshore price (Cspp) model of Canusl Port, Australia, is slightly higher, followed by the Inventory, OPEC, and OECD volatility effect models, while the H value of WTI crude oil noncommercial arbitrage holdings is the largest, and the estimation result is poor. There is almost no difference in the MAE and RMSE values of each model.

③ In addition, the volatility effect model of low-frequency factors is compared with the level effect model of low-frequency factors: Although the corresponding models have advantages and disadvantages from the criteria of BIC, LLH, H, MAE, and RMSE, the volatility model generally outperforms the level effect model. Based on the average value of each indicator of the five models, the BIC, LLH, H, MAE, and RMSE values of the horizontal effect model are −7.0128, −4967.73, 0.0188, 0.0005, and 0.0008, and the five types of indicator values of the volatility effect model are −7.0268, −5217.56, 0.0187, 0.0005, and 0.0008, so overall the volatility effect model slightly outperforms the horizontal effect model.

④ The θ_v coefficient estimation values of all low-frequency factors are greater than zero, indicating that the volatility of the six low-frequency factors is positively correlated with the volatility of Brent crude oil futures returns, but the θ_v values of the coal spot price and natural gas spot price model are not statistically significant.

⑤ The increase in the volatility of crude oil inventories, crude oil consumption, crude oil production, substitute prices, and speculative factors will increase the volatility of the Brent crude oil futures market. Relatively speaking, the impact of substitutes is not sufficiently obvious.

(2) Analysis of volatility effects of high-frequency variables

The volatility effects of high-frequency variables are estimated as shown in Table 6.

From the estimated results in Table 6, we can see the following:

① Similar to the low-frequency variable volatility effect model, the $\alpha + \beta$ value of each model of the high-frequency variable volatility effect is very close to 1, indicating that each model has a relatively strong ARCH effect and relatively strong volatility persistence. In addition, the H value of the trading volume and natural gas price fluctuation effect model is 0.0173, the H value of the open interest and the dollar index fluctuation effect model is slightly higher, and the value of the spot price fluctuation effect model is the largest. Regarding LLH, the natural gas price model has the highest accuracy, but its BIC value is the largest, so from these two standards, the evaluation results are inconsistent, indicating that the stability of this model is poor. The MAE and RMSE values of each model have almost no difference.

② In addition, based on the comparison between the high-frequency factors volatility effect model and the high-frequency factors horizontal effect model, although the corresponding models have advantages and disadvantages based on the BIC, LLH, and H criteria, the volatility model generally outperforms the level effects model. From the average value of the indicators of the five models, the BIC, LLH, and H values of the horizontal effect model are -7.0226 , -5280.11 , and 0.0174 , and the five types of indicators of the volatility effect model are -7.0230 , -5333.10 , and 0.0177 , respectively. Therefore, in general, the volatility effect model of high-frequency factors slightly outperforms the horizontal effect model.

③ Judging from the θ_v parameter values of various variables, whether it is the high-frequency factors on the market or the high-frequency factors outside the market, all the θ_v coefficients are greater than zero, indicating that the increase in volatility will increase the volatility of Brent crude oil futures prices. Judging from the significance of the θ_v coefficients of each variable, the θ_v coefficients of the next three factors are statistically significant, except for the position volume and the dollar index. This shows that the volatility of high-frequency factors is also an important factor affecting the volatility of Brent crude oil futures.

5.2. Analysis of the Single-Factor Associative Effect of Level and Fluctuation

5.2.1. Analysis of Combination Effects of Low-Frequency Variables

The combined effect analysis of low-frequency variables is shown in Tables 7 and 8.

Table 7. Estimation results of associative model of level and fluctuation effect in low-frequency factors.

Variable	Inventory	OPEC	OECD	GasSpot Price	Cspp	Wtioi
μ	0 [0.9056]	0 [0.9423]	0 [0.8930]	0 [0.9325]	0 [0.9610]	0 [0.1294]
α	0.0338 *** [0.0007]	0.0419 *** [0]	0.0458 *** [0]	0.0433 *** [0]	0.0438 *** [0]	0.7231 *** [0]
β	0.9662 *** [0]	0.9530 *** [0]	0.9482 *** [0]	0.9525 *** [0]	0.9522 *** [0]	0.2769 *** [0]
m	-13.3721 *** [0]	-3.1813 ** [0.0479]	-27.9798 *** [0]	-9.0165 *** [0]	-6.8672 *** [0]	-0.2209 [1.0000]
θ_l	0.0013 *** [0]	-0.0002 *** [0.0005]	0.3948 *** [0]	0.1837 [0.2872]	0.0306 ** [0.0180]	0.1942 [0.9614]
θ_v	0.0074 ** [0.0422]	0.0046 ** [0.0487]	0.0851 [0.5522]	0.0617 [0.8209]	0.0019 [0.9505]	0.0281 [0.9845]

Table 7. Cont.

Variable	Inventory	OPEC	OECD	GasSpot Price	Cspp	Wtioi
ω_l	4.5635 [1.0000]	1.0001 [0.3431]	12.8704 *** [0]	14.8551 *** [0]	1.3416 [0.1384]	15.6523 [0.8662]
ω_v	3.5401 [1.0000]	1.0564 *** [0.0019]	11.8110 *** [0]	17.1247 *** [0]	8.7272 *** [0.0000]	16.1098 [0.9835]
g_0	11.9098 * [0.0510]	1.9675 [0.1766]	1.5100 [0.1120]	1.7317 [0.3262]	1.7322 [0.2934]	7.8332 [0.9614]
η	6.4367 [0]	6.4931 [0]	6.4361 [0]	6.4588 [0]	6.4256 [0]	6.4345 [0.1195]
λ	−0.0404 [0.2105]	−0.0404 [0.2114]	−0.0401 [0.2138]	−0.0384 [0.2340]	−0.0406 [0.2091]	−0.0403 [0.2455]
BIC	−7.0387	−7.032	−7.0304	−6.9697	−7.0418	−6.9682
LLH	−4977.43	−4972.67	−4971.56	−6094.00	−4979.60	−5482.39
H	0.0182	0.0185	0.0182	0.0181	0.0168	0.0207
MAE	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005
RMSE	0.0008	0.0009	0.0009	0.0008	0.0008	0.0009

Note: ***, **, and * indicate statistically significant at the significance level of 1%, 5%, and 10%, respectively; p values are in brackets [].

Table 8. Comparison of average estimation results between associative effect model and the single-factor level and fluctuation effect model in the low-frequency factors.

