Energy Nexus 20 (2025) 100560

Contents lists available at ScienceDirect

Energy Nexus

journal homepage: www.elsevier.com/locate/nexus

Post-COVID dynamics of the refined–crude oil price spread in the us: Evidence from long memory and fractional cointegration models

Manuel Monge [a,b,*], Carlos Poza [a]

[a] Universidad Francisco de Vitoria, Madrid, Spain
[b] Universidad Europea de Madrid, Madrid, Spain

ARTICLE INFO

JEL Codes:
Q41
Q43

Keywords:
Refined oil
Crude oil
Spread refined-crude petroleum
Natural gas price

ABSTRACT

This study investigates the widening spread between refined and crude oil prices in the United States during the post-COVID period. The objective of the paper is to analyze the dynamics of this spread, with particular emphasis on the persistence of shocks and the existence of long-run equilibrium relationships. The study is motivated by concerns about inflationary pressures stemming from energy price dynamics, particularly in the context of persistent supply bottlenecks and shifting demand patterns. We employ fractional integration and fractional cointegration methodologies—specifically ARFIMA and FCVAR models—to analyze the persistence of shocks and the long-run relationships among key variables. Monthly data from 2003 to 2023 are used, including WTI crude prices, gasoline prices, industrial natural gas prices, petroleum consumption, and refinery capacity utilization. The findings indicate that, while the spread exhibits non-mean-reverting behavior after COVID, natural gas prices do not significantly affect the spread in the long run. In contrast, supply-demand imbalances and refinery constraints appear to be the main drivers. These results offer new insights into the structural determinants of refined oil prices in the post-pandemic energy landscape.

1. Introduction

1.1. Research background

The increase in refined petroleum prices in the international economy has caused a severe impact on inflation rates, in the context of weak economic growth, which has provoked problems of stagnation. This energy sector crisis has been spreading around the world quickly, impacting advanced and developing countries. One example of this concerns monetary policy tightening in the US and the Eurozone in 2023.

The cost of crude oil is the largest component of the retail price of gasoline and diesel (>50 % indeed). This cost varies over time because of the following main factors [1]: 1) Non-OPEC supply, 2) OPEC production, 3) balance or inventories, 4) spot prices, 5) financial markets (future position, for example), 6) Non-OECD demand, and 7) OECD demand.

Refining costs vary seasonally and by region. Additionally, the refined oil produced depends on the type of crude oil used and the type of processing technology available at the refinery where it is produced. However, refiners rely on natural gas as a crucial source of heat to distill crude oil into products such as diesel and gasoline [2].

Distribution and retail dealer costs are likewise included in the retail price of gasoline and diesel. Most refined oil is shipped from refineries by pipeline to terminals near consuming areas, where it may be blended with other products (such as fuel ethanol) to meet local government and market specifications. Gasoline and diesel are delivered to individual gasoline stations by tanker trucks. Moreover, the hydrocarbon and VAT taxes (the classical Pigouvian taxes), and their variations, also impact on the final refined oil prices.

Recent literature has emphasized the broader systemic, regulatory, and technological contexts that shape post-COVID economic dynamics. For instance, Yang et al. [3] show that sustainable urban development initiatives can generate both economic and environmental spillovers, underlining the interconnectedness between policy, growth, and sustainability. Similarly, Yang, Zheng, and Yang [4] highlight that environmental regulations influence technological complexity in manufacturing exports, reinforcing the role of regulatory frameworks in shaping competitiveness under global uncertainty. At the same time, Yang, Zhu, and Yang [5] demonstrate that urban digital construction significantly enhances economic growth, illustrating how digital transformation acts as a driver of resilience and long-term development.

* Corresponding author at: Universidad Francisco de Vitoria, Faculty of Law, Business and Governance, 28223 Pozuelo de Alarcón, Madrid, Spain.
E-mail address: manuel.monge@ufv.es (M. Monge).

https://doi.org/10.1016/j.nexus.2025.100560
Received 9 May 2025; Received in revised form 14 September 2025; Accepted 15 October 2025
Available online 17 October 2025

These findings suggest that the refined–crude oil price spread should not only be understood as a sector-specific phenomenon but also as part of wider systemic transformations in sustainability, regulation, and digitalization in the post-pandemic era.

Within this research stream, the main objective of this paper is to analyze the widening of the refined–crude petroleum price spread in the United States during the post-COVID period. To achieve this, we examine whether shocks to the spread exhibit persistence and whether long-run equilibrium relationships exist among the key explanatory variables, namely crude oil prices, refining costs, distribution costs, taxes, natural gas prices, petroleum consumption, and refining capacity utilization, using fractional integration and fractional cointegration models.

Many authors have tried to construct explanatory models using these factors to understand and forecast oil price fluctuations. Choi and Hammoudeh [6] tested for the presence of long memory in daily oil and refined products prices' absolute return, squared return, and conditional volatility. Choi et al. [7] studied the long- and short-term relationships between refined retail oil prices and the international crude oil price in some cities of China. Moreover, Monge and Infante [8] analyzed the crude oil prices using ARFIMA models to assess whether shocks in the series have temporary or persistent effects.

Within this research stream, the main aim of this paper is to analyze the increase in the spread refined-crude petroleum prices in the US in the post-COVID period, considering the main explanatory variables according to EIA [9]: crude petroleum prices, refining costs, distribution costs, and taxes.

The methodological approach, combining ARFIMA and FCVAR models, allows us to capture both persistence and long-run equilibrium relationships among structural variables—an approach aligned with recent research emphasizing causal mechanisms in economic transmission (e.g., [10]).

1.2. Literature review

Oil prices significantly affect the international economy [11,12]. When oil becomes unstable or shows bull trends, inflation and economic growth suffer, influencing monetary policy. Thus, oil price analysis is crucial both micro- and macroeconomically. Due to this relevance, many models have been developed to explain and forecast crude and refined oil prices.

Monge and Infante [13] used ARFIMA models to examine whether crude oil price shocks are temporary or persistent, using annual data from 1861 to 2019. Their results show mean reversion in the series through the estimated value of d.

Lu et al. [14] introduced a method to identify key price determinants and forecast oil prices. They applied GLMNET, spike-slab lasso, and Bayesian model averaging for selection, and LSTM for forecasting.

Choi et al. [7] studied the short- and long-term relationships between international crude oil and retail prices in Shenzhen, Hong Kong, and Macao. Using an asymmetric error correction model, they confirmed long-term links between gasoline, diesel, and crude oil prices in all three cities.

Bragoudakis et al. [15] investigated if retail gasoline prices respond asymmetrically to international crude oil prices, analyzing daily data across the supply chain. Their Greek case study showed asymmetric behavior in refinery markups, while retailers didn't adjust markups significantly.

Blair et al. [16] examined regional differences in the pass-through from crude oil to retail gasoline prices in the U.S. Using an error correction model, they found that a $1 change in crude oil leads to $2.36–$2.58 changes in retail prices, depending on region, confirming the "rockets and feathers" phenomenon.

Monge et al. [17] explored how the U.S. shale oil revolution affected WTI prices. Using wavelet analysis and long-run dependence methods, they showed structural differences in price behavior before and after the shale boom.

Gil-Alana et al. [18] analyzed crude oil prices using fractional integration on data from 1859 to 2015. They found nonstationary behavior in log prices, and long memory in squared and absolute returns.

Tetteh and Xu [19] predicted fuel prices in Ghana using cointegration, GARCH, and ANN techniques on monthly data. Their results indicated that ANN forecasts outperformed others in predicting stable price increases.

Kilian [20] highlighted the often-overlooked difference between U.S. gasoline and global crude oil prices. He analyzed this using a structural VAR model combining global crude and U.S. gasoline markets, emphasizing how demand and supply shocks affect both.

Choi and Hammoudeh [6] found strong evidence of long memory in daily returns and volatility for oil, gasoline, and heating oil using FIGARCH, showing varying degrees of persistence.

Kaufmann and Laskowski [21] examined price asymmetries in the U. S. petroleum market. They argued that refinery utilization and inventories explain crude-to-gasoline price gaps, and that long-term contracts could explain differences with heating oil prices. These factors imply that market efficiency doesn't eliminate price asymmetries and complicated policy responses.

This literature highlights the complexity of price dynamics in oil markets. However, our study addresses gaps by analyzing factors such as natural gas costs and supply-demand mismatches to better explain the refined–crude oil spread.

