ORIGINAL RESEARCH



Revisiting the relationship between spot and futures markets: evidence from commodity markets and NARDL framework

Hachmi Ben Ameur¹ · Zied Ftiti² · Waël Louhichi³

Accepted: 16 June 2021 / Published online: 27 July 2021 © The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

This study aims to investigate the relationship between the spot and futures commodity markets. Considering the complexity of the relationship, we use a nonlinear autoregressive distributed lag (NARDL) framework that considers the asymmetry and nonlinearity in both the long and short run. Based on the daily returns of six commodity indices reaggregated on three commodity types, our study reaches some interesting findings. Our analysis highlights a bidirectional relationship between both markets over the short and long run, with a greater lead for the futures market. This result confirms the future market's dominant contribution to price discovery in commodities. Changes in commodity prices appear first in the futures market, as informed investors and speculators prefer trading on this market that is characterized by low costs and a high-leverage effect. Then, the information is transmitted from the futures to the spot market through arbitrageurs' activity, which explains the nonlinearity of the relationship. These results are helpful to scholars, investors and policymakers.

Keywords Commodity markets \cdot NARDL \cdot Spot market \cdot Futures market \cdot Lead-lag relationship

JEL Classification $C58 \cdot G1 \cdot Q02$

1 Introduction

Since 2002, the commodity market has exhibited fundamental changes, as explained by financialization. Investment flow into commodity markets has grown substantially, since institutional as well as specialized investors have been interested in this market (Domanski et al. 2007). Consequently, various financial instruments and derivatives have been

Zied Ftiti zied.ftiti@edcparis.edu

¹ INSEEC Grande Ecole, INSEEC U, Paris, France

² EDC Paris Business School, Paris, France

³ ESSCA School of Management, Paris, France

developed, leading to more attention from investors, traders and hedgers. Additionally, commodity market mechanisms have evolved substantially since 2002, unlike in conventional markets, based on their physical features (storable and non-storable commodity) as well as their importance in worldwide economic activity.

In the two decades of commodity markets' financialization, various periods of instability, high volatility, extreme movements and seasonal patterns have been observed. These stylized facts revive the issue of price discovery. Financial instruments (derivatives, futures contracts, etc.) are known to have a positive effect on price discovery and the efficiency of resource allocation (Chan 1992; Schwarz et al. 1994; Henriksen et al. 2019). However, they might be sources of anomalies. Indeed, futures contracts are sources of various forwardlooking decisions of economic agents: producers might define their supply strategy based on the price of futures contracts and physical traders, investors might define their asset allocation strategy based on the trend of futures prices, etc. Therefore, the spot–futures relationship is still a topical issue, particularly for commodity markets that this study aims to revisit.

From a theoretical point of view, various mechanisms might highlight the spot-futures relationship. Based on market efficiency theory, the Law of One Price in futures markets stipulates that the spot and futures prices might be related based on spot-futures parity, which ends any trade-off opportunities. However, the commodity market is imperfect, characterized by market frictions such as transaction costs and news reactions. Moreover, the theory of storage, pioneered by Kaldor (1983), is susceptible to supporting the spot-futures relationship. This theory stipulates that in equilibrium, the futures prices are equal to the spot prices plus the cost of carrying. Specifically, the theory of storage concurs that the basis and futures spread are connected to risk premium, convenience yield and the cost of storage. The basis defines whether the commodity market is in a contango or backwardation situation, including the shortage or surplus of the physical commodity in the market. As the market approaches maturity, the basis reduces until it equals zero and spot and futures prices converge.

Both theories discussed above highlight the connection between spot and futures prices in the long term. However, the commodity market has some specifications and stylized facts, supporting theoretically and practically a divergence between both prices in the short term, such as seasonal dynamics of consumption, thin trading, lags in information transmission, capital market microstructure effects, etc. This potential short-run deviation between both prices supports a nonlinear connection that might be observed through a lead–lag relationship between spot and futures prices. This nonlinear relationship might be displayed, in the case of commodity prices, through a demand shock due to seasonal consumption. Indeed, if the consumption of a commodity increases, explained by the seasonal effect, the inventory level decreases. Consequently, the convenience yield will increase and influence the price. Consequently, the seasonal effect inducing a demand shock leads to the futures price being highly volatile during the life of the futures contract. During the life of the futures contract, particularly at the beginning of the contract, a lead–lag relationship might be present between both prices, with a magnitude decrease in the function of its maturity. At maturity (in the long run), the futures prices converge with the spot prices.

From a literature review, many studies have investigated the issue of spot-futures price connections, particularly for commodity markets, but they are inconclusive. Moreover, the literature is orientated to precious metals such as gold. Some studies show that futures markets lead spot markets and play an important role in price discovery (e.g., Aslan et al. 2018; Bopp et al. 1987; Brooks et al. 2001; Kawaller et al. 1987; Stoll et al. 1990; Talbi et al. 2020). Some others reach a contradictory conclusion, showing that spot prices lead to futures prices (e.g.,

173

Pradhan et al. 2020; Srinivasan 2012). Some studies support an unclear relationship between both prices (e.g., Turnovsky 1983; Pindyck 2001; Bohl 2020; Wang et al. 2017), while some others investigate the potential asymmetric relationship between spot and futures prices through causality analysis and inconclusive findings (e.g., Arouri et al. 2013; Chang et al. 2015; Dash et al. 2010; Hammoudeh et al. 2004; Jena et al.2019; Joseph et al. 2014; Mayer et al. 2017).

The literature discussed above is inconclusive considering the results and the nature of the commodity. Our study revisits the literature on the spot–futures relationship for the commodity market to explore its specific characteristics compared to the conventional market, with the aim to complete it at four levels. First, we assessed the role of spot/futures prices in price discovery by examining their dynamics of interaction. Second, this relationship is assessed in both long and short terms to find some insights into the theoretical background discussed above. Third, we deliver new insights on the lead–lag relationship by considering various types of commodities and delving into precious metals. Four, we investigated the spot-future relationship based on an empirical framework with several advantages compared to those used in previous studies.