Index Mean	Single Level Effect Model	Single Volatility Model	Combined Effect Model of Level and Volatility
BIC	−7.0128	−7.0268	−7.0135
LLH	−4967.73	−5217.56	−5246.28
H	0.0188	0.0187	0.0184
MAE	0.0005	0.0005	0.0005
RMSE	0.0008	0.0008	0.0008

According to the estimated results in Table 8, the following conclusions can be drawn:

① The average BIC value of the low-frequency factor combination effect model is equal to −7.0135, which is slightly larger than the average value of the volatility effect model, but less than the average value of the horizontal effect model. However, the average LLC value of the low-frequency factor combination effect model is −5246.28, which is significantly smaller than the horizontal effect model and the volatility effect model. In addition, the average H value of the low-frequency factor combination effect model is also less than the average value of the single-level horizontal effect model and the volatility effect model. The MAE and RMSE values are almost the same. Therefore, as a whole, the combined effect model outperforms the single-level horizontal effect model and the single-level volatility effect model.

② Combining the stratification research results of the level effect and the volatility effect of each low-frequency factor, as well as the combination research results of the level effect and the volatility effect: on the one hand, the direction of influence of each factor on the volatility of Brent crude oil futures has not changed; on the other hand, there may be a correlation between the level effect of each factor and the volatility effect, causing other factors, except for Inventory, OPEC, and OECD, to change the significance of the level effect or the parameter of the volatility effect.

③ In regard to substitutes, based on the results of the combined effect, natural gas prices have no significant impact on the volatility of Brent crude oil futures, while another type of substitute, coal spot prices, has a positive effect on the volatility of Brent crude oil futures. The volatility of coal spot prices has no significant impact on the volatility of crude oil futures returns. Overall, the horizontal effect of natural gas prices is more significant when studied separately, and the horizontal effect of coal prices is

more significant when studied in combination. As two important substitutes for oil, there is a strong correlation between these substitutes' horizontal price and the volatility of crude oil prices.

④ From the perspective of speculation: Generally speaking, speculative factors are one of the influencing factors of financial market price fluctuations. According to the stratification and combination research, the stronger the market speculation, the stronger the volatility of crude oil futures. When the number of speculative factors increases, the speculative power in the market is increasing, and the increase in speculative power will cause frequent price changes to a certain extent, which intensifies market volatility in the crude oil market. However, from the perspective of crude oil, as a global bulk product, speculative factors are not the main factor in market fluctuations.

⑤ From the supply and demand level: From the perspective of the level effect coefficients θ_l and volatility effect coefficients θ_v of Inventory, OPEC, and OECD, it is mainly supply and demand that can affect the fluctuation of crude oil prices. Current consumption and current inventory both represent the current demand for crude oil, and OPEC production represents the current crude oil output, that is, the current supply. Regarding the value and significance of the estimated coefficient, whether it is a single analysis or a combined effect study, supply and demand are both significant factors, which is consistent with economic theory. Generally speaking, when the supply increases, the price of crude oil tends to fall. When the supply is high, the crude oil market has sufficient supply and the market is stable, so the volatility of crude oil prices is low. The increase in demand for crude oil will increase the price of crude oil. When the demand for crude oil is high, its changes will increase the volatility of crude oil prices.

Based on the results of the hierarchical and combination analysis of low-frequency factors, the most important factor affecting the volatility of international crude oil futures returns is the supply and demand of crude oil. In the long run, other factors such as alternative energy sources, speculative factors, inventory changes and other long-term factors will eventually be reflected in the supply or demand level of the crude oil market; in the short term, on-market and off-market factors will affect short-term price fluctuations of the crude oil futures, but in fact they still affect the short-term supply and demand of crude oil. Therefore, supply and demand are the most important factors in the fluctuation of international crude oil prices.

5.2.2. Analysis of the Combined Effect of High-Frequency Variables

Analysis of high-frequency variables is shown in Tables 9 and 10. The estimated results according to Table 10 are discussed below:

① The mean values of H, BIC, and LLC of the combined effect model of high-frequency factors are 0.0182, -7.0298, and -344.15, respectively. It can be seen that, though the H value of the high-frequency factor combination effect model is higher than that of the single-level horizontal effect model and the volatility effect model, the BIC and LLC indicators outperform the single-level model. In addition, the MAE and RMSE values are basically equal. Therefore, as a whole, the combined effect model significantly outperforms the horizontal effect model and the volatility effect model.

② Combining the stratification research results of the level effect and the volatility effect of each low-frequency factor, as well as the combination research results of the level effect and the volatility effect: on the one hand, except for OI, the direction of the horizontal effect and the volatility effect of each high-frequency factor has not changed; on the other hand, from the perspective of significance, due to the possible correlation between the level effect and the volatility effect of each factor, the other three types of high-frequency factors, except for V and OI, have changed in the significance of the level effect or volatility effect.

③ From the perspective of high-frequency factors in the market: The high-frequency factor that has a significant and stable impact on the volatility of Brent crude oil futures is trading volume. The impact of open interest is not significant, and the result of the horizontal effect is unstable. However, the position level effect is positive during portfolio analysis, which is more in line with theory and practice. In regards to the comprehensive trading volume and open interest, the level value and

volatility of high-frequency factors in the market has a positive effect on the volatility of Brent crude oil futures.

Table 9. Estimation results of associative model of level and fluctuation effect in high-frequency factors.

Variable	V	OI	DI	GP	PS
μ	0.0001 [0.9645]	0 [0.9494]	0 [0.9341]	0.0024 ** [0.0362]	0 [0.9323]
α	0.9056 *** [0]	0.0421 [0.6607]	0.0425 *** [0]	0.8181 *** [0.0006]	0.0412 *** [0]
β	0.0940 [0.4174]	0.9554 *** [0]	0.9549 *** [0]	0.1375 [0.5845]	0.9563 *** [0]
m	0.2121 [0.9995]	−9.4252 [0.6304]	−9.4256 *** [0]	−0.6886 *** [0]	−9.4259 *** [0]
θ_l	0.0212 *** [0]	0.8479 [0.9649]	0.2187 [0.6479]	0.0033 *** [0]	−0.2594 [0.7957]
θ_v	2.5393 *** [0]	0.9963 [0.9512]	0.9693 ** [0.0332]	0.0239 *** [0]	0.9918 ** [0.0137]
ω_l	1.0989 * [0.0591]	17.8177 [0.1411]	10.5885 *** [0]	7.2627 *** [0]	10.6961 *** [0.0000]
ω_v	9.9031 [0.5319]	19.9077 [0.5599]	19.1271 *** [0]	7.8873 *** [0]	19.6052 *** [0]
g_0	3.3383 [0.7190]	0.9212 [0.9578]	0.8993 [0.1889]	5.8090 [0.7427]	0.9011 * [0.0897]
η	6.4362 [0]	6.4633 [0.1868]	6.4302 [0]	6.4376 [0]	6.4605 [0]
λ	−0.0406 [0.2091]	−0.0397 [0.3299]	−0.0402 [0.2148]	−0.0395 [0.2207]	−0.0394 [0.2214]
BIC	−7.0446	−7.0331	−6.9999	−7.0323	−7.0390
LLH	−4976.48	−4973.49	−5020.22	−6772.94	−4977.61
H	0.0172	0.0175	0.0199	0.0182	0.0182
MAE	0.0005	0.0005	0.0005	0.0005	0.0005
RMSE	0.0009	0.0008	0.0008	0.0009	0.0008

Note: ***, **, and * indicate statistically significant at the significance level of 1%, 5%, and 10%, respectively; p values are in brackets [].