Recent research has also emphasized the systemic and interconnected nature of commodity markets under global stress conditions. Anwer et al. [22] document how the COVID-19 pandemic heightened systemic risk linkages between energy and non-energy commodities, illustrating the fragility of cross-market dynamics. Mishra et al. [23] identify strong spillovers between crude oil and agricultural commodities in India during major global disruptions, underscoring intersectoral risk channels. In addition, Rao et al. [24] demonstrate that geopolitical risk exerts a persistent influence on energy spread behavior, even after accounting for financial variables such as interest rates. Complementarily, Liu et al. [25] show that pandemic-related sentiment significantly shaped extreme crude oil returns across multiple time scales, highlighting the role of behavioral drivers in amplifying volatility. Taken together, these studies suggest that energy price spreads are embedded in a broader web of multiscale, nonlinear, and persistent transmission mechanisms, with important implications for pricing, hedging, and policy design in increasingly fragile global systems.

1.3. Research contributions

This paper contributes to existing literature in several ways. First, it provides one of the few empirical analyses focused specifically on the behavior of the refined–crude oil price spread in the US during the post-COVID period—a time marked by sharp demand shocks and supply bottlenecks. Second, it employs advanced time-series methodologies, including ARFIMA and FCVAR models, which offer greater flexibility in capturing persistence and long-run relationships compared to traditional I(1)/I(0) based frameworks. Third, unlike previous studies that often isolate variables such as crude oil prices or refinery markups, our model integrates key structural factors—namely, refining capacity, natural gas prices, and demand–supply imbalances—to explain the spread dynamics in a more comprehensive way. Finally, our findings challenge some prevailing assumptions by showing that natural gas prices, while relevant for production costs, do not exhibit significant causal influence on the refined–crude spread in the long run. These contributions offer both methodological and policy-relevant insights into the complex mechanisms behind refined oil price formation in a post-pandemic economy.

1.4. Paper structure organization

The subsequent sections of the study are structured into four distinct parts. First, we recap the literature review. Second, we describe the methodology applied. Third, we provide the most important empirical findings. Finally, we include some concluding comments.

2. Research design

2.1. Measurement of variables

We examine the evolution of the crude oil (WTI Price, $/Barrel), the refined oil (Gasoline Price, USX/Gallon), the spread refined-crude oil price, that is defined as the gasoline price index minus the crude oil price index, with both series rebased to 100 (January 2003 = 100) to ensure unit consistency), the industrial natural gas price (index), the petroleum consumption (thousands barrel/day), and the industrial utilization capacity through crude processing (percentage) and considering constant taxes following EIA [9].

2.2. Explanation of data

The time series are monthly, and the sample period goes from January 2003 to October 2023; therefore, the total number of observations in each case is 250. The chosen time span includes the last two NBER-dated recessions [26]. The data source is the Thomson Reuters Eikon-Datastream (currently known as LSGE Datastream).

In our study, we define the start of the post-COVID period as March 2021, which corresponds to two overlapping criteria:

- Theoretical rationale: This period marks the beginning of a significant recovery in global economic activity following the widespread rollout of COVID-19 vaccination programs and the gradual lifting of mobility restrictions across major economies. It aligns with the transition from the acute health crisis phase to a reopening and normalization of energy demand, particularly in the transportation and industrial sectors. This interpretation is supported by the IMF and EIA, which identify early 2021 as a turning point in energy consumption patterns [9,27].
- Empirical evidence of structural change: In our dataset, a marked divergence appears starting in early 2021 between crude oil prices and refined petroleum prices, as well as in the spread between them. Specifically, our Chart 1 shows a structural shift in the refined–crude oil spread beginning in March 2021. This corresponds to a surge in petroleum demand that outpaced supply-side recovery, especially in refining capacity, due to lingering shutdowns and supply chain disruptions. These changes justify treating March 2021 as the beginning of a structurally distinct regime.

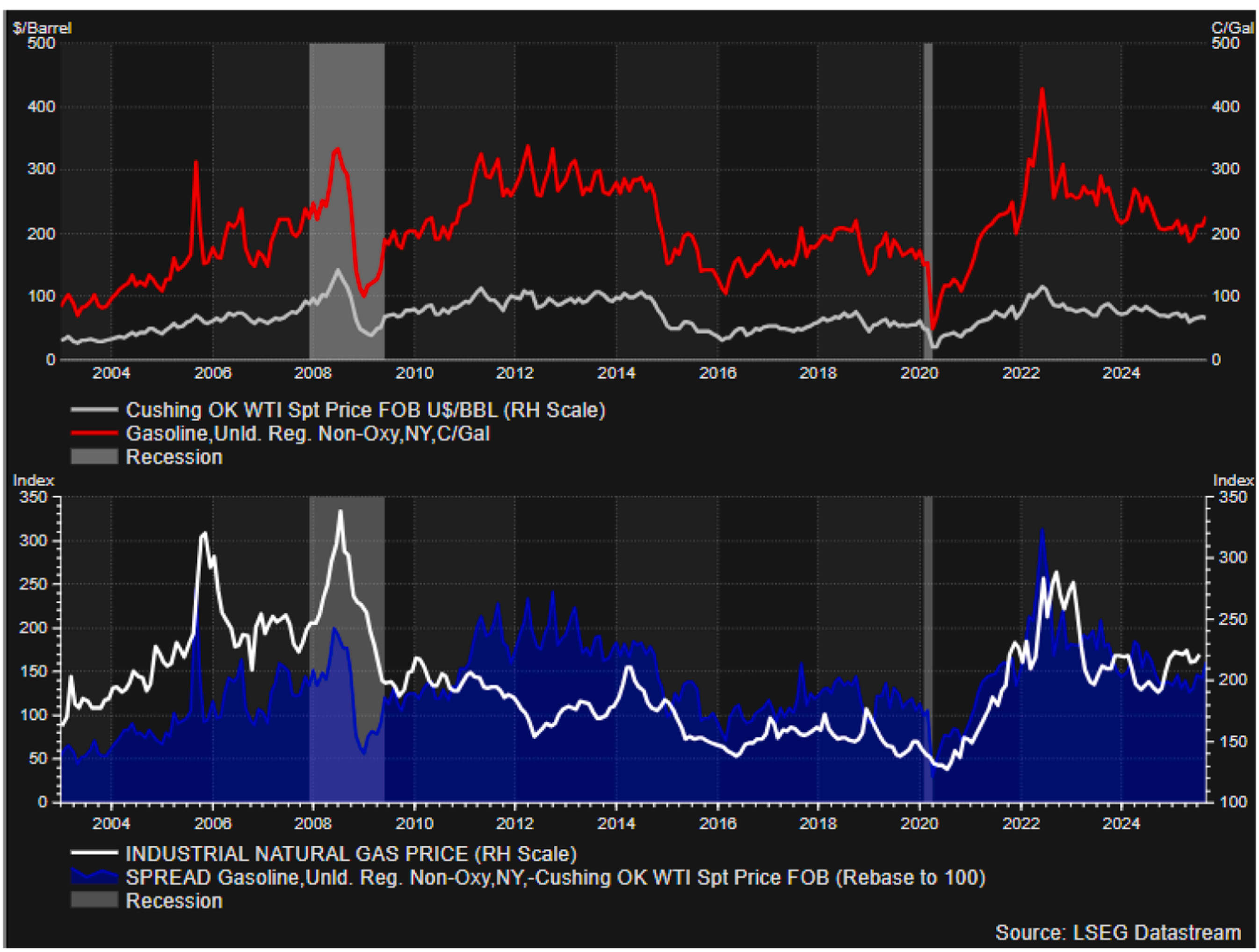

Fig. 1. The spread of refined-crude oil price was rocketing in the US after COVID, partly because of gas natural prices increasing.
Note: Panel 1 reports the price evolution of crude oil (USD/barrel) and gasoline (cents/gallon). Panel 2 displays the industrial natural gas price (index) and the gasoline–crude oil spread (both variables have been transformed into index numbers, rebased to 100). The spread is defined as the gasoline price index minus the crude oil price index, both series rebased to 100. All data are monthly and sourced from Thomson Reuters Datastream.

Additionally, although the WHO [28] formally declared the end of the COVID-19 global health emergency in May 2023, the economic and energy market recovery began earlier, which justifies selecting an empirical cutoff based on observed data behavior and economic conditions, rather than the official end of the pandemic itself.