In this study, we assess the spot and futures relationship for three types of commodities: metal markets (aluminum, gold and copper), energy markets (Brent and natural gas) and agriculture markets (wheat). Considering the stylized fact of the commodity markets (seasonal dynamics of consumption, thin trading, lags in information transmission and capital market microstructure effects) and the potential complexity of their co-dynamics, we propose a nonlinear autoregressive distributed lag (NARDL) framework. This approach has several advantages over methods used in the previous literature (GARCH family models, cointegration, Granger Causality, Wavelet, etc.). First, it allows us to assess of the dynamics of the relationship between spot and futures prices for the long and short run and to test the theoretical background discussed above, contrary to other methods (GARCH family models, Granger Causality, etc.). Second, it considers the potential asymmetry and nonlinearity in the spot-futures relationship. Third, it avoids loss of information by allowing a mix between stationary and nonstationary series, contrary to the cointegration approach. Fourth, the nonlinear ARDL approach avoids spurious regression treatments (unit root tests, cointegration testing, etc.) that may be capricious based on some specific decisions such as the choice of deterministic part, autoregressive lags and so on (Dimitriadis et al. 2020).

Our study has some noteworthy findings. We found a bidirectional relationship between both markets over the short and long run. However, leadership is more pronounced in the futures market, supporting the conclusion of price discovery in commodity markets. This study provides insights on the transmission mechanisms in commodity markets, as we confirm that changes in commodity prices appear first in the futures market, explained by the preference of investors and speculators to trade in this market behind the low costs and the high-leverage effect. Thus, the information is transmitted from the futures to the spot market through arbitrage activity, which explains the nonlinear relationship.

The remainder of our paper is organized as follows. Section 2 presents the econometric methodology. Section 3 discusses the data and the empirical results and Sect. 4 presents the conclusion.

2 Econometric methodology

In this study, we use the NARDL approach introduced by Shin et al. (2014) to examine the relationship between the futures and spot markets for several commodity markets, especially the metal markets (aluminum, gold and copper), energy markets (Brent and natural gas) and agriculture markets (wheat). The use of the NARDL approach allows us to consider the complexity of the interaction between spot and futures returns, especially the non-linearity and asymmetry over the short and long run.

Pesaran et al. (1999) and Pesaran et al. (2001) are the first to propose a linear ARDL framework presented in the following form:

$$\Delta J_t^{Sp} = \alpha_0 + \sum_{i=1}^{P} \alpha_1 \Delta J_{t-i}^{Sp} + \sum_{i=1}^{q} \alpha_2 \Delta J_{t-i}^{Ft} + \theta_1 J_{t-1}^{Sp} + \theta_2 J_{t-1}^{Ft} + \varepsilon_t \tag{1}$$

$$\Delta J_t^{Ft} = \alpha_0 + \sum_{i=1}^P \alpha_1 \Delta J_{t-i}^{Ft} + \sum_{i=1}^q \alpha_2 \Delta J_{t-i}^{Sp} + \theta_1 J_{t-1}^{Ft} + \theta_2 J_{t-1}^{Sp} + \varepsilon_t$$
(2)

where J_t is the log price series for the aluminum, gold, copper, Brent, natural gas and wheat markets. Ft and Sp are the indicators for the futures and spot markets, respectively. α_1 is the autoregressive parameter. θ_1 and θ_2 represent the long-run dynamic relationships. p and q are the optimal lag lengths to be used for the different terms of the dependent and independent variables, respectively. ε_t is an *iid* process with zero mean and finite variance.

To implement the nonlinear ARDL framework, the positive and negative partial sum decompositions detecting the asymmetric effects both in the short and long run, are integrated into the linear ARDL Eqs. (1) and (2). Following Shin et al. (2014), the nonlinear cointegrating regression has the following form for the spot market:

$$\Delta J_{t}^{Sp} = \alpha_{0} + \theta_{1} J_{t-1}^{Sp} + \gamma^{+} J_{t-1}^{Ft,+} + \gamma^{-} J_{t-1}^{Ft,-} + \sum_{i=1}^{p} \alpha_{1} \Delta J_{t-i}^{Sp} + \sum_{i=1}^{q} (\beta_{i}^{+} \Delta J_{t-i}^{Ft,+} + \beta_{i}^{-} \Delta J_{t-i}^{Ft,-}) + \varepsilon_{t}$$
(3)

where $J_t^{Sp,+}$ and $J_t^{Sp,-}$ are the partial sums of positive and negative changes in J_t^{Sp} , defined as follows:

$$\begin{cases} J_{t}^{Sp,+} = \sum_{i=1}^{t} \left(\Delta J_{t}^{Sp,+} \right) = \sum_{j=1}^{t} \max \left(\Delta J_{t}^{Sp}, 0 \right) \\ J_{t}^{Sp,-} = \sum_{i=1}^{t} \left(\Delta J_{t}^{Sp,-} \right) = \sum_{j=1}^{t} \min \left(\Delta J_{t}^{Sp}, 0 \right) \end{cases}$$

For the futures markets, the nonlinear cointegrating regression has the following form:

$$\Delta J_t^{Ft} = \alpha_0 + \sum_{i=1}^p \alpha_1 \Delta J_{t-i}^{Ft} + \sum_{i=1}^q (\beta_i^+ \Delta J_{t-i}^{Sp,+} + \beta_i^- \Delta J_{t-i}^{Sp,-}) + \theta_1 J_{t-1}^{Ft} + \gamma^+ J_{t-1}^{Sp,+} + \gamma^- J_{t-1}^{Sp,-} + \varepsilon_t$$
(4)

For futures markets, the partial sums of positive and negative changes in J_t^{Ft} are defined as follows:

$$\begin{cases} J_t^{Ft,+} = \sum_{i=1}^t \left(\Delta J_t^{Ft,+} \right) = \sum_{j=1}^t \max\left(\Delta J_t^{Ft}, 0 \right) \\ J_t^{Ft,-} = \sum_{i=1}^t \left(\Delta J_t^{Ft,-} \right) = \sum_{j=1}^t \min\left(\Delta J_t^{Ft}, 0 \right) \end{cases}$$

🖄 Springer

The parameters γ^+ and γ^- include the long-run asymmetry. The positive and negative long-run coefficients are measured as follows: $L^+ = \frac{-\gamma^+}{\theta_1}$ and $L^- = \frac{-\gamma^-}{\theta_1}$. The parameters β_i^+ and β_i^- are associated with the short-run asymmetry.