Table 10. Comparison of the mean value between associative effect model and the single-factor level and fluctuation effect model in the high-frequency factors.

Index Mean	Single Level Effect Model	Single Volatility Model	Combined Effect Model of Level and Volatility
BIC	−7.0226	−7.0230	−7.0298
LLH	−5280.11	−5333.10	−5344.15
H	0.0174	0.0177	0.0182
MAE	0.0005	0.0005	0.0005
RMSE	0.0008	0.0008	0.0008

④ The signs of the parameters of the combined effects of the high-frequency off-field factors are consistent with the results of the hierarchical analysis, but the significance of the parameters is different. The DI volatility effect changed from insignificant to a 5% significance level, the GP level effect changed from insignificant to a 1% significance level, and the PS level effect changed from significant to insignificant. In contrast, the DI level effect and the volatility effect of GP and PS are stabler. From the estimated results of the combined effect, the level value of GP is positively correlated with the volatility of crude oil futures. As one of the substitutes for crude oil, the higher the price of natural gas, the greater the market demand for natural gas or the decrease in supply, which will cause a change in the price of crude oil. The volatility of natural gas prices (up or down) is positively correlated with the volatility of crude oil futures prices. After considering the level effect of DI, the volatility effect

of DI is significant in the portfolio analysis, indicating that the correlation between the fluctuation of the US dollar index and the fluctuation of crude oil is enhanced after combining the level of the US dollar index. Changes in the exchange rate of the US dollar will cause fluctuations in the nominal price of crude oil, thereby increasing the volatility of crude oil futures. Whether it is a layered analysis or a combination effect analysis, the volatility effect of spot prices is always significantly positive, reflecting the close connection and linkage between the Brent crude oil futures market and the spot market.

Based on the results of the hierarchical analysis and combination analysis of high-frequency factors, both high-frequency on-market and off-market factors are important factors in the volatility of international crude oil futures returns, especially the on-market trading volume of the high-frequency factor, the off-market natural gas price, and the crude oil spot price of the high-frequency factor.

5.3. Analysis of the Multi-factors Associative Effect

According to the results of the aforementioned single-layer and combination analysis, there is a big difference between the results of single-factor single-effect analysis and combination effect analysis. This result is related to the type and number of factors included in the model. The modeling in the previous part is based on similar factors (a low-frequency or high-frequency factor), without considering the influence of other factors. In order to comprehensively consider the influence of the combination of different types of factors on the price volatility of Brent crude oil futures, this article attempts to conduct a multifactor combination analysis. However, if multiple combinations of various factors are considered, the number of combinations will be very large. Therefore, a combination of low-frequency and high-frequency factors is considered in a model.

According to the previous analysis, low-frequency factors such as the US crude oil inventory (Inventory), OPEC crude oil production (OPEC), and OECD crude oil consumption (OECD) and trading volume of high-frequency factors (V) have significant impacts on the volatility of Brent crude oil futures prices, and their conclusions are stable. Therefore, constructing a combined effect model from certain low-frequency factors (Inventory, OPEC, OECD) and high-frequency factors (V). The estimation results of the combined effect model of high-frequency and low-frequency factors are shown in Tables 11 and 12.

Firstly, in order to intuitively judge the estimation ability of the above-mentioned comprehensive low-frequency and high-frequency mixed effect model and the corresponding cofrequency effect model, the graphs of the realized volatility RVF estimated by the corresponding model and the realized volatility RV calculated using the original data were compared (graphics omitted). It was found that the RV and RVF of different models are basically consistent, indicating that the prediction effects of the models are acceptable.

Secondly, the MAE and RMSE values of each model are basically the same, indicating that the estimated RV and RVF errors of different models are basically the same, and the fluctuation trends are basically the same.

The out-of-sample prediction method was used to test the prediction accuracy of the GARCH-MIDAS model. The sample interval of the high-frequency variables in this article is from 1 January 2008 to 31 December 2017, and the total sample size is 2459 (excluding nontrading days). The sample includes 2376 data for modeling between 1 January 2008 and 31 August 2017, and 83 data for forecasting between 1 September 2017, and 31 December 2017. The developed GARCH-MIDAS model was used to calculate the MAE and RMSE indicators inside and outside the sample according to the results of parameter estimation to examine the prediction effect of the model. MAE and RMSE were used to compare the error between the estimated realized volatility and the actual realized volatility; the smaller the value, the more accurate the model's estimate of the volatility. Table 13 shows the estimated results of the predicted parameters of each model.

Table 11. Estimation results of associative effect model of high-frequency factors and low-frequency factors.