In Fig. 1 (upper part), we observe an enormous increase in refined oil prices from and after the COVID pandemic, with costs reaching above those in the Great Recession. We also see a rise in the price of crude oil. According to these data, we could assume crude oil prices are affecting refined oil prices; however, in Fig. 1 (lower part), the spread of refined-crude oil prices is rocketing, which reveals the main reason for the refined price rising.

Additionally, in Fig. 1 (lower part) we observe a strong increase in natural gas price, which is correlated to the spread refined-crude oil price spread. This natural gas trend is obviously impacting refined prices. In the Great Recession, natural gas prices were even higher, but they did not affect the spread refined-crude oil as much. This evidence invites us to continue investigating the main reason for the current refined oil price increase.

In Fig. 2, we see an important lag in the supply capacity to respond to the post-COVID demand increase. After the collapse during the pandemic, the petroleum demand has soared, but the industry capacity utilization (crude processing) has not been able to absorb such a demand in the short term. This has provoked a huge rise in refined oil prices due to the bottleneck effect. This consequence was not perceived in the Great Recession.

2.3. Research model

The modeling approach adopted in this study is grounded in well-established economic features of energy markets. Prices of crude and refined petroleum products are influenced by a complex interplay of structural and cyclical factors, including supply chain rigidity, regulatory frameworks, inventory adjustments, and exogenous shocks such as geopolitical crises or pandemics (see [17,29–33]; among others). These factors often lead to nonstationary behavior, persistent shocks, and long memory, especially in post-crisis contexts. In this sense, fractional integration models such as ARFIMA are well-suited to measure the degree of persistence and identify whether price movements exhibit mean-reverting tendencies or permanent shifts.

Moreover, economic theory suggests that while refined product prices are derived from crude oil prices, the transmission is not perfect due to refining costs, taxes, capacity constraints, and demand shifts. This motivates the use of cointegration techniques to test for long-run relationships between these markets. In particular, the FCVAR model allows for fractional orders of integration, enabling a more flexible and realistic modeling of equilibrium behavior than traditional I(1)-based frameworks. These methodological choices are thus directly aligned

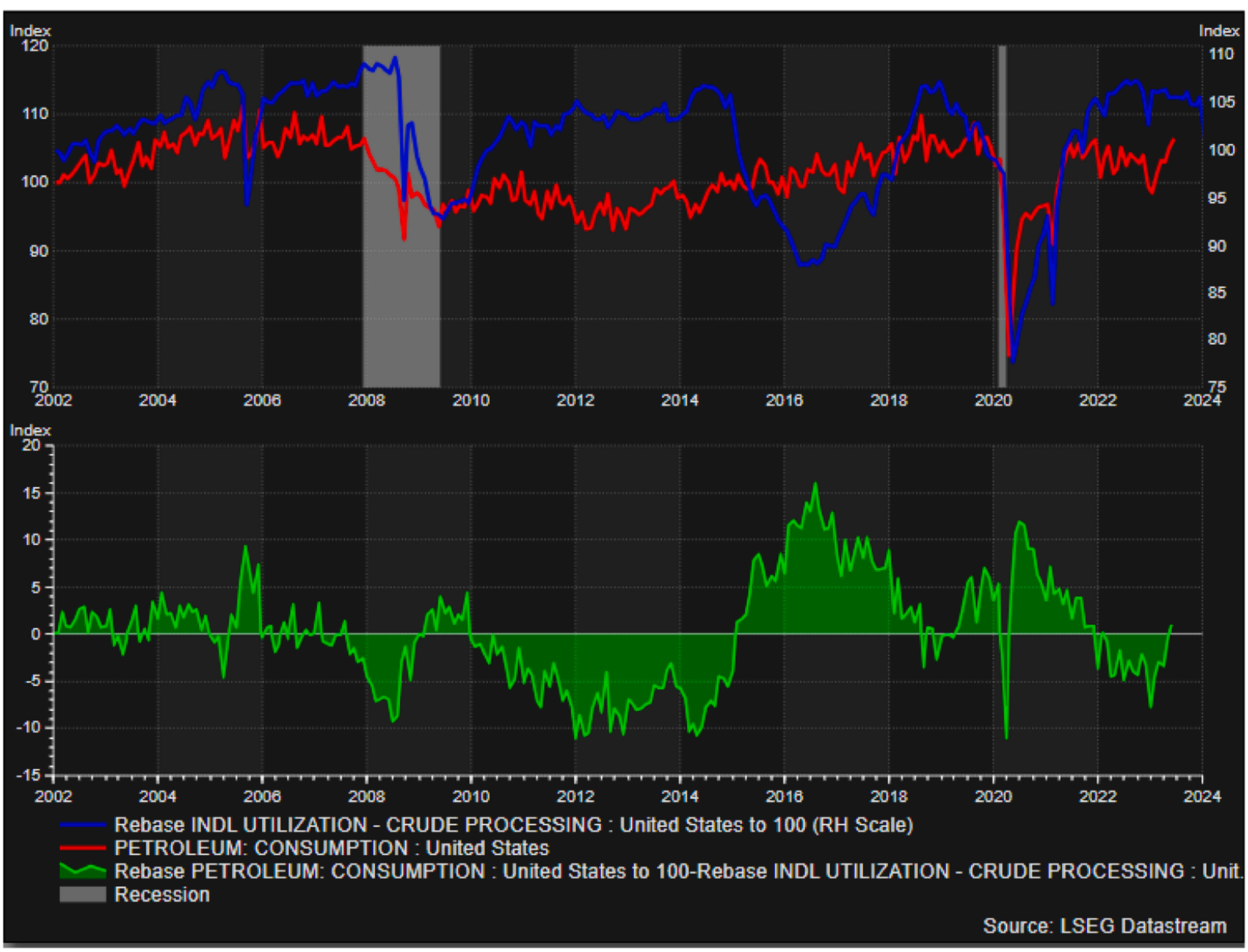

Fig. 2. The supply-demand oil imbalance also seemed an important reason.
Note: Panel 1 reports the industrial capacity utilization (supply-side) and the petroleum consumption (demand-side) (both variables have been transformed into index numbers, rebased to 100). Panel 2 displays the difference between supply and demand. All data are monthly and sourced from Thomson Reuters Datastream.

with the economic characteristics of the oil sector.

2.3.1. Unit roots

To model variables and comprehend interrelationships, statistics and econometrics we employ single or multi-equation regression models of time series (see [34]).

However, it is crucial to comprehend how these time series behave before using these kinds of models. To deal with the series, one must first determine if the process is stationary I(0) when it does not have a unit root or non-stationary I(1) when it does (see [35]).

Therefore, the analysis employs the conventional unit root test to ascertain each time series' integration order. The Dickey-Fuller test is the most popular and commonly used unit root test (see [36]). The Augmented Dickey-Fuller test is created if a non-systematic component in the Dickey-Fuller models is autocorrelated (see [37]). Because of their higher power, several different tests have been taken into consideration. These include Phillips [38] and Phillips and Perron [39], which employed a non-parametric estimate of the spectral density u_t at the zero frequency. The methodology for analyzing the deterministic trend is based on Kwiatkowski et al. [40].

2.3.2. ARFIMA (p, d, q) model

We use a more sophisticated method after utilizing typical unit root tests to evaluate the integration sequence of each time series. The number of differences does not always need to be an integer value to achieve stationarity I(0); instead, it can be any point on the real line, which makes it fractional I(d) (see [41–45]).

Therefore, we differentiate the time series using a fractional number to make the time series stationary I(0). Because of the lower power under fractional alternatives, this is a more sophisticated approach than unit root testing (see [46–48]).

Determining and capturing the persistence of the data is another element of the I(d) models. This is the situation where observations are strongly linked yet separated in time.

The ARFIMA (p, d, q) model is the fractional integrated approach that we employ in this study article. The mathematical notation for this model is:

$$(1-L)^d x_t = u_t,\ t = 1,\ 2, \tag{1}$$

The covariance stationary process I(0), which has a positive and finite spectral density function at zero frequency and a weak form of time dependence, is denoted by the symbol u_t in Eq. (1). The time series with an integrated process of order d is denoted by $(x_t \approx I(d))$, where d can be any real number and L is the lag-operator $(Lx_t = x_{t-1})$.