To identify the best model that estimates the relationship between the spot and futures markets, we use the Wald test to detect the long- and/or short-run symmetries. For long-run symmetry, the null hypothesis H_0 , defined as $\gamma^+ = \gamma^-$, is tested against the alternative hypothesis H_1 of long-run asymmetry. For the short run, the null hypothesis H_0 defined as $\beta_i^+ = \beta_i^-$, is tested against the alternative hypothesis H_1 of short-run asymmetry. The Wald test allows the selection of the optimal model between the following four specifications:

2.1 Case 1

 H_0 is rejected for short-run symmetry and cannot be rejected for long-run symmetry. In this case, the cointegrating NARDL model with short-run asymmetry can be expressed as follows:

$$\Delta J_{t}^{Sp} = \alpha_{0} + \sum_{i=1}^{p} \alpha_{1} \Delta J_{t-i}^{Sp} + \sum_{i=1}^{q} (\beta_{i}^{+} \Delta J_{t-i}^{Ft,+} + \beta_{i}^{-} \Delta J_{t-i}^{Ft,-}) + \theta_{1} J_{t-1}^{Sp} + \theta_{2} J_{t-1}^{Ft} + + \varepsilon_{t}$$

$$\Delta J_{t}^{Ft} = \alpha_{0} + \sum_{i=1}^{p} \alpha_{1} \Delta J_{t-i}^{Ft} + \sum_{i=1}^{q} (\beta_{i}^{+} \Delta J_{t-i}^{Sp,+} + \beta_{i}^{-} \Delta J_{t-i}^{Sp,-}) + \theta_{1} J_{t-1}^{Ft} + \theta_{2} J_{t-1}^{Sp} + \varepsilon_{t}$$
(6)

2.2 Case 2

 H_0 is rejected for the long-run symmetry and cannot be rejected for short-run symmetry. The cointegrating NARDL model with long-run asymmetry can be expressed as follows:

$$\Delta J_t^{Sp} = \alpha_0 + \sum_{i=1}^p \alpha_1 \Delta J_{t-i}^{Sp} + \sum_{i=1}^q \alpha_2 \Delta J_{t-i}^{Ft} + \theta_1 J_{t-1}^{Sp} + \gamma^+ J_{t-1}^{Ft,+} + \gamma^- J_{t-1}^{Ft,-} + \varepsilon_t \quad (7)$$

$$\Delta J_t^{Ft} = \alpha_0 + \sum_{i=1}^p \alpha_1 \Delta J_{t-i}^{Ft} + \sum_{i=1}^q \alpha_2 \Delta J_{t-i}^{Sp} + \theta_1 J_{t-1}^{Ft} + \gamma^+ J_{t-1}^{Sp,+} + \gamma^- J_{t-1}^{Sp,-} + \varepsilon_t \quad (8)$$

2.3 Case 3

 H_0 is rejected for both long- and short-run symmetry. The cointegrating NARDL model with long- and short-run asymmetry is defined in Eqs. (3) and (4).

2.4 Case 4

 H_0 cannot be rejected for both long- and short-run symmetry. The suitable model is the cointegrating linear ARDL model defined in Eqs. (1) and (2).

Finally, the existence of cointegration estimated by the previous models does not necessarily mean that the estimated coefficients are stable. To test the stability of the estimated coefficients, we use the cumulative sum (CUSUM) and cumulative sum of squares (CUSUMQ) tests proposed by Brown et al. (1975). The CUSUM and CUSUMQ statistics are updated recursively and plotted against the breakpoints. The test simply

identifies parameter instability if the CUSUM and CUSUMQ exceed the area between the two critical bounds of 5% significance.

3 Data and empirical results

3.1 Data and preliminary results

In this study, we investigate the interaction between the spot and futures markets for the metal, energy and agriculture markets. We obtained the daily price time series data from the Bloomberg database and for the period from January 2002 to November 2020. The use of daily frequency constitutes a contribution to the literature, as existing studies typically use the monthly frequency. This frequency might provide more information on the dynamics of the relationship between spot and futures markets, especially in the evolution of storage. The descriptive statistics for all variables are presented in Table 1.

Descriptive statistics are calculated based on the log-returns series. The mean returns are positive for all the indexes excepting the spot Brent index. The mean of the futures indexes is higher than the mean of the spot indexes for the gold, aluminum and Brent markets. The volatility of the futures indexes is higher than the volatility of the spot indexes for the gold, aluminum, copper and wheat markets. The skewness was often negative and statistically significant during this period for gold, aluminum, Brent futures and spot indexes and positive for the natural gas and wheat markets. For the copper market, the skewness was negative for the spot index and negative for the futures index. Kurtosis of all indexes suggests that distribution tails are higher than those of a normal distribution are. Normality was statistically rejected for all series.

Table 2 presents the results of the unit root tests. We used two different tests for the unit root to assess the integration order of the series. The null hypothesis for the augmented Dickey-Fuller (ADF) test is that the data are nonstationary, where the null hypothesis for the KPSS test is that the data are stationary. The two tests show that at log levels, both futures and spot log price series are nonstationary in all cases, indicating the presence of unit root. For the log difference, the null hypothesis for the ADF test was rejected for all the series. The results are confirmed by the KPSS test, which is in favor of stationarity for all log difference series. Finally, we note that for the ARDL specification, it does not matter whether the regressors are stationary processes I (0) or nonstationary processes I (1).

Table 3 presents the results of the corresponding best suitable model that estimates the relationship between the spot and futures markets for the metal, energy and agriculture markets. We note that the nature of the relationship between spot and futures markets is specific for each asset. Short-run asymmetry characterizes the relationship between the spot and futures markets since this hypothesis is confirmed for all estimations except for the model specifications of the aluminum and wheat futures prices and copper spot prices. Additionally, for natural gas, the interaction between the spot and futures returns is characterized by a short-run asymmetry in both directions. Finally, the nonlinearity in either the long or short run is present for all the estimations except for the model specifications of the NARDL framework, as they show that the use of a linear symmetric model leads to misspecification of the relationship between spot and futures returns.