Variable	InventoryL+VL+VV	InventoryL+InventoryV+ VL+VV	OPECL+VL+VV	OPECL+InventoryV++ VL+VV	OECDL+VL+VV	OECDL+OECDV+ VL+VV
μ	0.0001 [0.9156]	−0.0003 [0.2803]	−0.0016 *** [0]	0.0010 *** [0]	0.0002 [0.3474]	0.0005 * [0.0666]
α	0.0417 *** [0]	0.5932 *** [0]	0.8800 *** [0]	0.9999 *** [0]	0.9860 *** [0]	0.9900 *** [0]
β	0.9582 *** [0]	0.4067 *** [0]	0.1100 *** [0]	0.0001 [0.6153]	0.0140 [0.3583]	0.0001 [1.0000]
m	14.0424 [0.1203]	0.7543 [0.3725]	−0.7494 [0.1586]	−0.5405 [0.2541]	−0.4019 [0.1451]	18.0479 [0.2912]
θ_{i1}	0.0001 * [0.0788]	0.0585 *** [0.0027]	−2.1498 * [0.0718]	−0.9357 * [0.0737]	0.0057 ** [0.0383]	0.0032 ** [0.0192]
θ_{i2}		1.1645 * [0.0576]		3.4459 ** [0.031]		0.7361 * [0.0890]
θ_{v1}	0.0001 ** [0.0482]	1.4133 ** [0.0362]	3.3657 * [0.0924]	3.7302 * [0.0759]	1.0079 ** [0.0379]	6.5945 ** [0.0481]
θ_{v2}	0.0093** [0.0218]	3.8929 ** [0.0161]	3.3337** [0.0287]	0.4019 ** [0.0392]	1.2308 ** [0.0254]	6.5698 ** [0.0240]
ω_{i1}	3.7131 * [0.0708]	2.0958 * [0.0672]	7.2931 [0.8700]	7.3121 [0.1011]	6.7925 * [0.0819]	2.5936 * [0.0755]
ω_{i2}		10.9361 [0.2057]		7.4393 * [0.0568]		1.6733 * [0.0293]
ω_{v1}	1.5872 * [0.0727]	13.2489 * [0.0841]	17.2791 * [0.0899]	9.6018 * [0.0706]	15.0131 * [0.0614]	3.6233 * [0.0418]
ω_{v2}	1.0001 [0.8533]	8.1073 [0.6763]	14.0755 [0.3989]	1.2509 [0.4989]	4.3083 [0.1974]	3.6233 [0.1359]
g_0	0.0001 [0.9190]	7.3380 [0.8702]	3.0136 [0.8680]	1.0888 [0.9738]	5.9636 [0.9396]	77.1754 [1.0000]
η	6.4349 [0]	6.4445 [0.1205]	6.4260 [0]	6.4363 [0]	6.4224 [0]	6.4371 [0]
λ	−0.0404 [0.2105]	−0.0403 [0.2456]	−0.0415 [0.2168]	−0.0407 [0.2092]	−0.0411 [0.2149]	−0.0411 [0.2188]
BIC	−7.0537	−7.0465	−7.0537	−7.0326	−7.0337	−7.1299
LLH	−4983.881	−4973.03	−4983.87	−4973.12	−4993.8743	−6074.294
H	0.0171	0.0170	0.0168	0.0169	0.0172	0.0176
MAE	0.0005	0.0005	0.0005	0.0005	0.0005	0.0004
RMSE	0.0008	0.0008	0.0008	0.0008	0.0008	0.0008

Note: ***, **, and * indicate statistically significant at the significance level of 1%, 5%, and 10%, respectively; p values are in brackets [].

Table 12. Comparison between the associative effect model of high-frequency and low-frequency factors and single-factor effect model.

Model	BIC	LLH	H	MAE	RMSE
InventoryL	−7.044	−4973.91	0.0182	0.0005	0.0008
OPECL	−7.035	−4967.55	0.0195	0.0005	0.0009
OECDL	−7.0398	−4970.94	0.0195	0.0005	0.0009
InventoryL+InventoryV	−7.0387	−4977.43	0.0182	0.0005	0.0008
OPECL+OPECV	−7.032	−4972.67	0.0185	0.0005	0.0009
OecdL+OECDV	−7.0304	−4971.56	0.0182	0.0005	0.0009
VL+VV	−7.0446	−4976.48	0.0172	0.0005	0.0009
InventoryL+VL+VV	−7.0537	−4983.881	0.0171	0.0005	0.0008
InventoryL+InventoryV++VL+VV	−7.0465	−4973.03	0.0170	0.0005	0.0008
OPECL+VL+VV	−7.0537	−4983.87	0.0168	0.0005	0.0008
OPECL+InventoryV++VL+VV	−7.0326	−4973.12	0.0169	0.0005	0.0008
OECDL+VL+VV	−7.0337	−4993.874	0.0172	0.0005	0.0008
OECDL+OECDV++VL+VV	−7.1299	−6074.294	0.0176	0.0004	0.0008

Table 13. Predicted parameter estimation results for each model.

Model	MAEin	RMSEin	MAEout	RMSEout
InventoryL	0.0005	0.0009	0.0003	0.0003
OPECL	0.0005	0.0009	0.0002	0.0003
OecdL	0.0005	0.0009	0.0002	0.0002
InventoryL+InventoryV	0.0005	0.0009	0.0002	0.0003
OPECL+OPECV	0.0005	0.0009	0.0002	0.0002
OECDL+OECDV	0.0005	0.0008	0.0002	0.0002
VL+VV	0.0005	0.0009	0.0002	0.0003
InventoryL+VL+VV	0.0005	0.0009	0.0002	0.0002
InventoryL+InventoryV++VL+VV	0.0005	0.0009	0.0002	0.0002
OPECL+VL+VV	0.0005	0.0009	0.0001	0.0002
OPECL+InventoryV++VL+VV	0.0005	0.0008	0.0002	0.0002
OECDL+VL+VV	0.0005	0.0005	0.0003	0.0003
OECDL+OECDV++VL+VV	0.0005	0.0005	0.0001	0.0001

Note: MAEin: MAE in Sample; RMSEin: RMSE in Sample; MAEout: MAE out Sample; RMSEout: RMSE out Sample.

According to Table 13, the MAEout value of various models is lower than MSEin, and the RMSEout value is lower than RMSEin, indicating that the parameters estimated by the fitting sample have a good estimation effect on the data outside the fitting sample, and the model does not have the problem of overfitting. Parameter estimation has a strong ability to predict the volatility of financial time series. In addition, like the previous analysis, the MAEin and RMSEin parameter estimation results of the various models in the sample are not much different, and the mixed-factor model slightly outperforms the same frequency effect model. From the out-of-sample MAEout and RMSEout parameters, the mixed factor model also outperforms the same frequency effect model.

Finally, according to the quantitative indicators of each model in Table 13, comparing the high and low frequency model with the model before combining, it was found that the combination of LLC, BIC and H criteria in the hybrid low-frequency and high-frequency combination model considers more different types of factors that affect crude oil price fluctuations, and contains more effective information that affects the long-term components of high-frequency volatility and therefore reflects the advantages over the same-frequency factor combination model in terms of estimation accuracy. For example, the InventoryL+Inventory V+VL+VV model is composed of the Inventory level and volatility combined effect model and the volume combined effect model. Comparing this with the InventoryL+InventoryV model and the VL+VV model, it was found that its BIC value is −7.0465, lower than the other two types models of co-frequency combination effect, and its H value is 0.017,

which is also lower than the other two models of cofrequency combination effect. It can be seen that the combined effect model, which combines two different factors of low frequency and high frequency, outperforms the model that only considers the combined effect of the same frequency factors. The comparison results of other models are basically the same, so they are not repeated here.

Regarding the impact of various variables on the volatility of Brent crude oil futures, comparing Table 13 with the estimated results of the corresponding model, although the significance of the individual models has changed, the conclusion of the study on the impact of crude oil futures volatility is not changed, and the parameter symbols are consistent with the research results of the comparison model, which reflects the advantages of the high and low frequency mixed effect model in the stability of the estimation results.

5.4. Summary and Contrast of Fluctuation Effects Estimation Results of Mixed Frequency and the Same Frequency Data Models

This section will summarize and compare the volatility effects of crude oil futures prices analyzed by GARCH and GARCH-MIDAS models.