Consequently, we can say that x_t is $ARMA(p,\ d,\ q)$ if u_t is $ARMA\ (p,\ q)$.

Eq. (1) yields the binomial expansion for the polynomial $(1-L)^d$ where x_t for every real d depends on its entire history in addition to a finite number of previous observations. Accordingly, a larger value of d denotes a stronger degree of correlation between the series' observations.

The value of the parameter d determines which cases we can distinguish between. Table 1 presents a summary of the various outcomes of d.

Numerous techniques can be used to measure the degree of fractional integration and long-memory ([49–55]). However, this study applies the Akaike information criterion (AIC) (see [56]) and the Bayesian information criterion (BIC) (see [57]) to select the best ARFIMA model.

Table 1
Interpretation of the results of d for the ARFIMA model.

$d = 0$	x_t process is short memory
$d > 0$	x_t process is long memory
$d < 0.5$	x_t is covariance stationary
$d \geq 0.5$	x_t is nonstationary
$d < 1$	x_t is mean reverting
$d \geq 1$	x_t is not mean reverting

2.3.3. Breitung–Candelon test

The causality test proposed by Breitung and Candelon [58] contributes to providing an idea about whether the relationship between both time series is temporary or permanent (see [59–61]). Because it interprets Granger causality across several frequency domains, this test has an advantage over other frequently used causality tests. To this end, two-time series—one based on coherence and the other on the bivariate spectral-density matrix—are categorized according to their spectral associations. An overall count of immediate forward and backward causality mechanisms is then obtained from the categorization.

According to Breitung and Candelon [58], the VAR(p) model below can be used to specify the interdependence between two variables, x and y:

$$x_t = \alpha_1 x_{t-1} + \alpha_p x_{t-p} + \beta_1 y_{t-1} + \ldots + \beta_p y_{t-p} + \beta_{1t} \tag{2}$$

The null hypothesis, $H_0 : M_{y\to x}(w) = 0$, as tested by Geweke [62], matches the null hypothesis of linear restriction given as:

$$R(w)\beta = 0 \tag{3}$$

Where β denotes the coefficient vector of y. $R(w)$ is defined as:

$$R(w) = \begin{bmatrix} \cos(w)\cos(2w)\ldots\cos(pw) \\ \sin(w)\ \sin(2w)\ldots\sin(pw) \end{bmatrix} \tag{4}$$

The F-statistics for the null hypothesis in Eq. (5) has an approximated distribution of $F(2,\ T-2P)$ for $Fw \in (0,\ \pi)$. Furthermore, co-integration is frequently used as a framework for examining the frequency-based Granger causality test. Therefore, Breitung and Candelon [58] substitute x_t in Eq. (4) for Δx_t. As a result, the existence of cointegration between the series suggests that the primary long-term causation and zero-frequency causality share conceptual similarities. However, if there is no long-term link in the stationary case, the evidence of a causal association at a low frequency implies that the variable under consideration's frequency element can be predicted by a different variable.

2.3.4. FCVAR model

Following Johansen and Nielsen [63], this study applies their multivariate Fractional Cointegrated VAR (FCVAR) model to check the relationship of the variables in the long term. The FCVAR model is notated in the next equation:

$$\Delta^d X_t = \ \alpha\beta' L_b \Delta^{d-b} X_t + \sum_{i=1}^{k} \Gamma_i \Delta^b L_b^i Y_t + \varepsilon_t \tag{5}$$

Where ε_t is a term with mean zero and variance-covariance matrix Ω that is p-dimensional independent and identically distributed, and α and β are $p \times r$ matrices where $0 \leq r \leq p$. The relationship in the long-term equilibria in terms of cointegration in the system is due to the matrix β. Parameter Γ_i controls the short-term behavior of the variables. Finally, the deviations from the equilibria and their speed in the adjustment is due to parameter α.

Compared to traditional cointegration models (e.g., Engle-Granger or Johansen CVAR), the FCVAR model presents key advantages for our study. First, many variables are fractionally integrated, with d values significantly different from 1. Classical models require variables to be I (1) and do not accommodate fractional orders of integration, potentially leading to biased or inefficient estimates. Second, the FCVAR model permits long-run cointegrating relationships among I(d) variables, allowing us to capture persistence and partial mean reversion dynamics. This provides a more accurate framework for assessing the equilibrium behavior of refined–crude oil price spreads and their driving forces.

3. Basic analysis

The first analysis that we carried out in this research paper is the unit root/stationarity test to analyze the behavior of the time series mentioned above. To do this analysis we performed the Augmented Dickey-Fuller (ADF) test, the Phillips Perron (PP) test and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test.

Table 2 displays the results that we obtained using the unit roots that we have mentioned before. We find that the variables analyzed have a non-stationary I(1) behavior, except for petroleum consumption, and therefore the results indicate evidence that an exogenous shock has permanent effects.

While unit root tests provide a useful classification, they rely on a binary distinction (I(0) vs I(1)) and often lack power against fractional alternatives. Therefore, we extend our analysis using ARFIMA models to better capture the persistence and long-memory properties of the data.

Thus, in the following, we allow for fractional differentiation throughout the previously mentioned ARFIMA framework, thereby providing additional flexibility in capturing the persistence and long-memory properties of the time series. We employ ARFIMA (p, d, q) models to analyze the behavior of key variables of the original time series and after the COVID-19 shock, particularly because conventional unit root tests have lower power under fractional alternatives.

The appropriate AR and MA orders are selected using the Akaike Information Criterion (AIC) [56] and the Bayesian Information Criterion (BIC) [57]. Since the AIC and BIC might not be the ideal criteria for applications requiring fractional models, care should be taken in this case (see [45,64]). We estimate different ARFIMA model specifications using Sowell's [54] maximum likelihood estimator for combinations where $p, q \leq 2$. Table 3 reports the estimated fractional differencing parameter d as well as the selected AR and MA terms for each time series.

Although the full sample from January 2003 to October 2023 includes 250 monthly observations, ensuring robust estimation and inference, special attention is required when analyzing the post-COVID sub-period. This sub-sample, spanning from March 2020 to October 2023, includes >40 monthly observations, which may appear modest for complex time-series models. Nevertheless, this sample size is widely accepted in econometrics as sufficient to apply asymptotic methods. According to the Central Limit Theorem, sample sizes greater than 30 are generally adequate for normal approximation and valid inference using likelihood-based estimators.

Moreover, fractional integration models such as ARFIMA are particularly well suited to moderate sample sizes. Sowell [54] demonstrated that the maximum likelihood estimator for ARFIMA remains efficient even with fewer than 50 observations, provided model selection is guided by parsimonious criteria like AIC and BIC. Likewise, Beran et al. [64] showed that bias in the estimation of fractional models remains low when sample sizes exceed 30 and the differencing parameter dd does not approach the unit root boundary.

Analyzing the results of the original time series from Table 3, we see that the estimates of d in all cases are fractional. Focusing on each variable, we observe that the WTI crude oil prices, Industrial Natural Gas and the processing capacity of the industry in refining crude are 1 or higher than 1 ($d \geq 1$), observing that there is a high degree of persistence and finding evidence of non-mean reversion. For the rest of the variables, we observe that the values of d are in the range $(0, 1)$,

Table 3
Results of the long memory tests.

Data Analyzed		Model Selected	d	Std. Error	Interval	$I(d)$
Original Time Series						
WTI Price	250	ARFIMA (0, d, 0)	1.10	0.062	[1.00, 1.20]	I(1)
Refined Oil: Gasoline Price	250	ARFIMA (0, d, 0)	0.89	0.059	[0.79, 0.99]	I(d)
Spread Refined Crude Oil Price	250	ARFIMA (1, d, 2)	0.75	0.305	[0.25, 1.26]	I(1)
Industrial Natural Gas	249	ARFIMA (0, d, 0)	1.08	0.058	[0.98, 1.17]	I(1)
Petroleum Consumption	245	ARFIMA (2, d, 2)	0.64	0.054	[0.55, 0.73]	I(d)
Industrial Utilization Capacity: Crude Processing	249	ARFIMA (2, d, 2)	1.00	0.056	[0.90, 1.09]	I(1)
Imbalance d-S	249	ARFIMA (1, d, 2)	0.86	0.114	[0.68, 1.05]	I(1)
After COVID-19						
WTI Price	47	ARFIMA (1, d, 2)	0.65	0.229	[0.27, 1.03]	I(1)
Refined Oil: Gasoline Price	47	ARFIMA (0, d, 0)	0.88	0.142	[0.65, 1.12]	I(1)
Spread Refined Crude Oil Price	47	ARFIMA (2, d, 2)	0.72	0.157	[0.46, 0.97]	I(d)
Industrial Natural Gas	46	ARFIMA (1, d, 2)	0.95	0.295	[0.47, 1.44]	I(1)
Petroleum Consumption	42	ARFIMA (0, d, 0)	0.82	0.191	[0.50, 1.13]	I(1)
Industrial Utilization Capacity: Crude Processing	46	ARFIMA (0, d, 0)	0.97	0.142	[0.74, 1.21]	I(1)
Imbalance d-S	47	ARFIMA (2, d, 2)	0.73	0.447	[−0.01, 1.46]	I(0), I(1)

Table 2
Unit root tests.