	Gold		Aluminum		Brent		Copper		Gas		Wheat	
	Futures	Spot	Futures	Spot	Futures	Spot	Futures	Spot	Futures	Spot	Futures	Spot
Mean	0.00045	0.00044	0.00018	0.00017	0.00044	-0.00009	0.00048	0.00048	0.00052	0.00110	0.00034	0.00034
Med	0.00018	0.00058	0.00000	0.00000	0.00042	0.00000	0.00000	0.0000	0.00000	0.00000	0.00000	0.00000
Max	0.09008	0.10788	0.09753	0.06604	0.21019	0.32528	0.12349	0.12441	0.38314	1.01565	0.09193	0.14315
Min	-0.09354	-0.09074	-0.10775	-0.07924	-0.24404	-1.99086	-0.11050	-0.09840	-0.18045	-0.44530	-0.09492	-0.09982
Std	0.01110	0.01091	0.01383	0.01343	0.02265	0.04641	0.01711	0.01654	0.03292	0.04877	0.01978	0.01950
Skew	-0.22800	-0.25477	-0.15076	-0.10197	-0.10970	-25.44100	-0.11003	0.01601	1.09149	4.15016	0.27711	0.29296
Kurt	8.48803	8.99977	7.03263	5.54349	13.36809	1035.43100	7.71873	7.55832	12.22166	78.81827	4.88025	5.90809
J-B Prob,	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
This table 1 maximum (eports the dea Max), minimu	scriptive stati: um (Min), stau	stics of spot a ndard deviation	nd futures ret on (Std), skew	urns for the c ness (Skew),	ommodity mark Kurtosis (Kurt)	cets of our sa and Jarque–1	mple. For eac 3era (J-B)	sh series, we r	eport the mea	n (Mean), me	dian (Med),

 Table 1 Descriptive statistics for the return series

	Gold		Aluminum		Brent		Copper		Gas		Wheat	
	Futures	Spot	Futures	Spot	Futures	Spot	Futures	Spot	Futures	Spot	Futures	Spot
ADF												
Log levels	2.271	2.319	0.396	0.364	0.299	0.284	1.262	1.413	-0.715	-0.890	0.396	0.039
Log difference KPSS	-70,823***	$-71,324^{***}$	$-75,120^{***}$	-72,977***	-72,600***	-38,756***	-75,041 ^{***}	$-75,130^{***}$	-73,203***	$-21,529^{***}$	-70,928 ^{***}	-71,3365***
Log levels	7.176	7.175	0.832	2.171	2.171	1.693	3.723	3.725	5.338	5.225	2.386	1.607
Log difference	0,2763	0,2805	0,0989	0,0894	0,2142	$0,3779^{*}$	0,3371	$0,3540^{*}$	0,1084	0,0774	0,0770	0,0853
The ADF test c and 10% level (critical values 0.3470).* der	: 1% level (3.4314), 5% l nce at the 10°	level (-2.861 % level, ** de	9) and 10% 1 protes signific	evel (-2.567)	0). The KPSS % level and *	Asymptotic ** denotes si	critical value gnificance at t	s: 1% level ((he 1% level	0.7390), 5% 1	evel (0.4630)

 Table 2
 Results of unit root tests

Dependent variable	Long run	Short run	Corresponding best-fit model
Aluminum spot	5.385**	12.461***	NARDL, long- and short-run asymmetries
	(0.020)	(0.000)	
Aluminum futures	7.963***	0.850	NARDL, long-run asymmetry
	(0.005)	(0.493)	
Gold spot	7.650***	7.681***	NARDL, long- and short-run asymmetries
	(0.006)	(0.000)	
Gold futures	1.866	4.386***	NARDL, short-run asymmetry
	(0.172)	(0.000)	
Brent spot	4.453**	5.890***	NARDL, long- and short-run asymmetries
	(0.035)	(0.000)	
Brent futures	0.498	55.529***	NARDL, short-run asymmetry
	(0.481)	(0.000)	
Copper spot	3.419	1.782	ARDL
	(0.065)	(0.076)	
Copper futures	4.929**	2.789***	NARDL, long- and short-run asymmetries
	(0.027)	(0.007)	
Gas spot	0.006	3.847***	NARDL, short-run asymmetry
	(0.938)	(0.000)	
Gas futures	1.222	3.229***	NARDL, short-run asymmetry
	(0.269)	(0.001)	
Wheat spot	0.230	11.401***	NARDL, short-run asymmetry
	(0.631)	(0.000)	
Wheat futures	0.289	0.291	ARDL
	(0.591)	(0.747)	

 Table 3
 Wald test results for long- and short-run symmetries

The corresponding best-fit model is determined based on the estimation of Eqs. (1), (2), (3) and (4). For long-run symmetry, the null hypothesis H0 defined as $\gamma^+ = \gamma^-$, is tested against the alternative hypothesis H1: long-run asymmetry. For the short run, the null hypothesis H0 defined as $\beta_i^+ = \beta_i^-$, is tested against the alternative hypothesis H1: short-run asymmetry.

P-values are in brackets. ***, ** and * indicate the significance at 1%, 5% and 10% levels, respectively

3.2 Lead–lag relationship between the spot and futures commodity markets

Table 4 presents the estimation results for the specification of the best regression models, as defined in Table 3. In these models, spot returns of our sample of commodity indexes (aluminum, gold, copper, Brent, natural gas and wheat) are a function of spot market returns and past information from the futures market, which allow us to test the predictive power of futures prices in explaining the spot returns. If the coefficient of the past information from the futures market is significant, the futures market contributes to the spot market price discovery and speculation in futures markets plays an important role in price formation. From Table 4, we can see that the Breusch–Godfrey test for serial correlation and the autoregressive conditional heteroscedasticity (ARCH) test with a lag of eight indicate that the selected models are correctly specified at a level of 5%. Furthermore, Table 4 indicates that all the models exhibit short-run asymmetry except for the model explaining the copper spot returns. The asymmetric long-run coefficients L^+ and L^- are positive and significant at the 1% level, for aluminum, gold and