5.4.1. Summary and Comparison of Estimation Results of Crude Oil Futures Price Volatility Effects

As the GARCH model is the same frequency data modeling, and the GARCH-MIDAS model is the mixing data modeling, in order to compare these two models, the estimate results of the high-frequency data level and volatility combined effect model in the GARCH-MIDAS model (Table 9) and the independent effect model in the GARCH model (Table 14) were compared and analyzed (Table 15).

Table 14. Estimation results of the influence of various factors on the fluctuation of Brent crude oil futures.

Variable	LV	LOI	LDI	LGP	LPS
	(A)	(B)	(C)	(D)	(E)
ω	6.99×10^{-7} [0.2873]	0.000314 *** [0]	3.96×10^{-7}	1.48×10^{-6} ***	8.75×10^{-6} ***
α_1	0.043678 *** [0]	0.104152 *** [0.0004]	0.042164 *** [0]	0.043754 *** [0]	0.072994 *** [0]
β_1	0.953208 *** [0]	0.534416 *** [0]	0.943901 *** [0]	0.953893 *** [0]	0.567250 *** [0]
γ_1	3.96×10^{-5} *** [0]	−0.001474 *** [0.0001]	0.000873 ** [0.0214]	9.31×10^{-5} * [0.097]	−0.000973 *** [0.0055]
γ_2	1.11×10^{-5} [0.1854]	0.040357 *** [0]	0.196468 *** [0]	0.000104 * [0.0907]	0.359574 *** [0]
R^2	0.001900	0.002526	0.002051	0.001944	0.000947
Adjusted R^2	0.001494	0.002120	0.001645	0.001537	0.000540
Log likelihood	6332.2	5915.735	6337.712	6329.677	6503.577
AIC	−5.146624	−4.807759	−5.151108	−5.187671	−5.286068
SC	−5.130086	−4.791221	−5.134571	−5.171134	−5.269530
Q(36)	41.685 [0.237]	42.432 [0.213]	42.177 [0.221]	42.944 [0.198]	42.122 [0.223]
ARCH-LM(1)	0.999609 [0.3174]	0.197764 [0.6565]	1.115448 [0.2909]	1.355851 [0.2443]	0.073082 [0.7869]
T-DIST	7.026739 *** [0]	7.117727 *** [0]	7.094364 *** [0]	7.066860 *** [0]	7.094174 *** [0]
$\alpha_1 + \beta_1$	0.996886	0.638568	0.986065	0.997647	0.640244

Note: ***, **, and * indicate statistically significant at the significance level of 1%, 5%, and 10%, respectively; p values are in brackets [].

Table 15. Comparison of estimation results between GARCH model and GARCH-MIDAS model.

Variable	GARCH Model				GARCH-MIDAS Model			
	Level Effect		Volatility Effect		Level Effect		Volatility Effect	
	Correlation	Significance	Correlation	Significance	Correlation	Significance	Correlation	Significance
V	+	***	+		+	***	+	***
OI	−	***	+	***	+		+	
DI	+	**	+	***	+		+	**
GP	+	*	+	*	+	***	+	***
PS	−	***	+	***	−		+	**

Note: + and − indicate positive correlation and negative correlation, respectively; *, **, *** mean statistically significant at the significance level of 10%, 5%, and 1%, respectively.

Regarding the direction of action, there is only a difference between the GARCH and GARCH-MIDAS models in the direction of the horizontal effect of the position volume. The GARCH model estimation result is negative, while the GARCH-MIDAS estimation result is positive but not significant, but the estimation result of the position fluctuation effect model is consistent with that of the GARCH model, that is, negative (not significant). The level of other variables and the direction of action of the volatility effect are consistent.

From the results of the significant difference, the significance of the volume fluctuation effect, the level and volatility effect of open interest, the dollar index level effect, and the spot price level effect have undergone significant changes. There is no significant difference between the level effect and the volatility effect of other variables.

5.4.2. Comparison of Estimation Accuracy between GARCH and GARCH-MIDAS Models

Furthermore, in order to compare the difference in estimation accuracy between GARCH and GARCH-MIDAS-t models, the GARCH-MIDAS level and volatility effect combined model was compared with the GARCH model.

Firstly, according to the GARCH model estimation results of the combined effect of volume level and volatility, the MAE and RMSE values were calculated; then according to the estimated value of the corresponding GARCH-MIDAS-t model, divide the MAE (RMSE) based on the GARCH-MIDAS-t model by the MAE (RMSE) of the GARCH model of the same sample to calculate RMAE (relative ratio value of MAE) and RRMSE (relative ratio value of RMSE), which can be used as the inspection index. When the RMAE (RRMSE) value is less than 1, it means that the GARCH-MIDAS model has the effect of improving the accuracy of model prediction and estimation compared with the GARCH (1,1) model.

According to the calculation results (Table 16), the RMAE and RRMSE values of trading volume, open interest, dollar index, natural gas price, and crude oil spot price are all less than 0.05. It shows that the GARCH-MIDAS-t model, compared with the GARCH model, has a very obvious role in improving the accuracy of prediction and estimation, which reflects the advantages of the mixed frequency model compared to the same frequency model.

Table 16. RMAE and RRMSE values of GARCH model and GARCH-MIDAS model.

Variables	GARCH Model		GARCH-MIDAS Model		RMAE	RRMSE
	MAE	RMSE	MAE	RMSE		
V	0.015042	0.021219	0.0005	0.0009	0.032131	0.041224
OI	0.015040	0.021213	0.0005	0.0008	0.031542	0.03938
DI	0.015040	0.021217	0.0005	0.0008	0.031299	0.039288
GP	0.015041	0.021218	0.0005	0.0009	0.032129	0.041074
PS	0.015069	0.021229	0.0005	0.0008	0.030623	0.03846

6. Conclusions

Crude oil is the most important energy product in the world and one of the major bulk commodities in the international market. It has a very close relationship with the economic and social development of all countries in the world and has attracted the attention and participation of many capitals. Given the impact of two oil shocks in the 1970s, led by OPEC who controls the pricing power on the economies of the developed countries in the West, in order to reduce the international crude oil spot price fluctuations, the United States and Britain in 1983 and in 1988 launched the WTI crude oil futures trading and the north sea Brent crude oil futures trading, hoped to ensure their goals of crude oil price security, and to grasp the international crude oil pricing power again through two basic functions of the futures market price discovery and hedging. In this way, the price of crude oil futures is not only affected by traditional factors, but also closely related to financial factors, resulting in frequent price fluctuations and very complex patterns, which has become a widespread concern for the government, academia and industry. In this paper, a GARCH-MIDAS-Skewed-t model based on mixed data was constructed, and the effects and differences of each single factor, multiple factors, and their combinations on the volatility of Brent crude oil futures were examined.