	ADF			PP		KPSS	
	(i)	(ii)	(iii)	(ii)	(iii)	(ii)	(iii)
Original Data							
WTI Price	−0.5731	−3.0479*	−3.0381	−2.9325*	−2.9204	0.3994*	0.4142
Refined Oil: Gasoline Price	−0.6179	−3.2514*	−3.3233	−3.1697*	−3.2468	0.5457	0.3765
Spread Refined Crude Oil Price	−0.7698	−3.5276*	−3.6994*	−3.5753*	−3.7901*	0.6894	0.3435
Industrial Natural Gas	−0.3	−2.3561	−2.5994	−2.3473	−2.6162	1.2135	0.3813
Petroleum Consumption	−0.1035	−3.7229*	−3.6988*	−4.3952*	−4.4094*	0.5526	0.4594
Industrial Utilization Capacity: Crude Processing	0.1046	−2.5312	−2.4734	−2.6531	−2.6064	0.478	0.1791
Imbalance d-S	−2.8485*	−2.8337	−2.8341	−3.2943*	−3.3004	0.5062	0.3285
After COVID-19							
WTI Price	0.0007	−1.2898	−2.3818	−1.2073	−2.0603	0.8498	0.1859
Refined Oil: Gasoline Price	−0.0477	−1.3418	−2.2306	−1.3676	−2.185	0.9077	0.1713
Spread Refined Crude Oil Price	−0.1151	−1.4221	−2.2486	−1.4881	−2.3453	0.9224	0.1625
Industrial Natural Gas	0.3273	−1.2887	−1.2351	−1.2066	−1.3076	0.9328	0.1792
Petroleum Consumption	0.0014	−2.4129	−3.9295*	−2.2937	−3.3765	0.6272	0.0985
Industrial Utilization Capacity: Crude Processing	0.3701	−1.0527	−3.124	−1.0522	−2.8532	0.9413	0.1063
Imbalance d-S	−2.4202*	−2.5356	−3.0116	−2.6329	−3.0355	0.3972*	0.2303

* Denotes a statistic significant at the 5 % level.

implying mean reversion.

Focusing on the period after COVID-19, we observe that all variables present a mean reversion behavior ($d < 1$) in all cases. Focusing on the confidence intervals, we cannot reject the hypothesis of I(1) in the most variables due to the great uncertainty, except for the spread refined-crude oil prices. So, we conclude that extraordinary measures will be necessary to reestablish trends caused by this shock.

It is true that after the appearance of COVID-19, "industrial utilization capacity: crude oil transformation" shows more persistence than oil consumption and the industrial price of natural gas. Reuters [65] indicated that *"since the start of the global pandemic, the United States has lost nearly a million barrels a day of oil refining capacity, and more will be shut down in the coming years"*. With the results obtained, the results indicate that this is not the case, as the refining-crude oil price differential shows a reversion to the mean.

Having established the fractional integration properties of each series, we proceed to examine the causal dynamics between variables across different time horizons (short, medium, long) using the frequency-domain approach proposed by Breitung and Candelon [58].

These results will be very important, since by means of the fractional cointegration analysis we will see how the variables significant to this test are related in the long run. The results are displayed in Table 4.

We find different results using the frequency domain causality test for the full-time series. Focusing on the results of the Wald test statistics and the p-value (in brackets) that are in Table 4, the results indicate that the spread refined-crude oil prices cause effects in the long term to refined-oil prices, natural gas prices, and the industrial utilization capacity (crude processing). On the other hand, the results indicate that the WTI crude oil prices affect the spread.

Given the importance of the relationship between the variable "Spread" and "Imbalance" for this study, we see in our results that the relationship between both variables is statistically significant in a bidirectional manner at a frequency of 0.05. That is, the spread has causal effects on the imbalance variable and vice versa only in the long run. This result is consistent with economic theory which posits that an increase in an imbalance demand-supply puts pressure on final prices: the refined oil prices tend to rise as a consequence to impact spread refined-crude oil prices.

Once we have measured the causal relationship between the variable considered in the time-frequency domain and turn our attention to the long-term, this study applies a methodology based on Fractional Cointegrating VAR model to understand the relationship that exists in the long-term between the variables that we found statistically significant.

To do this, we follow the model introduced by Johansen [66] and expanded by Johansen and Nielsen [63]. The model is called the Fractional Cointegrated Vector AutoRegressive (FCVAR) model, and it is a step ahead of the Cointegrated Vector AutoRegressive (CVAR) model proposed by Johansen [67].

Before presenting the FCVAR estimations, we first employed the Johansen–Nielsen [63] likelihood ratio test to determine the appropriate cointegration rank. This step is essential, because the FCVAR framework assumes the existence of at least one long-run relation among the variables. Without formally establishing cointegration, the estimation of fractional cointegration vectors and adjustment coefficients would not be meaningful.

The results of the LR test in Table 5 confirm that in all systems the null of no cointegration ($r = 0$) is rejected, while the null of one cointegrating relation ($r = 1$) cannot be rejected. This implies that the optimal rank is one across all specifications, which validates the FCVAR estimations reported in Table 6.

From an economic perspective, this means that the spread between refined crude oil and unrefined crude oil shares a long-term equilibrium with refined products, WTI, and processing capacity. Conversely, although the test suggests a cointegration relationship with natural gas and imbalance, the estimated relationship is weak, indicating that both variables do not play a structural role in the long term.

The results of the FCVAR model have been summarized in Table 6.

According to the causality test and the results that we get using the FCVAR model to analyze the relationship in the long-term between several variables after COVID-19, we are going to focus on two terms, the integrating and cointegrating part ($d \neq b$) and the beta term, to analyze the behavior of the time series.

The analysis reveals from Panel I to V that the order of integration of the individual series is lower than 1 in all cases ($d < 1$) as we get the same magnitude in the reduction in the degree of integration in the cointegrating regression. These results imply I(0) cointegrating errors. Therefore, from the results of our cointegration analysis, we do not rule out the hypothesis that the effects of the shock disappear in the short run.

On the other hand, it is observed through the beta parameters that a growth of the imbalance demand-supply (1.000) is preceded by a slightly negative impact of the Spread (–0.050). This behavior is observed because the initial increasing in spread due to the rise in imbalance immediately after COVID has been corrected step by step for the supply adaptation. When supply adapts to demand, the rise in spread practically disappears.

According to our results, supply–demand imbalances, refining

Table 4
Breitung and candelon frequency domain causality test results.

Hypothesis	Long Term ($\omega = 0.05$)	Medium Term ($\omega = 1.5$)	Short Term ($\omega = 2.5$)
Original Time Series			
d_Spread →d_Refined	26.78*	13.27*	5.04
	(0.00)	(0.00)	(0.08)
d_Refined →d_Spread	24.62*	10.74*	5.32
	(0.00)	(0.00)	(0.07)
d_Spread →d_WTI	5.20	5.33	2.47
	(0.07)	(0.07)	(0.29)
d_WTI →d_Spread	24.62*	10.74*	5.32
	(0.00)	(0.00)	(0.07)
d_Spread →d_Gas	7.34*	3.70	0.59
	(0.03)	(0.16)	(0.74)
d_Gas→d_Spread	0.188	0.042	0.309
	(0.91)	(0.98)	(0.85)
d_Spread →d_Consumption	1.00	0.50	1.20
	(0.60)	(0.77)	(0.55)
d_Consumption→d_Spread	0.96	0.88	0.47
	(0.61)	(0.64)	(0.79)
d_Spread →d_Processing	7.38*	4.69	0.66
	(0.02)	(0.09)	(0.71)
d_Processing→d_Spread	2.99	2.53	0.12
	(0.22)	(0.28)	(0.94)
d_Spread →d_Imbalance	13.11*	0.59	0.06
	(0.00)	(0.74)	(0.96)
d_Imbalance→d_Spread	6.62*	3.05	3.44
	(0.03)	(0.22)	(0.18)

* This shows that there is a significant causality relationship at the 5 % significance level. The values in the brackets are the probability value of the F statistics calculated for the relevant ω values.