Table 4 Resu	lts for the NAF	RDL estimat	tion for the spot	returns								
Variable	Aluminum spo	t	Gold spot		Brent spot		Copper spot		Gas spot		Wheat spot	
	NARDL-LR&	SR	NARDL-LR&:	SR	NARDL-LR&	SR	ARDL		NARDL-SR		NARDL-SR	
	Coef	Prob.*	Coef	Prob.*	Coef	Prob.*	Coef	Prob.*	Coef	Prob.*	Coef	Prob.*
$\alpha_{1,t-1}$	-0.362***	0.000	-0.256^{***}	0.000	-0.091***	0.000	-0.395^{***}	0.000	-0.135^{***}	0.000	-0.178^{***}	0.000
$\alpha_{1,t-2}$	-0.130^{***}	0.000	-0.192^{***}	0.000	-0.011	0.432	-0.244^{***}	0.000	-0.237^{***}	0.000	-0.024^{*}	0.084
$\alpha_{1,t-3}$	-0.037^{***}	0.007	-0.166^{***}	0.000	-0.081^{***}	0.000	-0.142^{***}	0.000	-0.147^{***}	0.000	-0.022^{**}	0.024
$\alpha_{1,t-4}$			-0.122^{***}	0.000	0.041^{***}	0.002	-0.096***	0.000	-0.138^{***}	0.000		
$\alpha_{1,t-5}$			-0.069***	0.000	-0.018	0.153	-0.059^{***}	0.000	-0.046^{***}	0.001		
$\alpha_{1,t-6}$			-0.027^{***}	0.002	-0.011	0.376	-0.041^{***}	0.003	-0.100^{***}	0.000		
$\alpha_{1,t-7}$					-0.026^{**}	0.013			-0.095^{***}	0.000		
$\alpha_{1,t-8}$									-0.043^{***}	0.000		
$\alpha_{2,t}$							0.868^{***}	0.000				
$\alpha_{2,t-1}$							0.379^{***}	0.000				
$\alpha_{2,t-2}$							0.237^{***}	0.000				
$\alpha_{2,t-3}$							0.130^{***}	0.000				
$\alpha_{2,t-4}$							0.080^{***}	0.000				
$\alpha_{2,t-5}$							0.059^{***}	0.000				
$\alpha_{2,t-6}$							0.036^{***}	0.008				
$\alpha_{2,t-7}$							0.014^{**}	0.011				
$\alpha_{2,t-8}$							0.015^{***}	0.00				
β_t^+	0.951***	0.000	0.962^{***}	0.000	1.013^{***}	0.000			0.469^{***}	0.000	0.764^{***}	0.000
β_{t-1}^+	0.334^{***}	0.000	0.205^{***}	0.000	0.228^{***}	0.000			0.567^{***}	0.000	0.273***	0.000
β_{t-2}^+	0.165^{***}	0.000	0.206^{***}	0.000	-0.130^{***}	0.0020			0.185^{***}	0.000	0.003	0.985
β_{t-3}^+	0.047***	0.001	0.147^{***}	0.000	0.244^{***}	0.0000			0.222^{***}	0.000	0.037**	0.016
β^+_{t-4}			0.097^{***}	0.000	-0.118*	0.086			0.285^{***}	0.000		
β_{t-5}^+			0.080^{***}	0.000	0.008	0.826			-0.005	0.903		
β_{r-6}^+					0.036	0.285			0.145^{***}	0.000		
β_{t-7}^+					0.063***	600.0			0.105**	0.010		

		5	Gold spot		Brent spot		Copper spot		Gas spot		Wheat spot	
	NARDL-LR6	č SR	NARDL-LR&	SR	NARDL-LR&	č SR	ARDL		NARDL-SR		NARDL-SR	
	Coef	Prob.*	Coef	Prob.*	Coef	Prob.*	Coef	Prob.*	Coef	Prob.*	Coef	Prob.*
g^+_{r-8}									0.070**	0.023		
8- °	0.936^{***}	0.000	0.997***	0.000	1.106^{***}	0.000			0.405^{***}	0.000	0.862^{***}	0.000
θ_{t-1}^-	0.400^{***}	0.000	0.223^{***}	0.000	-0.014	0.735			0.739^{***}	0.000	0.129^{***}	0.000
θ_{r-2}^-	0.123^{***}	0.000	0.135^{***}	0.000	0.087^{**}	0.034			0.152^{***}	0.000	0.083^{***}	0.000
β_{-1}^{-1}	0.039^{***}	0.005	0.164^{***}	0.000	0.031	0.442			0.153^{***}	0.000		
θ_{r-4}^-			0.112^{***}	0.000	0.125^{***}	0.000			0.131^{***}	0.001		
θ^{-}			0.030^{*}	0.061					0.233^{***}	0.000		
β_{-6}^{-6}			0.032^{**}	0.021					0.022	0.588		
θ_{t-7}^-			-0.015	0.083					0.117^{***}	0.004		
β_{-8}^{-1}			0.020	0.001					0.090	0.000		
9, 5	-0.263 * * *	0.000	-0.681^{***}	0.000	-0.007	0.169	-0.093^{***}	0.000	-0.130^{***}	0.000	-0.008***	0.000
\mathfrak{I}_2							0.093***	0.000	0.131^{***}	0.000	0.009***	0.000
+	0.265^{***}	0.000	0.682^{***}	0.000	0.003^{***}	0.000						
	0.263^{***}	0.000	0.678^{***}	0.000	0.014^{***}	0.000						
χ ⁰	1.895^{***}	0.000	3.850***	0.000	-0.002*	0.089	0.287***	0.000	-0.113^{***}	0.000	-0.045^{***}	0.000
+	$1,006^{***}$	0.000	$1,001^{***}$	0.000	$0,049^{***}$	0.000						
-1	0,999***	0.000	$0,996^{***}$	0.000	$1,942^{***}$	0.000						
AIC	-8.374		-8.186		-4.341		-7.102		-3.583		-6.206	
SIC	-8.355		-8.153		-4.310		-7.079		-3.545		-6.189	
3-G (8)	0.000***		0.000^{***}		0.001^{***}		0.000^{***}		0.000***		0.018^{**}	
ARCH(8)	0.000***		0.000^{***}		0.000***		0.000***		0.000 ***		0.000^{***}	
AKCH(8)	0.000*** 0.000***	in back mitor			0.000***		0.000***		0.000***		0.000***	