(1) The crude oil futures series and the high-frequency explanatory variable series do not follow a normal distribution and have “spikes and thick tails” and the phenomenon of bias. Therefore, this paper combines the skewed-t distribution and the GARCH-MIDAS model to construct a new GARCH-MIDAS-Skewed-t model. Empirical results show that, compared with the GARCH-MIDAS model of the normal distribution, the GARCH-MIDAS-Skewed-t model can better describe the volatility of the Brent crude oil futures return series.

(2) Regarding single-factor level effects, among the low-frequency factors, in addition to coal spot prices and speculative factors, crude oil production, consumption, inventory, and natural gas prices are the main factors that have a major impact on the fluctuation of Brent crude oil futures prices after the financial crisis. Although coal is also one of the important substitutes for petroleum, the substitution effect of coal on crude oil requires certain conditions. When the price of coal gradually increases, its advantage of replacing crude oil will not exist. In addition, from a long-term perspective, speculation factors are not the main factor in the price fluctuation of crude oil futures. Regarding the level effect of high-frequency factors, the trading volume in the high-frequency factors on the market and the spot prices in the off-market high-frequency factors have a significant explanatory effect on the price volatility of the Brent crude oil futures market. As trading volume reflects market liquidity and trading activity, and spot prices are the basis of futures prices. therefore, from a short-term perspective, trading volume and spot prices have more important impacts on price fluctuation in the crude oil futures market. This is consistent with the analysis result of the existing theory.

(3) Regarding single-factor volatility effects, among the low-frequency factors, the volatility of all low-frequency factors is positively correlated with the volatility of Brent crude oil futures returns, but the effects of coal and natural gas spot price models are not statistically significant. The volatility of crude oil inventories, crude oil consumption, crude oil production, substitute prices, and speculative factors will increase the volatility of the Brent crude oil futures market. From a long-term perspective, relatively speaking, the impact of substitutes is not sufficiently obvious, indicating that the main factor affecting the volatility of Brent crude oil futures is its own demand and supply. Regarding the high-frequency data, whether it is on-market high-frequency factors or off-market high-frequency factors, the volatility of all high-frequency factors will increase the volatility of Brent crude oil futures. In addition to open interest and the US dollar index, the other three factors, namely crude oil spot price, substitute price, and trading volume, are all significantly positively correlated with the volatility of Brent crude oil futures, indicating that the volatility of Brent crude oil futures are susceptible to both on-market and off-market factors.

(4) Regarding the combined effect of single factor level and volatility, considering the low-frequency data, the most important factor influencing international crude oil futures return volatility is the supply and demand of crude oil, while other factors such as the change in alternative energy and

speculative factors, will ultimately be reflected in the supply or demand level of the crude oil market. Therefore, supply and demand are the most important factors in the fluctuation of international crude oil prices. This shows that the financial crisis has not changed the long-term core position of supply and demand. Supply and demand are the basic cause of crude oil fluctuations. However, supply and demand are quietly undergoing structural changes: Firstly, the financial crisis has accelerated the trend of crude oil pricing power from being concentrated to decentralized. Before the financial crisis, the international pricing power of crude oil was basically concentrated in the hands of developed countries such as Europe and the United States. After the outbreak of the economic crisis in 2008, the economic strength and financial control of the developed countries was weakened. On the issue of international oil pricing, OPEC, Europe, America, Japan, China, and Russia all have a certain degree of discourse in international market pricing. The United States and the Organization of Petroleum Exporting Countries have both similarities and conflicts of interest. Secondly, the shale oil and gas revolution in the United States, Russia's huge oil reserves, and China's huge oil imports are all affecting international crude oil prices. It can be seen from this that, after the financial crisis, the tendency of international market pricing power to shift from concentration to decentralization is relatively obvious, and it is concluded that crude oil production and consumption have a major influence on the price fluctuation of Brent crude oil futures.

Regarding high-frequency data, high-frequency on-market and off-market factors are important factors in the volatility of international crude oil futures returns, especially the trading volume of the on-market high-frequency factors, and the natural gas price and crude oil spot price of the off-market high-frequency factors. There are two main theories on the relationship between price fluctuations and trading volume in financial market: one is the assumption of continuous information arrival; the other is the hypothesis of mixed distribution. The assumption of continuous information arrival holds that market information is spread out step by step, and traders get information step by step. Only when all traders get relevant information, the market can reach the final equilibrium. There is a positive correlation between price fluctuations and trading volume, and volume information is helpful for forecasting price fluctuations. The mixed distribution hypothesis holds that trading volume and price fluctuations both depend on the same variable which can be described by the speed of market information transmission. When market information changes, price and trading volume respond to the information at the same time, that is, the mixed distribution hypothesis holds all traders receive new information and respond at the same time. The new balance point does not need to pass through the intermediate balance point and is completed in an instant, and the price changes and the adjustment of trading volume are carried out simultaneously. According to this theory, the sequence-related information arrival process causes the ARCH behavior in the futures rate of return, namely the continuity of volatility, and the futures contract transaction volume can be used as a proxy indicator of information arrival, explaining the time-varying nature of the variance of the return rate. Investors can take advantage of the fact that there is a correlation between futures price fluctuations and trading volume, and using the joint information between crude oil futures price fluctuations and trading volume can improve the accuracy and effectiveness of analysis and inference; it also helps to understand the role of trading volume and open interest in futures price fluctuations, and strengthen risk management in the crude oil futures market.

Based on the above conclusions, the fundamental reason for crude oil price fluctuations is the increasing diversification of dominant forces, and the direct cause is that the game of pricing power has triggered the structural disorder of the crude oil supply and demand system. A historical investigation of the structural transformation process of the crude oil market found that: spatial isolation between scarcity and supply and demand, small demand elasticity and supply which is vulnerable to be monopolized, the contradiction between supply and demand caused by strong demand growth, the gap control of storage, supply, demand, and inventory in the transaction process, changes in the price of petrodollars, excessive speculations by traders, and the impact of emergencies all have triggered violent and frequent fluctuations in crude oil prices. Crude oil futures prices are the combination of natural

laws (laws of sustainable development), micro-market laws, financial market laws, and emergencies and policy regulations in different time and space. According to the scope of action, degree of influence, and frequency of occurrence, the main controlling factors that affect crude oil futures price fluctuations are mainly divided into three categories: First, the basic factors that affect the medium and long-term trend, mainly including crude oil supply, crude oil demand and crude oil transaction settlement currency, while basic factors have a direct impact on the international crude oil futures prices; second, structural factors that affect the short-term trend, mainly include crude oil inventories in the spot market and changes in the prices of substitutes (natural gas, coal), changes in positions (arbitrage and speculation, etc.) on the financial market (future market); third, incidental factors that trigger short-term fluctuations, mainly including international political turbulence, geopolitical events, wars, and sudden climate changes. In the market, the various factors and categories of crude oil prices are independent and interactive with each other, showing integrity, organic correlation, dynamic development, orderliness, and complexity, forming a “transaction–market–environment” system. In this system, supply, demand, and crude oil transaction settlement currency are the underlying factors of crude oil as a commodity transaction, interacting in the transaction subsystem of the system; crude oil spot inventory in the commodity market, alternatives to coal and natural gas, on-site factors such as hedging and speculation in the crude oil futures market complete the structural activities in the market subsystem; international politics and economic turmoil, conflicts and wars in major oil producing areas, and extreme abnormal climates constitute the external environmental subsystem in which information is fully exchanged, therefore triggered sharp fluctuations in crude oil futures prices.