Table 5
Likelihood ratio tests for cointegration rank in the FCVAR models.

	LR Statistic for Rank = 0	LR Statistic for Rank = 1	Decision
Panel I: Spread (Var 1) vs Refined (Var 2)	3.641	0.483	Do not reject Rank =1
Panel II: WTI (Var 1) vs Spread (Var 2)	3.641	0.483	Do not reject Rank =1
Panel III: Spread (Var 1) vs Gas (Var 2)	5.082	–2.566	Do not reject Rank =1
Panel IV: Spread (Var 1) vs Processing (Var 2)	3.819	–5.111	Do not reject Rank =1
Panel V: Imbalance (Var 1) vs Spread (Var 2)	2.066	–1.293	Do not reject Rank =1

Table 6
Results of the FCVAR model.

	$d \neq b$	Cointegrating equation beta	
		Var 1	Var 2
Panel I: Spread (Var 1) vs Refined (Var 2)	$d = 0.010\ (0.122)$	1.000	–0.691
	$b = 0.010\ (0.001)$		
	$\Delta^d\left(\begin{bmatrix} Spread \\ Refined \end{bmatrix} - \begin{bmatrix} 99.078 \\ 149.096 \end{bmatrix}\right) = L_d \begin{bmatrix} 820902.048 \\ 1076719.781 \end{bmatrix} \nu_t + \sum_{i=1}^{2} \widehat{\Gamma}_i \Delta^d L_d^i (X_t - \mu) + \varepsilon_t$		
Panel II: WTI (Var 1) vs Spread (Var 2)	$d = 0.477\ (0.000)$	1.000	3.079
	$b = 0.477\ (0.000)$		
	$\Delta^d\left(\begin{bmatrix} Spread \\ WTI \end{bmatrix} - \begin{bmatrix} 31.975 \\ 57.371 \end{bmatrix}\right) = L_d \begin{bmatrix} 0.003 \\ 0.017 \end{bmatrix} \nu_t + \sum_{i=1}^{2} \widehat{\Gamma}_i \Delta^d L_d^i (X_t - \mu) + \varepsilon_t$		
Panel III: Spread (Var 1) vs Gas (Var 2)	$d = 0.010\ (0.105)$	1.000	−1.184
	$b = 0.010\ (0.000)$		
	$\Delta^d\left(\begin{bmatrix} Spread \\ Gas \end{bmatrix} - \begin{bmatrix} 103.011 \\ 155.371 \end{bmatrix}\right) = L_d \begin{bmatrix} -173208.295 \\ -55407.236 \end{bmatrix} \nu_t + \sum_{i=1}^{2} \widehat{\Gamma}_i \Delta^d L_d^i (X_t - \mu) + \varepsilon_t$		
Panel IV: Spread (Var 1) vs Processing (Var 2)	$d = 0.010\ (0.160)$	1.000	−3.865
	$b = 0.010\ (0.001)$		
	$\Delta^d\left(\begin{bmatrix} Spread \\ Processing \end{bmatrix} - \begin{bmatrix} 142.948 \\ 82.381 \end{bmatrix}\right) = L_d \begin{bmatrix} 27993.285 \\ 10695.113 \end{bmatrix} \nu_t + \sum_{i=1}^{2} \widehat{\Gamma}_i \Delta^d L_d^i (X_t - \mu) + \varepsilon_t$		
Panel V: Imbalance (Var 1) vs Spread (Var 2)	$d = 0.010\ (0.150)$	1.000	–0.050
	$b = 0.010\ (0.000)$		
	$\Delta^d\left(\begin{bmatrix} Imbalance \\ Spread \end{bmatrix} - \begin{bmatrix} -1.978 \\ 112.283 \end{bmatrix}\right) = L_d \begin{bmatrix} 122235.185 \\ 209609.381 \end{bmatrix} \nu_t + \sum_{i=1}^{2} \widehat{\Gamma}_i \Delta^d L_d^i (X_t - \mu) + \varepsilon_t$		

capacity constraints, and production costs (proxied by natural gas prices) affect the refined–crude oil price spread, aligned with classical and recent literature in energy economics (see [2,6,20,21,65,68]):

1. Supply–demand imbalance as a transmission channel: Economic theory posits that when aggregate demand for refined oil rises more quickly than the available refining capacity can respond, the resulting bottleneck leads to a higher spread between crude and refined oil prices [20,68]. Our results confirm this: the results indicate a significant long-run causal relationship between the imbalance variable and the spread, which is consistent with the theoretical expectation that demand shocks translate into price increases when supply is inelastic in the short term.
2. Refining capacity constraints: The role of refining infrastructure as a price amplifier has been well-documented [21,65]. Our post-COVID analysis shows that industrial utilization capacity remained persistently below pre-pandemic levels, which helps explain the non-mean-reverting behavior of the spread immediately after the shock. The use of fractional integration and cointegration allows us to capture this persistence more accurately than standard models.
3. Natural gas as an input cost: While some industry reports (e.g., [2]) suggest a transmission from natural gas to refined product prices, our empirical findings show that natural gas prices do not significantly influence the spread in the long run. This result supports the view that cost-push factors may be less relevant than capacity or demand-side constraints during crisis periods—a nuance that adds to the theoretical discussion on pass-through effects in the energy sector [6].
4. Transmission mechanism modeled via ARFIMA and FCVAR: Our methodology directly maps onto the economic transmission channels. The ARFIMA framework models the persistence of shocks in key variables (e.g., WTI, natural gas, spread), while the FCVAR framework identifies long-run equilibrium relationships and causal linkages. These econometric techniques provide a formal structure to capture how changes in fundamentals propagate through the market over time.

4. Discussion and conclusions

The main goal of this paper was to analyze the rise in the spread refined-crude petroleum prices in the US after COVID, considering the key explanatory variables according to EIA [9].

Regarding the Long Memory Analysis applied, we observe a high degree of persistence in variables such as WTI crude oil prices, industrial natural gas, and crude processing capacity, with estimates of the fractional differencing parameter $d \geq 1$, indicating non-mean reversion. However, in the post-COVID period, all variables exhibit mean reversion behavior ($d < 1$), although we cannot reject the hypothesis of I(1) in most variables due to great uncertainty. Thus, extraordinary measures will be necessary to reestablish trends caused by this shock.

Using the Breitung and Candelon frequency domain causality test, the study finds that WTI crude and refined oil prices influence the spread, and there is a significant causality in the long-run between the imbalance of demand-supply and the spread.

The long-term relationships among variables are further explored using the FCVAR model. The analysis indicates that the order of integration of individual series is <1, implying I(0) cointegrating errors. This suggests that the effects of the COVID-19 shock are likely to disappear in the short run. The beta parameters show that an increase in the demand-supply imbalance is initially accompanied by an increase in the spread, but this effect diminishes as supply adapts to demand over time.

While our econometric framework focuses on the persistence and long-run relationships among core energy variables, our findings should also be interpreted in light of recent evidence on systemic risk and cross-commodity spillovers. Studies such as Anwer et al. [22], Rao et al. [24], Liu et al. [25], and Mishra et al. [23] highlight that energy spreads are shaped not only by supply–demand imbalances and refining constraints but also by contagion effects, geopolitical tensions, and sentiment-driven volatility across interconnected commodity markets. Future research could therefore extend our approach by incorporating systemic risk measures, geopolitical risk indices, agricultural price linkages, or sentiment variables into fractional integration and cointegration frameworks. This would allow for a richer assessment of how global disruptions transmit through multiple channels to affect refined–crude oil price spreads.