🖄 Springer

Brent spot markets indicating that both positive and negative variation in the futures markets allows a variation in the same direction in the long run. Similarly, asymmetric short-run coefficients β_i^+ and β_i^- for aluminum, gold and Brent spot markets are positive and significant indicating that last variations are induced in the same direction. For the copper, spot returns estimated using the symmetric ARDL model, the coefficients θ_2 and $\alpha_{2,i}$ associated with the long and short run, are positive and significant indicating that a change in the futures market is induced in the same direction for the copper spot market. For the natural gas and wheat spot markets, the asymmetry is relevant only for the short run. Overall, our results show that the futures market contributes to price discovery in the spot market. This implies that the process of price discovery is based on the transmission of new information from the futures prices to the spot prices through the transactions initiated by arbitrageurs who trade in different markets. This arbitrage activity and transaction costs characterizing the functions of commodity markets can be an explanation for the nonlinear relationship between the spot and futures returns.

Table 5 presents the estimation results of the best specifications (see Table 3) for the futures returns of the commodity markets of our sample (aluminum, gold, copper, Brent, natural gas and wheat). These regressions aim to test the predictive power of past information from the spot market and past futures returns in explaining futures returns. If the coefficient of the past information from the spot market is significant, we conclude that there is a bidirectional relationship between the spot and futures commodity markets and market fundamentals play an important role in price discovery. From Table 5, we note that the Breusch–Godfrey test for serial correlation and the ARCH test for conditional heteroscedasticity with a lag of eight indicate that the selected models are correctly specified at a level of 5%. The results reported in Table 5 justify the use of the NARDL model, which considers the complexity of the relationship between both markets characterized by nonlinearity and asymmetry in the short and long run. Interestingly, all the selected models exhibit short-run asymmetry except for the model explaining aluminum and wheat futures returns. By comparing the parameter values of regressions estimating the spot returns (Table 4) and regressions estimating futures returns (Table 5) for gold, Brent and gas, we note that, in the short run, the parameter values are greater when the spot return is the dependent variable. This conclusion indicates that there is bidirectional feedback between both markets but the futures markets have a greater impact on the spot markets than the contrary. Similarly, for the wheat market, by comparing the relationship based on the symmetric long run, the relationship is also stronger when considering the impact of the futures index on the spot index.

Overall, our results show that futures returns lead to spot returns, which confirms the dominant contribution of the futures market toward price discovery in commodity markets. Our results are in line with the strand of literature supporting bidirectional causality (such as Bhatia et al. 2018; Dash et al. 2010) and contradicts the strand supporting unidirectional causality from the futures market to the spot market and otherwise (e.g., Jena et al. 2018; Joseph et al. 2014; Srinivasan, 2012). However, our findings complete this literature from at least two aspects. First, we propose an investigation of a large set of commodity markets. Second, our empirical framework offers insights to explain the previous mixed results; although our results support a bidirectional relationship, we show that the relationship is more pronounced from the futures markets to spot markets.

Table 5 Resu	lts for the NAI	RDL estimat	tion for the futu	res returns								
Variable	Aluminum fut	ures	Gold futures		Brent futures		Copper future:	~	Gas futures		Wheat futures	
	NARDL-LR		NARDL-SR		NARDL-SR		NARDL-LR&	SR	NARDL-SR		ARDL	
	Coef	Prob.*	Coef	Prob.*	Coef	Prob.*	Coef	Prob.*	Coef	Prob.*	Coef	Prob.*
$\alpha_{1,t-1}$	-0.373^{***}	0.000	-0.224***	0.000	-0.159^{***}	0.000	-0.388***	0.000	-0.201^{***}	0.000	-0.170^{***}	0.000
$\alpha_{1,t-2}$	-0.159^{***}	0.000	-0.172^{***}	0.000	-0.016	0.251	-0.242^{***}	0.000	-0.046^{***}	0.007	-0.019^{**}	0.016
$\alpha_{1,t-3}$	-0.054^{***}	0.000	-0.150^{***}	0.000	0.022	0.112	-0.127^{***}	0.000	-0.040^{**}	0.018		
$lpha_{1,t-4}$			-0.108^{***}	0.000	-0.017	0.230	-0.089^{***}	0.000	-0.042^{**}	0.013		
$\alpha_{1,t-5}$			-0.069***	0.000	-0.005	0.707	-0.066***	0.000	-0.082^{***}	0.000		
$\alpha_{1,t-6}$			-0.039^{***}	0.005	0.033	0.017	-0.043^{***}	0.002	-0.034^{**}	0.038		
$lpha_{1,t-7}$					-0.026*	0.056	-0.015^{**}	0.012	-0.002	0.879		
$lpha_{1,t-8}$					0.023^{**}	0.044	-0.010	0.096	-0.050^{***}	0.001		
$\alpha_{2,t}$	0.980^{***}	0.000									0.844^{***}	0.000
$lpha_{2,t-1}$	0.364^{***}	0.000									0.150^{***}	0.000
$\alpha_{2,t-2}$	0.144***	0.000										
$\alpha_{2,t-3}$	0.045***	0.001										
$\alpha_{2,t-4}$												
$\alpha_{2,t-5}$												
$\alpha_{2,t-6}$												
$\alpha_{2,t-7}$												
β_{r}^{+}			0.884^{***}	0.000	0.639^{***}	0.000	0.931^{***}	0.000	0.258^{***}	0.000		
β_{r-1}^+			0.284^{***}	0.000	-0.079***	0.000	0.397***	0.000	0.030	0.160		
β_{r-2}^+			0.163^{***}	0.000	0.069***	0.000	0.264^{***}	0.000	0.079^{***}	0.000		
β_{r-3}^+			0.160^{***}	0.000	0.023*	0.070	0.139^{***}	0.000	0.027	0.202		
β_{r-4}^+			0.134^{***}	0.000	-0.024*	0.059	0.081^{***}	0.000	0.082^{***}	0.000		
β_{r-5}^+			0.071^{***}	0.001	0.060***	0.000	0.095***	0.000	0.016	0.442		
β_{r-6}^+			0.053***	0.002	-0.030^{***}	0.009	0.034^{**}	0.038	0.005	0.790		
β^+_{t-7}					-0.023***	0.004			0.072***	0.000		