(5) Regarding high-frequency and low-frequency combined effects, although the results of single-factor and multifactor models are different, there is no significant difference in general. The research results are basically the same, reflecting the advantages of the mixed effect model of high and low-frequency data in the stability of the estimation results. This shows that with the advancement of technology, the trend of information integration will become more and more obvious. Therefore, in the management of crude oil futures volatility risk, attention should be paid to the rich information contained in high-frequency data, especially in risk hedging, which can play more important role.

Author Contributions: C.W. and F.W. conceived the idea of this paper; X.W. and F.W. provided financial support; C.W. and F.W. performed the model; X.W. analyzed the data; C.W. finished the first draft; S.L. and J.S. assisted in modifying the paper and translating papers. All the authors read and confirmed the final manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 71871215 and grant number 71673270; Think Tank of Energy Mining Economy (2018 Project for Cultural Evolution and Creation of CUMT, China), grant number 2018WHCC01. And the APC was funded by 71871215 and 71673270.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Aroui, M.E.H.; Jouini, J.; Nguyen, D.K. On the impacts of oil price fluctuations on European equity markets: Volatility spillover and hedging effectiveness. *Energy Econ.* **2012**, *34*, 611–617. [\[CrossRef\]](#)
2. Ji, Q.; Fan, Y. How does oil price volatility affect non-energy commodity markets? *Appl. Energy* **2012**, *89*, 273–280. [\[CrossRef\]](#)
3. Elder, J.; Serletis, A. Oil price uncertainty. *J. Money Credit. Bank* **2010**, *42*, 1137–1159. [\[CrossRef\]](#)
4. Qiang, W.; Lin, A.; Zhao, C.; Liu, Z.; Liu, M.; Wang, X. The impact of international crude oil price fluctuation on the exchange rate of petroleum-importing countries: A summary of recent studies. *Nat. Hazards* **2018**, *95*, 227–239. [\[CrossRef\]](#)
5. Fong, W.M.; See, K.H. A Markov switching model of the conditional volatility of crude oil futures prices. *Energy Econ.* **2002**, *24*, 71–95. [\[CrossRef\]](#)
6. Yang, C.W.; Hwang, M.J.; Huang, B.N. An analysis of factors affecting price volatility of the us oil market. *Energy Econ.* **2002**, *24*, 107–119. [\[CrossRef\]](#)

7. Demirer, R.; Kutan, A.M. The behavior of crude oil spot and futures prices around OPEC and SPR announcements: An event study perspective. *Energy Econ.* **2010**, *32*, 1467–1476. [CrossRef]
8. Saif Ghouri, S. Assessment of the relationship between oil prices and US oil stocks. *Energy Policy* **2006**, *34*, 3327–3333. [CrossRef]
9. Hayat, A.; Narayan, P.K. The oil stock fluctuations in the United States. *Appl. Energy* **2010**, *87*, 178–184. [CrossRef]
10. John, E.P. Black Gold and Fool's Gold: Speculation in the Oil Futures Market. *Economía* **2010**, *10*, 81–116. [CrossRef]
11. Blair, B.F.; Rezek, J.P. The effects of Hurricane Katrina on price pass-through for Gulf Coast gasoline. *Econ. Lett.* **2008**, *98*, 229–234. [CrossRef]
12. Liu, J.; Ma, F.; Tang, Y.; Zhang, Y. Geopolitical risk and oil volatility: A new insight. *Energy Econ.* **2019**, *84*, 104548. [CrossRef]
13. Gil-Alana, L.A. A fractionally integrated model with a mean shift for the us and the uk real oil prices. *Econ. Model.* **2000**, *18*, 643–658. [CrossRef]
14. Heath, D. Macroeconomic Factors in Oil Futures Markets. *Manag. Sci.* **2019**, *65*, 4407–4421. [CrossRef]
15. Sadik, Z.A.; Date, P.M.; Mitra, G. Forecasting crude oil futures prices using global macroeconomic news sentiment. *Ima J. Manag. Math.* **2019**, *31*, 191–215. [CrossRef]
16. Bildirici, M.; Ersin, Ö. Forecasting volatility in oil prices with a class of nonlinear volatility models: Smooth transition RBF and MLP neural networks augmented GARCH approach. *Pet. Sci.* **2015**, *12*, 534–552. [CrossRef]
17. Chen, H.; Liu, L.; Li, X. The predictive content of CBOE crude oil volatility index. *Phys. A Stat. Mech. Its Appl.* **2018**, *492*, 837–850. [CrossRef]
18. Wei, Y.; Wang, Y.; Huang, D. Forecasting crude oil market volatility: Further evidence using GARCH-class models. *Energy Econ.* **2010**, *32*, 1477–1484. [CrossRef]
19. Herrera, A.M.; Hu, L.; Pastor, D. Forecasting crude oil price volatility. *Int. J. Forecast.* **2018**, *34*, 622–635. [CrossRef]
20. Pan, Z.; Wang, Y.; Wu, C.; Yin, L. Oil price volatility and macroeconomic fundamentals: A regime switching GARCH-MIDAS model. *J. Empir. Financ.* **2017**, *43*, 130–142. [CrossRef]
21. Gong, X.; Lin, B. Predicting the volatility of crude oil futures: The roles of leverage effects and structural changes. *Int. J. Financ. Econ.* **2020**, 1–31. [CrossRef]
22. Yun, W.C. Comparison of Price Predictive Ability between Futures Market and Expert System for WTI Crude Oil Price. *Environ. Resour. Econ. Rev.* **2005**, *14*, 201–220.
23. Yanagisawa, A. Usefulness of The Forward Curve in Forecasting Oil Prices. 2009. Available online: <https://econpapers.repec.org/paper/eabenergy/22969.htm> (accessed on 25 November 2020).
24. Baruník, J.; Malinská, B. Forecasting the term structure of crude oil futures prices with neural networks. *Appl. Energy* **2016**, *164*, 366–379. [CrossRef]
25. Cai, Y.; Xu, G.; Hongwei, Z. Volatility forecasting of crude oil futures: The role of investor sentiment and leverage effect. *Resour. Policy* **2019**, *61*, 548–563.
26. Sévi, B. Forecasting the volatility of crude oil futures using intraday data. *Eur. J. Oper. Res.* **2014**, *235*, 643–659. [CrossRef]
27. Sadorsky, P. Modeling and forecasting petroleum futures volatility. *Energy Econ.* **2006**, *28*, 467–488. [CrossRef]
28. Pindyck, R.S. Volatility in natural gas and oil markets. *Work. Pap.* **2003**, *30*, 312.
29. Cansado-Bravo, P.; Rodríguez-Monroy, C. Persistence of oil prices in gas import prices and the resilience of the oil-indexation mechanism. The case of spanish gas import prices. *Energies* **2018**, *11*, 3486. [CrossRef]
30. Villar, J.A.; Joutz, F.L. The Relationship between Crude Oil and Natural Gas Prices. *Energy Inf. Adm.* **2006**, *35*, 1–43. Available online: <https://www.mendeley.com/catalogue/bf4c95bf-5f4f-3581-bf12-a0c1801457cf/> (accessed on 25 November 2020).
31. Brigida, M. The switching relationship between natural gas and crude oil prices. *Energy Econ.* **2014**, *43*, 48–55. [CrossRef]
32. Caporin, M.; Fontini, F. The long-run oil–natural gas price relationship and the shale gasrevolution. *Energy Econ.* **2017**, *64*, 511–519. [CrossRef]
33. Ghysels, E.; Sinko, A.; Valkanov, R. MIDAS Regressions: Further Results and New Directions. *Econom. Rev.* **2007**, *26*, 53–90. [CrossRef]