Our findings are consistent with previous studies that identified long memory and persistence in energy markets. For example, Choi and Hammoudeh [6] and Choi et al. [7] documented strong persistence and spillover dynamics in crude oil and refined product markets, while Monge and Infante [8] applied fractional integration techniques to commodity spreads with similar evidence of long-run dependence. However, our results differ in showing that, in the post-COVID period, the refined–crude oil spread exhibits non-mean-reverting behavior, which suggests a structural change in the dynamics of U.S. energy markets. This conclusion also resonates with recent work on systemic risks and spillovers in commodity markets during global disruptions, thereby positioning our contribution within the broader literature.

These findings provide critical insights into the dynamic behavior of oil-related time series data, emphasizing the impact of COVID-19 on market variables and the necessity for adaptive measures in response to significant economic shocks.

According to Ding et al. [68], the widening of the price differential between refined and crude oil prices in the aftermath of the COVID-19 epidemic may be attributed to many reasons, including interruptions in the supply chain, shifts in demand patterns, and deliberate profit-seeking strategies employed by refiners.

Regarding disruptions in the supply chain, the COVID-19 pandemic resulted in substantial disturbances to worldwide supply networks, impacting both the production of crude oil and the operations of refineries. Simultaneously, logistical difficulties and the closure of refineries caused constraints in the supply chain. As global conditions improved, there was a rapid increase in demand that exceeded the refineries' capacity to raise production fast. This resulted in higher spreads.

Concerning shifts in demand trends, the period of economic recovery after the epidemic witnessed a rapid increase in the demand for gasoline and diesel, due to the reopening of businesses and the resumption of travel. Nevertheless, the supply of crude oil did not match the sharp surge in demand. Geopolitical tensions, OPEC+ production cutbacks, and slower-than-anticipated growth in shale oil production all had a role in reducing the availability of crude oil, leading to a decrease in the supply of refined oil products and a rise in prices [69].

These two factors, that we have condensed in our research as an imbalance in demand and the supply of oil, have had certain impact on spread after COVID. The initial imbalance demand-supply has been reducing little by little because of the reaction of supply and have alleviated the pressure on spreads. Furthermore, natural gas (part of cost of production) does not seem to have a key impact on spreads.

Our analysis incorporates a direct supply-side indicator through the variable "industrial utilization capacity in crude processing", which captures the extent to which existing refining infrastructure is employed at the national level. The results show that this measure plays a central role in explaining the persistence and cointegration dynamics of the refined–crude spread during the post-COVID period. Nevertheless, we acknowledge that this aggregate indicator does not account for refinery-specific or regional heterogeneity, such as PADD-level utilization, planned outages, or seasonal maintenance cycles. Future research could extend our framework by incorporating such disaggregated supply-side variables or by simulating counterfactual scenarios (e.g., higher utilization rates), which would provide an even richer basis for assessing the policy implications of refining bottlenecks.

As far as strategic profit-seeking is concerned, refiners have been actively optimizing their processes to maximize profitability in response to shifting market dynamics. According to Ding et al. [68], the crack spread has considerably expanded, suggesting increased profitability in the refining sector. Refiners have improved their profitability in a turbulent market by improving output and controlling inventory levels. This variable might be incorporated in our models in future research.

This study has several unique strengths. First, it applies advanced fractional integration (ARFIMA) and fractional cointegration (FCVAR) methods to capture the persistence and long-run equilibrium dynamics of the refined–crude oil price spread in the post-COVID period. Second, it incorporates frequency-domain Granger causality testing, which allows us to distinguish between short- and long-run drivers of the spread. Third, it focuses on the refined–crude oil spread as a policy-relevant but relatively underexplored indicator of energy-related inflationary pressures in the United States.

At the same time, some limitations must be acknowledged. The analysis is constrained by the availability of disaggregated refinery-level and regional supply-side data, which limits our ability to incorporate capacity utilization and outage effects directly into the econometric framework. In addition, the study is restricted to the U.S. market, and future work could extend the analysis to other economies for comparative purposes. Finally, while fractional integration provides valuable insights into persistence and cointegration, complementary approaches such as regime-switching models, nonlinear specifications, or machine learning techniques could further enrich the understanding of refined–crude price dynamics.

CRediT authorship contribution statement

Manuel Monge: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Carlos Poza:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

All authors have participated in (a) conception and design, or analysis and interpretation of the data; (b) drafting the article or revising it critically for important intellectual content; and (c) approval of the final version.

This manuscript has not been submitted to, nor is under review at, another journal or other publishing venue.

The authors have no affiliation with any organization with a direct or indirect financial interest in the subject matter discussed in the manuscript

Data availability

Data will be made available on request.

References

[1] EIA (2022b). What drives crude oil prices? U.S. Energy Information Administration. Gasoline and Diesel Fuel Update. July 12, 2022.
[2] Bloomberg (2022). Global natural gas crunch amplifies U.S. Refining boom, for now. Bloomberg, Markets. May 4, 2022.
[3] W. Yang, Y. Yang, Z. Chen, Y. Gu, Systemic impacts of national civilized cities on sustainable development: a quasi-experimental analysis of economic and environmental outcomes in China, Systems 13 (1) (2025) 23.
[4] W. Yang, X. Zheng, Y. Yang, Impact of environmental regulation on export technological complexity of high-tech industries in Chinese manufacturing, Economies 12 (2) (2024) 50.
[5] W. Yang, C. Zhu, Y. Yang, Does urban digital construction promote economic growth? Evidence from China, Economies 12 (3) (2024) 59.
[6] K. Choi, S. Hammoudeh, Long memory in oil and refined products markets, Energy J. 30 (2) (2009) 97–116, https://doi.org/10.5547/ISSN0195-6574-EJ-Vol30-No2-5.
[7] K. Choi, M.H. Liu, Y. Zhang, The relationship between refined retail oil prices and crude oil prices: a tale of three cities in the Greater Bay area of China, Chin. Econ. 54 (3) (2021) 157–175, https://doi.org/10.1080/10971475.2020.1848471.
[8] M. Monge, J. Infante, A fractional ARIMA (ARFIMA) model in the analysis of historical crude oil prices, Energy Res. Lett. 4 (1) (2023), https://doi.org/10.46557/001c.36578.
[9] EIA (2022a). Gasoline explained: factors affecting gasoline prices. U.S. Energy Information Administration. Independent Statistics and Analysis, March 15, 2022.
[10] E. Nkansah, Does climate change exposure affect the intensive and extensive margins of corporate R&D investment? Quant. Finance Econ. 9 (2) (2025) 317–354, https://doi.org/10.3934/QFE.2025011.
[11] Arezki, R., Jakab, Z., Laxton, D., Matsumoto, A., Nurbekyan, A., Wang, H. & Yao, J. (2017). Oil prices and the global economy. IMF Working Paper, WP/17/15. International Monetary Fund.
[12] L.A. Gil-Alana, O.S. Yaya, The relationship between oil prices and the Nigerian stock market. An analysis based on fractional integration and cointegration, Energy Econ. 46 (2014) 328–333.
[13] M. Monge, J. Infante, A fractional ARIMA (ARFIMA) model in the analysis of historical crude oil prices, Energy Res. Lett. 4 (1) (2023).
[14] Q. Lu, S. Sun, H. Duan, Analysis and forecasting of crude oil prices based on the variable selection-LSTM integrated model, Energy Inform. 4 (2021) 47, https://doi.org/10.1186/s42162-021-00166-4. Suppl. 2.
[15] Z. Bragoudakis, S. Degiannakis, G. Filis, Oil and pump prices: testing their asymmetric relationship in a robust way, Energy Econ. 88 (2020), https://doi.org/10.1016/j.eneco.2020.104755.