183

Variable	Aluminum fut	ures	Gold futures		Brent tutures		Copper future	s	Gas futures		Wheat future	
	NARDL-LR		NARDL-SR		NARDL-SR		NARDL-LR8	cSR	NARDL-SR		ARDL	
	Coef	Prob.*	Coef	Prob.*	Coef	Prob.*	Coef	Prob.*	Coef	Prob.*	Coef	Prob.*
$\beta_{r=8}^+$									0.024	0.121		
β_t^{-}			0.877^{***}	0.000	0.290^{***}	0.000	0.956***	0.000	0.256***	0.000		
β_{t-1}^-			0.239^{***}	0.000	0.356^{***}	0.000	0.391***	0.000	0.116^{***}	0.000		
β_{t-2}^-			0.227 * * *	0.000	0.028	0.094	0.230^{***}	0.000	-0.000	0.982		
β_{t-3}^-			0.169^{***}	0.000	-0.078***	0.000	0.134^{***}	0.000	0.046^{**}	0.035		
β_{t-4}^-			0.113^{***}	0.000	0.072***	0.000	0.136^{***}	0.000	0.025	0.249		
β_{t-5}^-			0.096^{***}	0.000	-0.023	0.136	0.035*	0.080	0.034	0.111		
β_{t-6}^-			0.035^{**}	0.034	-00.00	0.563	0.064^{***}	0.000	0.063***	0.003		
β_{r-7}^-			0.015*	0.098	0.063^{***}	0.000			0.010	0.602		
β_{t-8}^-			-0.021^{***}	0.001	-0.065***	0.000			0.046^{***}	0.000		
θ_1	$-0,286^{***}$	0.000	-0.700***	0.000	-0.010^{***}	0.000	-0.092^{***}	0.000	-0.021^{**}	0.018	-0.005^{***}	0.001
θ_2			0.700***	0.000	0.011^{***}	0.000			0.015^{*}	0.078	0.004^{***}	0.004
γ^+	$0,275^{***}$	0.000					0.094***	0.000				
γ^{-}	$0,285^{***}$	0.000					0.089***	0.000				
α_0	1.998^{***}	0.000	0.002	0.109	-0.024***	0.000	0.374^{***}	0.000	0.018^{***}	0.000	0.024^{***}	0.001
L^+	$0,961^{***}$	0.000					$1,021^{***}$	0.000				
L^{-}	$0,996^{***}$	0.000					0,967***	0.000				
AIC	-8.327		-8.289		-5.450		-7.019		-4.122		-6.158	
SIC	-8.312		-8.256		-5.413		-6.985		-4.084		-6.149	
B-G (8)	0.000^{***}		0.000^{***}		0.000^{***}		0.004***		0.018^{**}		0.039^{**}	
ARCH(8)	0.000^{***}		0.000 * * *		0.000^{***}		0.000***		0.000^{***}		0.000^{***}	

 $\underline{\textcircled{O}}$ Springer





Fig. 1 The CUSUM test. Cusum denotes the CUSUM test results and 5% significance denotes the upper and lower range of significance based on 5%. The figures on the left and right show the CUSUM test results for the time series futures as explanatory variables and the CUSUM test results for the spot time series as explanatory variables, respectively



Fig.2 The CUSUMQ test. The cusum of squares denotes the CUSUMQ test results and 5% significance denotes the upper and lower range of significance based on 5%. The figures on the left and right show the CUSUMQ test results for the time series futures as explanatory variables and the CUSUMQ test results for the spot time series as explanatory variables, respectively

It seems that informed investors and speculators prefer trading in the futures market characterized by low costs and a high-leverage effect, which explains the lead of this market in price discovery. Subsequently, the information is transmitted to the spot market and in equilibrium, both markets are rich through arbitrage activity. These results show that speculation is an important determinant of commodity prices rather than market fundamentals.

Figure 1 reports the results of the CUSUM test for the stability of the relations between the different spot and futures commodity markets over the long and short run. According to Fig. 1, the estimated specifications are moderately stable for all regressions except the model specification for the copper spot returns. For the latter, the CUSUM exceeds the area between the two critical bounds of 5% significance. Similarly, the CUSUMQ test reported in Fig. 2 considers the variance, allows for the detection of random movements rather than systematic changes and is a perfect complement to the CUSUM test (Brown et al. 1975). The CUSUMQ test is in favor of the stability parameters of the model specifications of the wheat futures, spot returns and gas futures returns. The above results can be explained by the fact that our sample period includes several turmoil periods characterized by high volatility, such as the subprime crisis (2007–2009), the debt crisis (2010–2012) and the ongoing COVID-19 crisis. These turbulent periods that caused abrupt changes can explain the instability of the parameters revealed by the CUSUMQ test.

4 Conclusion

This study aims to investigate the relationship between spot and futures markets for a sample of commodity markets, especially the metal markets (aluminum, gold and copper), energy markets (Brent and natural gas) and agriculture markets (wheat). Considering the complexity of the interaction between both markets characterized by nonlinearity and asymmetry in the short and long run, we model this relationship on a NARDL framework. Our results highlight bidirectional feedback between both markets, but the lead of the futures market on the spot returns is greater than the contrary. This confirms the leadership of the futures commodity market in the process of price discovery. Changes in commodity prices appear first in futures markets and then the information is transmitted to spot markets through arbitrageurs' trades. This result shows that speculation is an important determinant of commodity prices rather than market fundamentals. The arbitrageurs' activity and transaction costs characterizing the functions of the commodity market explain the nonlinear relationship.

The results of our study are useful for several reasons. First, it informs regulators if the futures market plays an initial role in risk management or if it is dominated by speculative activity. Second, determining the leader market in which commodity prices change helps investors determine their investment and arbitrage strategy. Third, our results are helpful to scholars in the information efficiency field, as it helps them understand the process of price formation and transmission of information from one market to another.