34. Engle, R.F.; Ghysels, E.; Sohn, B. Stock market volatility and macroeconomic fundamentals. *Rev. Econ. Stat.* **2013**, *95*, 776–797. [\[CrossRef\]](#)
35. Girardin, E.; Joyeux, R. Macro fundamentals as a source of stock market volatility in China: A GARCH-MIDAS approach. *Econ. Model.* **2013**, *34*, 59–68. [\[CrossRef\]](#)
36. Conrad, C.; Loch, K.; Rittler, D. On the macroeconomic determinants of long-term volatilities and correlations in U.S. stock and crude oil markets. *J. Empir. Financ.* **2012**, *29*, 26–40. [\[CrossRef\]](#)
37. Pan, Z.; Wang, Q.; Wang, Y.; Yang, L. Forecasting u.s. real gdp using oil prices: A time-varying parameter midas model. *Energy Econ.* **2018**, *72*, 177–187. [\[CrossRef\]](#)
38. Feng, M.; Yin, L.; Zhang, Y.J. Harnessing jump component for crude oil volatility forecasting in the presence of extreme shocks. *J. Empir. Financ.* **2019**, *52*, 40–55.
39. Amendola, A.; Candila, V.; Scognamiglio, A. On the influence of US monetary policy on crude oil price volatility. *Empir. Econ.* **2017**, *52*, 1–24. [\[CrossRef\]](#)
40. Libo, Y.; Zhou, Y.M. What Drives Long-term Oil Market Volatility? Fundamentals versus Speculation. *Econ. Open Access J.* **2016**, *10*, 1–26.
41. Ma, Y.R.; Ji, Q.; Pan, J. Oil financialization and volatility forecast: Evidence from multidimensional predictors. *J. Forecast.* **2019**, *38*, 564–581.
42. Mei, D.; Ma, F.; Liao, Y. Geopolitical risk uncertainty and oil future volatility: Evidence from MIDAS models. *Energy Econ.* **2020**, *86*, 104624. [\[CrossRef\]](#)
43. Ma, R.; Zhou, C.; Cai, H.; Deng, C. The forecasting power of epu for crude oil return volatility. *Energy Rep.* **2019**, *5*, 866–873. [\[CrossRef\]](#)
44. Lei, L.; Shang, Y.; Chen, Y. Does the financial crisis change the economic risk perception of crude oil traders? A MIDAS quantile regression approach. *Financ. Res. Lett.* **2018**, *30*, 341–351. [\[CrossRef\]](#)
45. Baumeister, C.; Guerin, P.; Kilian, L. Do high-frequency financial data help forecast oil prices? The MIDAS touch at work. *Int. J. Forecast.* **2015**, *31*, 238–252. [\[CrossRef\]](#)
46. Zhang, Y.J.; Wang, J.L. Do high-frequency stock market data help forecast crude oil prices? Evidence from the MIDAS models. *Energy Econ.* **2019**, *78*, 192–201. [\[CrossRef\]](#)
47. HANSEN, B.E. Autoregressive Conditional Density Estimation. *Int. Econ. Rev.* **1994**, *35*, 705–730. [\[CrossRef\]](#)
48. Manickavasagam, J.; Visalakshmi, S.; Apergis, N. A novel hybrid approach to forecast crude oil futures using intraday data. *Technol. Soc Chang.* **2020**, *158*, 120–126. [\[CrossRef\]](#)
49. Perifanis, T.; Dagoumas, A. Living in an era when market fundamentals determine crude oil price. *Energy J.* **2019**, *40*, 317–335. [\[CrossRef\]](#)
50. Kim, S.; Heo, E.Y. Crude oil inventories: The two faces of Janus? *Empir. Econ.* **2020**, *59*, 1003–1018. [\[CrossRef\]](#)
51. Manzoor, D. Are crude oil, gas and coal prices cointegrated? *Iran. Econ. Rev.* **2011**, *16*, 29–51.
52. Moosa, I.A. School of Business. Arbitrage, hedging, speculation and the pricing of crude oil futures contracts. *Keio Econ. Stud.* **2000**, *37*, 53–61.
53. Samanta, S.K.; Zadeh, A.H.M. Co-movements of Oil, Gold, the U.S. Dollar, and Stocks. *Mod. Econ.* **2011**, *3*, 111–117. [\[CrossRef\]](#)
54. Sequeira, J.M.; Mcaleer, M. A market-augmented model for simex brent crude oil futures contracts. *Appl. Financ. Econ.* **2000**, *10*, 543–552. [\[CrossRef\]](#)
55. Silvennoinen, A.; Thorp, S. Financialization, crisis and commodity correlation dynamics. *J. Int. Financ. Mark. Inst. Money* **2013**, *24*, 42–65. [\[CrossRef\]](#)
56. Hatipoğlu, H.; Boran, F.E.; Avci, M.; Akay, D. Linguistic Summarization of Europe Brent Spot Price Time Series Along with the Interpretations from the Perspective of Turkey. *Int. J. Intell. Syst.* **2014**, *29*, 946–970. [\[CrossRef\]](#)

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).