[16] B.F. Blair, R.C. Campbell, P.A. Mixon, Price pass-through in US gasoline markets, Energy Econ. 65 (2017) 42–49, https://doi.org/10.1016/j.eneco.2017.04.011.
[17] M. Monge, L.A. Gil-Alana, F.P. de Gracia, Crude oil price behaviour before and after military conflicts and geopolitical events, Energy 120 (2017) 79–91.
[18] L.A. Gil-Alana, R. Gupta, O.E. Olubusoye, O.S. Yaya, Time series analysis of persistence in crude oil price volatility across bull and bear regimes, Energy 109 (2016) 29–37, https://doi.org/10.1016/j.energy.2016.04.082.
[19] A. Tetteh, Q. Xu, Forecasting refined petroleum products prices in Ghana, Br. J. Econ. Manag. Trade 8 (2015) 94–107, https://doi.org/10.9734/BJEMT/2015/17766.
[20] L. Kilian, Explaining fluctuations in gasoline prices: a joint model of the global crude oil market and the U.S. Retail gasoline market, Energy J. 31 (2) (2010) 87–112. http://www.jstor.org/stable/41323283.
[21] R.K. Kaufmann, Ch. Laskowski, Causes for an asymmetrical relation between the price of crude oil and refined petroleum products, Energy Policy 33 (12) (2005) 1587–1596.
[22] Z. Anwer, A. Khan, M.A. Naeem, A.K. Tiwari, Modelling systemic risk of energy and non-energy commodity markets during the COVID-19 pandemic, Ann. Oper. Res. 345 (2) (2025) 1193–1227.
[23] A.K. Mishra, A.K. Mishra, K. Anand, S. Ranjan, Risk spillovers between crude oil and agricultural commodities in India: insights from two major global disruptions, Eur. Rev. Agric. Econ. (2025) jbaf011.
[24] A. Rao, B. Lucey, S. Kumar, Temporal dynamics of geopolitical risk: an empirical study on energy commodity interest-adjusted spreads, Energy Econ. 141 (2025) 108066.
[25] X. Liu, C. Xu, Y. Hong, H. Xu, Dynamic multiscale relationships between COVID-19 sentiment and extreme crude oil returns: evidence from wavelet coherence analysis, Emerg. Mark. Finance Trade 60 (11) (2024) 2533–2548.
[26] NBER (2022). Business cycle dating. National Bureau of Economic Research, July 16, 2022.
[27] IMF (2021). World Economic Outlook Update – July 2021. International Monetary Fund.
[28] WHO (2023). Statement on the fifteenth meeting of the IHR Emergency Committee regarding the COVID-19 pandemic. May 5, 2023.
[29] L.A. Gil-Alana, M. Monge, Crude oil prices and COVID-19: persistence of the shock, Energy Res. Lett. 1 (1) (2020).
[30] M. Monge, L.A. Gil-Alana, F.P. de Gracia, U.S. shale oil production and WTI prices behavior, Energy 141 (2017) 12–19, https://doi.org/10.1016/j.energy.2017.09.055.
[31] M. Monge, L.A. Gil-Alana, Spatial crude oil production divergence and crude oil price behaviour in the United States, Energy 232 (2021) 121034.
[32] M. Monge, E. Cristobal, L.A. Gil-Alana, A. Lazcano, Oil extraction and crude oil price behavior in the United States: a fractional integration and cointegration analysis, Energy Sources B: Econ. Plan. Policy 17 (1) (2022) 2149900.
[33] M. Monge, A. Lazcano, Crude oil price behavior and the adoption of sustainable measures in companies before and after COVID-19, J. Infraestructure Policy Dev. 8 (11) (2024) 7406, https://doi.org/10.24294/jipd.v8i11.7406.
[34] G.E.P. Box, G.M. Jenkins, 1970: Time Series analysis, Forecasting and Control, Holden-Day, San Francisco, 1970.
[35] C.R. Nelson, C.R. Plosser, Trends and random walks in macroeconomic time series: some evidence and implications, J. Monet. Econ. 10 (2) (1982) 139–162.
[36] D. Dickey, W. Fuller, Distribution of the estimators for Autoregressive Time-Series with a unit root, J. Am. Stat. Assoc. 74 (1979) 427–431.
[37] D.A. Dickey, W.A. Fuller, Likelihood ratio statistics for autoregressive time series with a unit root, Econometrica 49 (1981) 10571072.
[38] P.C.B. Phillips, Time series regression with a unit root, Econom.: J. Econom. Soc. (1987) 277–301.
[39] P. Phillips, P. Perron, Testing for a unit root in time-series regression, Biometrika 75 (1988) 335–346.
[40] D. Kwiatkowski, P.C. Phillips, P. Schmidt, Y. Shin, Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time-series have a unit root? J. Econ. 54 (1992) 159–178.
[41] R.K. Adenstedt, On large-sample estimation for the mean of a stationary random sequence, Ann. Stat. (1974) 1095–1107.
[42] C.W. Granger, Long memory relationships and the aggregation of dynamic models, J. Econ. 14 (2) (1980) 227–238.
[43] C.W. Granger, Some properties of time series data and their use in econometric model specification, J. Econ. 16 (1) (1981) 121–130.
[44] C.W. Granger, R. Joyeux, An introduction to long-memory time series models and fractional differencing, J. Time Ser. Anal. 1 (1) (1980) 15–29.
[45] J.R. Hosking, Modeling persistence in hydrological time series using fractional differencing, Water Resour. Res. 20 (12) (1981) 1898–1908.
[46] F.X. Diebold, G.D. Rudebush, On the power of Dickey-Fuller tests against fractional alternatives, Econ. Lett. 35 (1991) 155–160.
[47] U. Hassler, J. Wolters, On the power of unit root tests against fractional alternatives, Econ. Lett. 45 (1) (1994) 1–5.
[48] D. Lee, P. Schmidt, On the power of the KPSS test of stationarity against fractionally integrated alternatives, J. Econ. 73 (1) (1996) 285–302.
[49] J. Geweke, S. Porter-Hudak, The estimation and application of long memory time series models, J. Time Ser. Anal. 4 (4) (1983) 221–238.
[50] P.C.B. Phillips, Discrete Fourier transforms Fractional Processes, Department of Economics, University of Auckland, 1999.
[51] P.C.B. Phillips, Unit root log periodogram regression, J. Econ. 138 (1) (2007) 104–124.
[52] P.M. Robinson, Efficient tests of nonstationary hypotheses, J. Am. Stat. Assoc. 89 (1994) 1420–1437.
[53] P.M. Robinson, Gaussian semi-parametric estimation of long-range dependence, Ann. Stat. 23 (1995) 1630–1661.
[54] F. Sowell, Maximum likelihood estimation of stationary univariate fractionally integrated time series models, J. Econ. 53 (1–3) (1992) 165–188.
[55] P.M. Robinson, Log periodogram regression of time series with long range dependence, Ann. Stat. 23 (3) (1995) 1048–1072.
[56] H. Akaike, Maximum likelihood identification of Gaussian autoregressivemoving average models, Biometrika 60 (2) (1973) 255–265.
[57] H. Akaike, A Bayesian extension of the minimum AIC procedure of autoregressive model fitting, Biometrika 66 (2) (1979) 237–242.
[58] J. Breitung, B. Candelon, Testing for short and long-run causality: a frequency domain approach, J. Econ. 132 (2) (2006) 363–378.
[59] Ç. Ciner, Eurocurrency interest rate linkages: a frequency domain analysis, Int. Rev. Econ. Finance 20 (2011) 498–505.
[60] M. Kırca, Ş. Canbay, K. Pirali, Is the relationship between oil-gas prices index and economic growth in Turkey permanent? Resour. Policy 69 (2020) 101838.
[61] H. Tastan, Testing for spectral Granger causality, Stata J. 15 (4) (2015) 1157–1166.
[62] J. Geweke, Measurement of linear dependence and feedback between multiple time series, J. Am. Stat. Assoc. (1982) 304–313.
[63] S. Johansen, M.O. Nielsen, Likelihood inference for a fractionally cointegrated vector autoregressive model, Econometrica 80 (2012) 2667–2732.
[64] J. Beran, R. Bhansali, D. Ocker, On unified model selection for stationary and nonstationary short- and long-memory autoregressive processes, Biometrica 85 (4) (1998) 921–934.
[65] Reuters, Factbox: which U.S. refineries have shut since the global pandemic, and why? Energy (2022). June 17, 2022Available here, https://www.reuters.com/business/energy/which-us-refineries-have-shut-since-global-pandemic-why-2022-06-17/.
[66] S. Johansen, A representation theory for a class of vector autoregressive models for fractional processes, Econ. Theory 24 (2008) 651–676.
[67] S. Johansen, Likelihood-based Inference in Cointegrated Vector Autoregressive Models, Oxford University Press, 1996.
[68] Ch. Ding, T. Fitzgibbon, H. Govindahari, B. Roth, The Near-Term Outlook for refining: Where to know? McKinsey & Company. Oil & Gas Practice, 2023.
[69] J.C. Vides, J. Feria, A.A. Golpe, J.M. Martin-Alvarez, How do supply or demand shocks affect the US oil market? Financ. Innov. 10 (16) (2024) https://doi.org/10.1186/s40854-023-00561-8.