We should note that in this study, we investigated the spot-futures relationship between various assets. It would be interesting for future research to draw some portfolio investment strategies that account for our findings using assets exhibiting only a short-run relationship and those exhibiting both short- and long-run relationships.

Funding None

Availability of data and material upon authors request

Code availability upon authors request

Declarations None

Conflict of interest None

References

- Arouri, M. E. H., Hammoudeh, S., Lahiani, A., & Nguyen, D. K. (2013). On the short-and long- run efficiency of energy and precious metal markets. *Energy Economics*, 40, 832–844.
- Aslan, S., Yozgatligil, C., & Iyigun, C. (2018). Temporal clustering of time series via threshold autoregressive models: Application to commodity prices. *Annals of Operations Research*, 260, 51–77. https://doi. org/10.1007/s10479-017-2659-0
- Bhatia, V., Das, D., Tiwari, A. K., Shahbaz, M., & Hasim, H. M. (2018). Do precious metal spot prices influence each other? Evidence from a nonparametric causality-in-quantiles approach. *Resources Policy*, 55, 244–252.
- Bohl, M. T., Siklos, P. L., Stefan, M., & Wellenreuther, C. (2020). Price discovery in agricultural commodity markets: Do speculators contribute?. *Journal of Commodity Markets*, 18, 100092.
- Bopp, A. E., & Sitzer, S. (1987). Are petroleum futures prices good predictors of cash value?. The Journal of Futures Markets (1986–1998), 19(4), 705.
- Brooks, C., Rew, A. G., & Ritson, S. (2001). A trading strategy based on the lead–lag relationship between the spot index and futures contract for the FTSE 100. *International Journal of Forecasting*, 17(1), 31–44.
- Brown, R. L., Durbin, J., & Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society: Series B (methodological)*, 37(2), 149–163.
- Chan, K. (1992). A further analysis of the lead–lag relationship between the cash market and stock index futures market. *Review of Financial Studies*, 5, 123–152.
- Chang, C. P., & Lee, C. C. (2015). Do oil spot and futures prices move together? *Energy Economics*, 50, 379–390.
- Dash, M., & Andrews, S. B. (2010). A study on market behaviour and price discovery in Indian commodity markets. Available at SSRN 1722770.
- Domanski, D., & Heath, A. (2007). Financial investors and commodity markets. BIS Quarterly Review, March.
- Dimitriadis, D., & Katrakilidis, C. (2020). An empirical analysis of the dynamic interactions among ethanol, crude oil and corn prices in the US market. Annals of Operations Research, 294, 47–57. https:// doi.org/10.1007/s10479-018-2832-0
- Hammoudeh, S., Dibooglu, S., & Aleisa, E. (2004). Relationships among US oil prices and oil industry equity indices. *International Review of Economics & Finance*, 13(4), 427–453.
- Henriksen, T. E. S., Pichler, A., Westgaard, S., & Frydenberg, S. (2019). Can commodities dominate stock and bond portfolios? *Annals of Operations Research*, 282, 155–177. https://doi.org/10.1007/ s10479-018-2996-7
- Jena, S. K., Tiwari, A. K., Hammoudeh, S., & Roubaud, D. (2019). Distributional predictability between commodity spot and futures: Evidence from nonparametric causality-in-quantiles tests. *Energy Economics*, 78, 615–628.
- Jena, S. K., Tiwari, A. K., & Roubaud, D. (2018). Comovements of gold futures markets and the spot market: A wavelet analysis. *Finance Research Letters*, 24, 19–24.
- Joseph, A., Sisodia, G., & Tiwari, A. K. (2014). A frequency-domain causality investigation between futures and spot prices of Indian commodity markets. *Economic Modelling*, 40, 250–258.
- Kaldor, N. (1983). The role of commodity prices in economic recovery. *Lloyds Bank Review*, 149, 21–34.
- Kawaller, I. G., Koch, P. D., & Koch, T. W. (1987). The temporal price relationship between S&P 500 futures and the S&P 500 index. *The Journal of Finance*, 42(5), 1309–1329.

- Mayer, H., Rathgeber, A., & Wanner, M. (2017). Financialization of metal markets: Does futures trading influence spot prices and volatility? *Resources Policy*, 53, 300–316.
- Pesaran, M. H., & Shin, Y. (1999). An autoregressive distributed lag modeling approach to cointegration analysis. S. Strom (Ed.), *Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium (11)*. Cambridge University Press, Cambridge (1999)
- Pesaran, M. H., Shin, Y., & Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics*, 16, 289–326.
- Pindyck, R. S. (2001). The dynamics of commodity spot and futures markets: A primer. The Energy Journal, 22(3), 1–30.
- Pradhan, R. P., Hall, J. H., & du Toit, E. (2020). The lead–lag relationship between spot and futures prices: Empirical evidence from the Indian commodity market. *Resources Policy*, 101934.
- Shin, Y., Yu, B., & Greenwood-Nimmo, M. (2014). Modelling asymmetric cointegration and dynamic multipliers in a nonlinear ARDL framework. *Festschrift in Honor of Peter Schmidt* (pp. 281–314). Springer.
- Schwarz, T. V., & Szakmary, A. C. (1994). Price discovery in petroleum markets: Arbitrage, cointegration and the time interval of analysis. *The Journal of Futures Markets* (1986–1998), 14(2), 147.
- Srinivasan, P. (2012). Price discovery and volatility spillovers in Indian spot–futures commodity market. *The IUP Journal of Behavioral Finance*, 9(1), 70–85.
- Stoll, H. R., & Whaley, R. E. (1990). The dynamics of stock index and stock index futures returns. Journal of Financial and Quantitative Analysis, 25(4), 441–468.
- Talbi, M., de Peretti, C., & Belkacem, L. (2020). Dynamics and causality in distribution between spot and futures precious metals: A copula approach. *Resources Policy*, 66, 101645.
- Turnovsky, S. J. (1983). The determination of spot and futures prices with storable commodities. *Econometrica: Journal of the Econometric Society*, 51(5), 1363–1387.
- Wang, D., Tu, J., Chang, X., & Li, S. (2017). The lead–lag relationship between the spot and futures markets in China. *Quantitative Finance*, 17(9), 1447–1456.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.